



Received: July 29, 2021
Accepted: September 13, 2021
Available online: September 25, 2021

Anorkhon Ahmedova

Independent researcher, Uzbekistan State World
Languages University
Tashkent, Uzbekistan
E-mail: anorxonimomaliyeva@mail.ru

STAGES OF CREATING PARALLEL CORPUS OF ENGLISH-UZBEK SIMILES

ANNOTATION

In recent years, corpus linguistics has been mentioned in the scientific literature as the main tool for the elaboration of dictionaries and grammar manuals, the creation of corpus analysis, the practical use of corpus, as well as the statistical study of linguistic analysis using corpus. Articles, textbooks, manuals on the creation of general and special electronic corpus have been published, but as the Uzbek language is a relatively new field of corpus linguistics, there is a need for research on some issues. The development of corpus linguistics and the increasing focus on statistical methods of processing linguistic materials have led to the development of a number of techniques related to the use of parallel or similar texts in different languages. This article is devoted to the stages of creating bilingual parallel corpus and explains the peculiarities of creating parallel corpus of texts. The creation of parallel corpus also has its own research stages. Statistical research in linguistics has a long history, especially since the advent of computers in the fifties of the twentieth century, research in this area has grown rapidly. The purpose of the study is to analyze the scientific views on the creation of the linguistic supply of the program of translation of similes on the basis of parallel texts (English-Uzbek, Uzbek-English), to study the linguistic basis and to study the lexical-semantic relationship of similes in bilingual vocabulary. This article deals with the creation of a parallel corpus of literary texts with English and Uzbek similes, analysis of lexical-semantic relations of similes, formation of a database of

Анорхон Аҳмедова

Ўзбекистон давлат жаҳон тиллари университети
мустақил изланувчиси
Тошкент, Ўзбекистон

ИНГЛИЗЧА-ЎЗБЕКЧА ЎХШАТИШЛАР ПАРАЛЛЕЛ КОРПУСИНИ ЯРАТИШ БОСҚИЧЛАРИ

АННОТАЦИЯ

Сўнгги йилларда корпус тилшунослиги илмий адабиётларда луғатлар ва грамматик қўлланмаларни тайёрлаш, корпус таҳлилларини яратиш, корпуслардан амалий фойдаланиш, шу билан бирга лингвистик таҳлилларни корпуслар ёрдамида статистик жиҳатдан ўрганишнинг асосий воситаси сифатида тилга олинмоқда. Умумий ва махсус электрон корпусларни яратишга бағишланган мақолалар, дарсликлар, ўқув қўлланмалари чоп этилган, бироқ ўзбек тили корпус лингвистикаси нисбатан янги соҳа бўлгани боис айрим масалалар устида илмий тадқиқот олиб бориш зарурати мавжуд. Ушбу мақола икки тилли параллел корпус яратиш босқичларига бағишланган ва унда матнларнинг параллел корпусини яратишнинг ўзига хос хусусиятлари тушунтирилади. Параллел корпуснинг яратилишининг ҳам ўз тадқиқот босқичлари бор. Тилшуносликда статистик тадқиқотлар узоқ тарихга эга, айниқса, XX асрнинг 50-йилларида компьютерлар пайдо бўлганидан буён ушбу соҳадаги тадқиқотлар тез суръатлар билан ўсди. Тадқиқотнинг мақсади параллел матнлар (инглизча – ўзбекча, ўзбекча – инглизча) асосида ўхшатишларнинг таржима қилиш дастури лингвистик таъминотини яратишга доир илмий қарашларни таҳлил қилиш, лингвистик асосларини ўрганиш ҳамда билингвал луғатдаги ўхшатишларнинг лексик-семантик муносабатларини тадқиқ қилишдан иборат. Ушбу мақолада инглизча ва ўзбекча ўхшатишлар қатнашган бадий матнларнинг

text segments based on translation alternatives of similes, scientific proof of linguistic models of English similes in translation and the effect of mentality is scientifically substantiated in the example of the translation of similes. A bilingual simile dictionary also plays an important role in creating a parallel corpus of similes. There is a great need for electronic dictionaries in the field of literary translation. The creation of bilingual and multilingual dictionaries, which embody the subtle nuances of the word, serves as a very important resource for translators, writers, poets, and language users.

Key words: corpus linguistics, parallel corpus, simile, role of similes in translation, English-Uzbek parallel texts.

параллел корпуси яратилиш, ўхшатишларнинг лексик-семантик муносабатлари таҳлил қилиш, ўхшатишларнинг таржима муқобиллари асосида матн сегментлари базаси шакллантириш, матнларни таржима қилишда инглизча ўхшатишларнинг лингвистик моделлари мувофиқлигини илмий далиллаш, таржимада эквивалентлик ва адекватлик масаласи, таржимага миллий маданият ва менталитетнинг таъсири ўхшатишлар таржимаси мисолида илмий жиҳатдан асослаб берилган. Шунингдек, ўхшатишлар параллел корпусини яратишда икки тилли ўхшатишлар луғати катта аҳамият касб этади. Бадиий таржима соҳасида ҳам электрон луғатларга эҳтиёж катта. Сўзнинг турли маъно нозикликларини ўзида мужассам этган икки тилли ва кўп тилли луғатларни тузиш, таржимон, ёзувчи, шоир ҳамда тилдан фойдаланувчилар учун жуда муҳим манбаа сифатида хизмат қилади.

Калит сўзлар: корпус лингвистикаси, параллел корпус, ўхшатиш, ўхшатишларнинг таржимадаги ўрни, инглизча-ўзбекча параллел матнлар.

INTRODUCTION

The first parallel corpus was “The Hansard” corpus (French-English), which consisted of more than fifty million words collected in the 1980s by working groups from Bell Communications Research and The IBM T.J. Watson Research Center. This corpus covered debates in the Canadian Parliament from the mid-1970s to 1988. The corpus also includes French-English legislative speeches, drafts, and oral presentations on the prepared legal field.

In 1992-93, the *European Corpus Initiative (ECI)* was developed. The *International Telecommunications Union CCITT Handbook* (13.5 million words) and the International Labor Organization (5 million words) jointly compiled a database for it. It covers 3 languages (French, English, Spanish). Multilingual corporations such as The MULTEXT and MLCC have also been created. It included bills of 10 million words in 9 languages of the European Community and negotiations in Parliament of 10 million words [Kenny D., 2001; 51-52].

Also today, MULTEXT-East language resources, a multilingual data set for the research of linguists, are designed to determine the morphosyntactic level of linguistic description. The MULTEXT-East database contains morphosyntactic classification, morphosyntactic lexicons, and annotated multilingual corporations based on The EAGLES. The parallel corpus in this corpus, George Orwell’s “1984” novel, consists of sentence-adapted and hand-crafted morphosyntactic descriptions and lemmas. The

Text Encoding Initiative Guidelines for Text Encoding, using the TEI P5 platform, resources are coded in XML, Bulgarian, Croatian, Czech, English, Estonian, Hungarian, Macedonian, Persian, Polish, Romanian, Russian, Serbian, Slovak, covers 16 languages such as Slovenian and Ukrainian. This data set is freely available for research purposes. However, in the corpus, the development history of MULTEXT-East sources, their coding and components are given, and some conclusions are given [Atwell E., 2008; 54].

MAIN PART

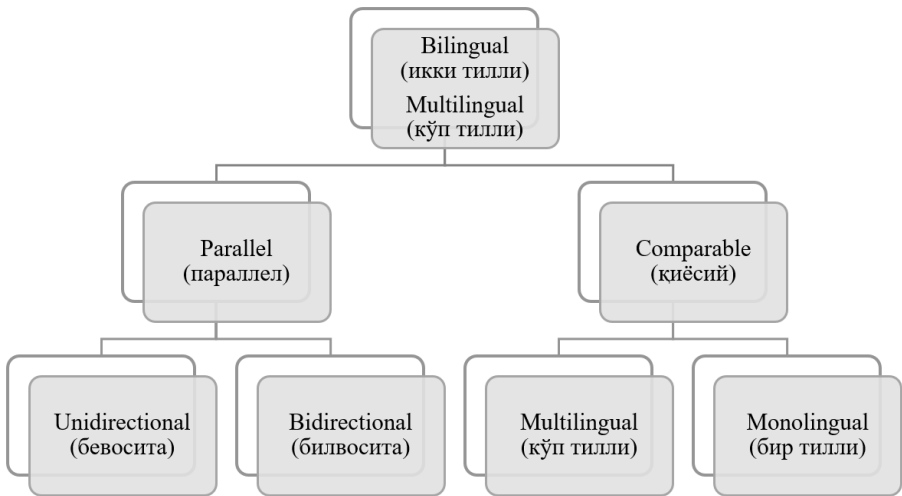
Russian scholars M.Zelko [Zelko M., 1991; 165] analysed Chinese-Russian parallel texts, Yu.N. Marchuk [Marchuk Yu., 1983; 232] encountered machine translation and the problems in it, Yu.A. Motorin [Motorin Yu.A., 1970] conducted research on automated translation technologies.

Although no scientific research has been conducted on parallel corpus in Uzbekistan, the site <http://uzbekcorpus.uz/> contains the text of laws and projects based on the concept of the family, led by researcher N. Abdurahmanova. This site contains a database of Uzbek-English, English-Uzbek and Uzbek-Russian, Russian-Uzbek parallel texts. This corpus has the ability to search for a single token (word) in the search engine that applies to each segment unit. The process of operation of this corpus is as follows: first, each word goes through the stage of morphological analyzer and its (word) basis is determined.

Once the word base is determined, the word is translated and the pairs of words with the *ID* value attached to the segments according to the translation alternative of the word are identified, and the text fragments in parallel corpus along the segment *ID* units attached to each word are displayed. In the initial stage, each word, phrase, stable compound, and sentence is divided into segment units. Many software has been developed for the implementation of this phase, using technologies such as WordFast, Dejabu, BOOTCAT, CAT in the process.

Kenny [Kenny D., 2001; 51-52] describes the parallel corpus as follows:

Table 1. Corpus classification in translation studies



2) by groups of semantic fields.

In a word search, words are taken as a linguistic benchmark and given to a search engine. As a linguistic standard, alternatives to all the similes associated with it are presented on a paired basis [Bakalo D., 2003; Bazylev V., 2007]. The search engine for word groups reflects the semantic relationship of any word in the alphabet, and the text alternatives of all the similes associated with the group in which it is written are sent to the screen (interface) [Vinson L., 2017; 320].

In order to collect the simulations, we paid special attention to whether the original text language of the works was Uzbek or English. The similes were collected in English and Uzbek from the original text language and translation language versions. In creating the parallel corpus of Uzbek-English, English-Uzbek similes, we used the following works of art translated from Uzbek into English and vice versa: *Alpomish*, *Navoi*, *Starry nights of Babur*, 1984, *Animal farm*, *Alchemist*, *A Farewell to Arms!*, *The Old Man and the Sea*, *The Da Vinci Code*, *Gulliver's Travels*, *Love of Life*, *White silence*, *The law of life*, *The story of Keesh*, *Martin Eden*, *White Fang*, *J. Gerxardt*, *An American Tragedy* and other fiction books with their translation.

During the stages of creating Uzbek-English parallel corpus, English-Uzbek similes, special attention was paid to the following:

1. Creating the content of translated works.
2. Identify translation alternatives and segment units.
3. Development of text creation technology between the author and the translator in the translation process.
4. Place similes in the original and translation in the database.
5. Provide metadata of bilingual texts.
6. Identify the query unit in the search engine (single-content and multi-content).

These methods are mainly based on statistics and mathematical linguistics (information theory, algorithm theory, etc.), which enable the user to obtain objective information about the composition, structure and occurrence of language units [Piatrovskiy R., 1999; 196].

Parallel corpora consist of translations or translated texts. In the process of compiling the parallel corpus, the presence of similes in the text was mainly taken into account. To create a parallel corpus, the texts are sorted based on the internal properties of the texts.

The study and analysis of parallel texts Tasks related to the creation of machine translation systems, automated and automated dictionaries, terminological databases require. The following tasks are performed when selecting a series of parallel texts:

- 1) identify specific goals and objectives for the study and design of linguistic-based automatic translation;
- 2) compile a sample from the total number of texts. Determining the homogeneity or diversity of total coverage at this stage and its limitation or infinity;
- 3) find the value of the allowable error in the calculations for the required size of the given sample;
- 4) auxiliary dictionaries determining the frequency of dictionaries.

In this section, we consider the main components of the analysis of such texts in relation to machine translation research.

To create a parallel corpus, we need to convert the data to a tabular format. That is, we can cite files in simple formats such as *Excel*, *TMX*, *XML*, *XLIFF*.

The requirements for a spreadsheet format are as follows: The spreadsheet includes language names. The information is in one column and its translation is given in the second column in the case where the sequence of words is matched in two languages. The spreadsheet makes the bilingual data customized in two columns. All other columns are left blank.

To create a corpus, special criteria are created depending on the purpose for which the corpus is created. In particular, texts must be decided whether to include a static or dynamic set and whole texts or text samples. In the corpus of authorship, size, theme, genre and style are also taken into account [Vlakhov S., Florin S., 1980; 342].

In general, the compilation of the case is based on the following requirements:

1. Contains auscultatory (natural process) information of the language.
2. Samples will consist of different dikurs data.

The source language and the translation language are alternated so that we can create a parallel corpus correctly. This means that words, phrases, sentences are adapted to the original language and the language of translation [Beldiyan V., 2009; 136; Bibler V., 1989; 43-47]. The reason for alternating the parallel corpus is that in the process of translation sentences can be shortened, phrases can be expressed in a single word, sentences can be dropped, two sentences can be combined into one sentence, sentences can be dropped or added, and sentence structure can change [Breus E., 2000; 208]. To compare texts, we need to match the text in the original language and the language of translation. In the process of alternation, the place of famous nouns, numbers, in quoted texts is equalized.

In a parallel corpus, the degree of alternation of the texts will be different. This is due to different genres of texts. For example, the compatibility of works of art with translation and scientific-technical or sectoral translations varies. That is, the translation of works of art is more freely translated than the translation of legal works. While scientific works are translated literally, works of art are translated with some modifications to express emotional feelings [Ataboev N., 2020; 189].

Corpus linguistics remains a topical theoretical and practical task in the field of machine translation. First, it is necessary to create effective dictionaries of input and output languages in order to translate vocabulary correctly. Such dictionaries should include a contextological dictionary that translates polysemantic words according to the context in which they are used in the introductory context. This dictionary is a set of context query algorithms for the presence of certain determinants, i.e., contextual properties that affect the translation of indefinite words. Marchuk dwells on this in his work [Abdurakhmonova N, Isroilov J., 2018;].

The statistical approach to machine translation facilitates manual and time-consuming tasks in traditional lexicography. Machine translation methods based on the first example served as an impetus for further steps in the field of obtaining linguistic

information from parallel texts. On the one hand, with the development of statistical machine translation systems, it was necessary to develop complex statistical models, reject the idea of absolute independence of algorithms from processed languages, develop morphological analyzers [Brown R., 1999; 20].

Naturally, all types of statistical research need to be interpreted qualitatively. However, they are the basis on which appropriate linguistic research and development procedures need to be built and can be established.

In the process of creating parallel corpus, the machine translation steps are initially used. In order for the texts to correspond to each other and to be adequate in the translation, the words adjacent to the main word are taken into account when selecting the one from several meanings of a word [Abdurakhmonova, N., 2018].

Automatic compilation of translated concord is one of the tasks to be solved by any statistical machine translation system, but to date it has become a separate area of a research.

Research on statistical processing of parallel texts is the most successful in the field of automatic dictionary compilation. In the mid-1990s, the automatic creation of translation dictionaries from parallel texts began to be considered as a separate task, and to date has created a new, promising approach in the form of setting up lexical translation analysis for unrelated texts [Fung P., 1998; 50].

Statistical methods of analyzing parallel texts have become part of information retrieval systems in different languages.

Translation of fiction is a process in which a translator is given greater opportunities than a scientific and technical translator. The main reason for this is that the literary translation text is original and to give aesthetic pleasure to the reader in the process of reading the translated text. In influencing emotions, the writer or poet uses a variety of art forms, exaggerated, artistically colored words. In translating such specific processes that occur in works of art, the translator uses a variety of translation transformations [Vasilyeva A., 2002; 171].

It is difficult to imagine the translation process without translation transformations, because the most important and highest goal of translation is to achieve adequacy. The task of the translator is to convey the content of the original text as fully as possible in the translated text on the basis of all the features of the language, style, genre, using various translation transformations effectively and appropriately.

Lexical transformations are used when there is a disproportion in the lexical meaning of words. The main reason for the use of lexical transformations often arises as a result of incompatibility of semantic units in two contact languages, because the word used in the original language can have one, two and many meanings, and in the translation language the lexical unit can contain only one meaning. Or conversely, a word in the original text and in the language of translation may have the same meaning. This opens the way for literal translation. Lexical transformations are addressed when the style and application of word meanings do not match the features of word formation [Vinogradov V., 1978; 58].

Grammatical transformation involves making changes in sentence structures in

accordance with the norms of the language of translation in the translation process.

Morphological transformations change almost nothing in the content of the text. The syntactic changes that are made have very little effect on the content of the original text.

Syntactic transformations involve changing the syntactic function of a word or phrase. Changing the syntactic function of words and phrases requires a grammatical transformation, that is, a replacement of syntactic constructions. In most cases, this involves converting the passive ratio to the self ratio and vice versa [Vavilova, L., 2004; 40].

Semantic transformation takes place on the basis of the various cause-and-effect relationships that exist between the elements that describe the situation.

It should be noted that in translation, in rare cases, translation transformations occur in pure form, which usually come in a mixed state. This indicates that translation transformations can be used as a whole system. At the same time, translation transformations consist of a step-by-step combination of complex changes that take place in the language of translation [Vereshchagin E.M., Kostomarov, V.G., 1990].

In our comparative analysis of simile of the two languages, we explored similarities and differences between Uzbek and English.

As a result of our research, a parallel subcorpus <http://uzbekcorpus.uz/> was created on the platform <http://uzbekcorpus.uz/>. Its content consists of texts in scientific and formal style. The fact that the formal and scientific style does not have a linguo-cultural and pragmalinguistic character applies to such norms as the use of words in their meaning, adaptation to the structure of the text, not to deviate from the requirements of the style. However, the fact that sensitivity in works of art has situations related to the human psyche makes it difficult to identify segment units in a parallel corpus. In this case, it is also a bit difficult to identify the similes in the search engine and find a second language alternative. However, it is possible to solve this problem by means of linguistic units, such as *like...* in English, *-dek* in Uzbek. But these tools are not always evaluated as means of simulation. The simulations were therefore placed in the database as a segment unit of translation. This makes it easier to search and use the case for the user interface of the case.

The process of accessing the parallel body of simulations:

1. Access via the link <http://uzbekcorpus.uz/>.
2. A parallel body is selected from the interface.
3. In the Style section, a parallel body of artistic texts is selected.
4. In the parallel corpus of similes 2 sections of the search are given:
 - 1) searching by group;
 - 2) searching by word.
5. Simile appears in the following view:

While translating a translator utilizes different types of stylistic devices. One of them is simile. The simile is one of the few studied stylistic tropes in linguistics. The term “simile” is differently described by various scholars. I.R. Galperin, D.U. Ashurova, Q.Musaev, G.Salomov and others conducted research in this area. In Uzbek simile has the following definition: if something or event is bound to a second thing or event, this comparison is called simile. There are also specific laws of using simile. If a creator wants to describe the beauty of a sweetheart, the poet uses certain elements specific of the *tashbeh*. For example, “*Mening yorim go’zallikda Shirin kabi*” will be translated “*My sweetheart’s in her beauty look like Shirin*”. The one selected for identification in verse is called “*mushabbah*” because ‘Shirin’ is the symbol of beauty in Uzbek literature. To describe it clearly, what is to be compared is that it is called “*mushabbahi bihi*” and “*sweetheart*” is an example. Then it is said that her face is likened to Shirin, meaning “*in beauty*” in literature is called “*vajhi shabah*”. In Uzbek the means of comparison *-dek*, *-day*, *kabi*, *misoli*, *yanglig’* are used and it’s called “*odati tashbeh*”. There are many similarities between English (simile) and Uzbek (o’xshatish) languages and their usage in poetry. In our research, we will try to find out problems related to simile.

In English this term is described as follows:

Simile is a figure of speech used in describing or explaining something. It points out a likeness between two different objects or ideas by using a connective word. This connective word is usually *like* or *as*. An example of a simile would be, ‘He is as bear today’.

Often simile becomes so compact that we drop the connecting word. Then, the simile becomes a metaphor. For example, the simile ‘She looked like an angel’. A better one would be this, from Hardy’s poem ‘Apostrophe to an Old Psalm Tune’,

‘...sweet as angels’ laughter’s’. That is, we are using the name of an animal to describe a man.

If the face of a well-described girl is ‘cherry’ or ‘roses’ in the English language, then the phrase ‘apple’ refers to the Uzbeks: ‘As red as a cherry (rose)’ - *olmaday qizil - red as apple*. ‘As no harmless as a dove’ - ‘*Musichaday beozor*’, in the English language the word ‘pigeon’ means a mild person but in Uzbek can transform it, ‘*turtle-dove-musicha*’ in the English language. The English symbol of ‘feather’ is ‘bird’ in Uzbek: ‘As light as a feather’ - ‘*Qushday yengil bo‘ling*’. When compared physical strong person with the English horse, Uzbeks are likened to elephants. In translation ‘*As strong as a horse*’ will be translated into Uzbek ‘*Filday baquvvat*. Such a discrepancy between the languages of expression can, undoubtedly, lead to frustration. The difficulties encountered by skilled translators are often solved through the functional harmony of the two languages of the linguistic means. Here is the song of Ophelia, who is crazy in ‘Hamlet’ tragedy:

Ophelia:

They bore him barefas’d on the bier;
Hey non nonny, nonny, hey nonny;
And in his grave rain’d many a tear, -
Fare you well, my dove!

On the basis of M.Loizinskiy’s translation, the tragedy is translated into the Uzbek language by Shaykhzoda the song as follows.

Ophelia (singing song):

Ochiq edi tobuti,
Olib ketishdi darhol,
Ko‘p yig‘lashdi odamlar,
Qo‘zichog‘im yaxshi qol!

My dove is in Uzbek means the bird, which is a kind of a bird. There is no emotional meaning in the word. For this reason, M.Shaikhzoda renounces my dove’s pigeon translation and replaces it with the image of the lamb. But in Uzbek, it does not mean that older people are lambs. At the same time, the translator has enough reason: the exception is the case, the appeals of Ophelia to the dead father, which further aggravates the tragedy and increases the effectiveness.

Metaphor is a poetic genius based on a similarity between objects: similar to the one that exists. While metaphor is always based on similarity, it differs from the simile by the following points: a) Simile and similarity are also used in their meanings with certain grammatical means. In metaphor, the likeness of the likeness of what is likened to is that it comes in a portable sense; b) When two or three components are involved in the simile interpretation, it can only be expanded as long as one component is involved in metaphor. For example: In ‘Captain Gastello’s’ novel by M.Sheikhzod:

Sher yurakli bu lochin
Qoqib qanot-qulochin,
Quzg‘unlardan asradi –
Elning xotin-xalochin.

Captain Gastello attacked the largest Fascists' arsenal and bombed it with this aircraft. Though our hero sacrifices himself, he will save thousands of his fellow citizens. In this, Gastello's braveness assimilated to the falcon.

A manifestation of literary imagery: a characteristic of the phenomenon, and the characteristic of the phenomenon (epithet) in the composition of the particular phrase, to come by his own character. For example: *golden valley*, *emerald spring*, *pure love*. Metaphorical epithet comes from similar and similar objects. Sometimes the base of simile, such as Tenor (*o'xshatish asosi*) and Vehicle (*o'xshatish vositasi*) are dropped. The basis of simile, that is, the symbol is hidden in the portable meaning of metaphorical expression. For example, in Hamid Olimjon's poem:

O'zing so'yla, shu qora ko'zlar,
Unutarmi shu uchrashishni?
Shuncha o'sib, *nashtar kipriklar*,
Ko'rganmikan shu go'zal tushni?

Each word in the language that is used is meant by the fact that the same alternative word in a different language often substitutes those meanings in part or somewhere, but often differs significantly. In particular, such an increase in the proportions that are combined in something else will increase. For example, English word **white**, Uzbek **oq** words are the symbol of lightness, clarity, and light are all things in all nations, its opposite is darkness, wickedness, and misery. But it does not mean that all the languages used in the "white" are exactly the same.

English phrase "White as a sheet" means "chayshabday", "qog'ozday oq". In the Uzbeks, paper is the symbol of whiteness was the basis of its natural characteristic of the notion of imagination. However, there is no phrasological pattern in Uzbek that is called "choyshabday oq" and we are called "dokaday oq". *White as milk – sutday oq* in the English, Uzbek languages, is a symbol of milk – **whiteness** in two languages. *White as snow – qorday oq*. Although *white as chalk* are closely intertwined in English and Russian, comparisons of "*bo'rday oppoq*" are not typical of Uzbek. *White as wool – junday oppoq* belongs only to the English language because in Uzbek the word *jun* – *white* does not serve to represent the notion of whiteness (not called "junday oppoq").

We explain a piece of Shakespeare's 'Romeo and Juliette'. After Romeo and Juliette were forced to marry, Juliette sent her sister-in-law to Romeo to determine the time and place of marriage. A mother, who is extremely tired of the path, begins to complain about her headache before telling Juliette the good news:

Lord, how my head aches!
What a head have I!

It beats as it would fall in twenty pieces.

Here the 'twenty' number is used to describe the pain of a hero's headache with the word 'piece'. In this case, Uzbeks use the word "chil – forty" (tajik) in Tajik. In order to clarify this or that idea, different nations use different means, which are norms for their own languages.

The Uzbek translator who understood well the original meaning translated the

language with full utilization:

Oh boshginam og‘rib ketdi!
Voy-voy boshginam,
Boshim go‘yo tars yorilib chilpora bo‘ldi.

Here the translator uses a great metaphor. That is, the crack is used to treat watermelons and melons. The word “*chilpora*” of the head with a wobble had increased the sensitivity.

There is no doubt that the following translation corresponds to the original meaning of both semantic and stylistic aspects. The simile *as* had been translated “go‘yo” so it was adequate translation.

The poet is Muhammad Ali, a translator who fulfilled his responsibilities as a poetical translator, and carried out a number of tasks in this area.

Muhammad Ali translated works of several Russian and other poets into Uzbek. He also featured the famous Scottish bard Robert Burns and the English poet George Gordon Byron, who was an outstanding representative of English poetry.

The nature of Scotland is in line with our mountain valley’s life, our sensual emotions, and, frankly, the song of the Mountains. For example: take a look at the poem “A Red, Red Rose”. Muhammad Ali translated as “Love”.

If we take “A Red, Red Rose”, it means “red-rose”. The poet used that word in the sense of ‘love’, but he also tried to make his poem a song of ‘love’ in English with the symbolic name ‘A Red, Red Rose’.

In the following verse used a striking simile:

O my luve is like a red, red rose
That’s newly sprung in June
O my luve is like the melodie²
That’s sweetly play’d in tune.

In Russian:

Любовь, как роза, роза красная,
Цветет в моем саду.
Любовь моя – как песенка,
С которой в путь иду.

Uzbek translation from Russian:

Sevgi atirgul yanglig‘,
Yashnar mening bog‘imda.
Sevgi men aytar qo‘shiq,
Yo‘lga chiqqan chog‘imda.

Gaybulla Salomov analyzed the translation in depth. In the first line of the poem, “O my luve is like a red, red rose”. It is possible to say literally the prose translation, “*Eh mening muhabbatim qip-qizil lolaga o‘xshaydi*”. Marshak described as a red, red rose – тюльпан. When the poet mentioned the redness of tulip, the idea became clearer. However, it is called “*как роза роза красная*” and in structural terms the Russian language is similar to English.

As a result, when the English version was translated into Russian, it remained

intact in the spirit of Byron. In Uzbek, it cannot be called a “gul” like a “qip-qizil gul”.

Therefore, Muhammad Ali lost his sense of elegance when it came to make it simple. That is, he wanted to put the word “qip-qizil gul” in the word “atirgul”. That inevitably led to the weakening of the purpose of the poem. Because in the reader’s concept, roses are not only red flowers or red blood, but it is also in different colors. In this respect, the translator would be more appropriate to use the word “tulip” in English. Thus, in this passage, there are two types of poetry artifacts that are emotional and beautiful. If the interpreter did not have enough experience, he would translate the word “twenty” into “yigirma”, which would result in orthography mistake. As a result, both hyperbole and simile ones disappeared.

CONCLUSION

In other words, the creation of aesthetic effect, as well as a complete picture of the artistic image, in some cases forced the translator may be to change the sentence construction and use different translation methods.

As a result of our research, a parallel subcorpus (<http://uzbekcorpus.uz/parallel>) was created on the platform <http://uzbekcorpus.uz/>. Its content is a parallel database of texts on scientific and official style, as well as similes from works of translation of fiction from English into Uzbek and vice versa. Translation alternatives in works of translation of fiction art can be identified and segment units of texts can be entered for translation memory. There are comparable corpora, which are closely related to parallel corpus, and consist of similar texts in terms of genre, subject, register, and so on. In a parallel corpus, multiple text translation options of the same text are used, while comparative corpora are a collection of the same but different texts according to the same subject, genre, register feature.

The presentation of similes as a bilingual dictionary on the basis of parallel text fragments serves as an object for lexicographic research. Linking translation theory and practice using corpus linguistics serves as a factor in solving complex situations that are difficult to understand or translate in the translation process. Parallel-based analysis is also important for enriching and developing the scope of lexicographic research. Gathering a database of texts into the translation memory using a parallel body and constantly updating it will serve as an object for future research and will have a positive impact on the quality and speed of translation. The provision of syntactic models of simulations serves as an important resource for automatic analysis in the form of free compounds and fixed compounds in the grammatical analysis phase of machine translation for text analysis in the field of computer linguistics.

REFERENCES

1. Abdurakhmonova, N. (2018). *Ontological model of Uzbek language (as example morphology)*. Computational Models in Language and Speech, Kazan.
2. Abdurakhmonova, N., Isroilov, J. (2018). Personal names spell-checking – a study related to Uzbek. *Iranian Journal of Social Sciences and Humanities Research*, 6, 1. Iran.
3. Abdurakhmonova, N. (2017). Modeling analytic forms of verb in Uzbek as stage of morphological analysis in machine translation. *Iranian Journal of Social Sciences and*

- Humanities Research*, 5, 4. Iran.
4. Abdurakhmonova, N.Z. (2020). *Computer Linguistics* (400 p.). Globe Edit.
 5. Ataboev, N.B. (2020). Functional features of the English corpus (in the example of COCA). *PhD thes. in phil.* (189 p.). Tashkent.
 6. Atwell, E. (2008). *Development of tag sets for part-of-speech tagging*. Corpus Linguistics: Mouton de Gruyter.
 7. Bakalo, D.I. (2003). Formation of readiness for translation of linguoculturological features of oral expression. *PhD thes. in phil.* (226 p.). Tomsk: Language University.
 8. Barkhudarov, L.S. (1969). *Levels of the language hierarchy and translation*. Moscow.
 9. Barkhudarov, L.S. (1973). *Language and translation* (240 p.). Moscow: International Relations.
 10. Bazylev, V.N. (2007). *General linguistics*. Moscow: Gardariki.
 11. Beldiyan, V.M. (2009). *Introduction to translation studies* (136 p.). Omsk.
 12. Bibler, V. (1989). Dialogue of cultures. *Problems of Philosophy*, 6, 43 – 47. Moscow.
 13. Breus, E.V. (2000). *Fundamentals of the theory and practice of translation from Russian into English* (2nd ed.) (208 p.). Moscow: Publishing house of URAO.
 14. Brown, R. (1999). *Adding linguistic knowledge to a lexical example-based translation system*. Proceedings of the Eighth International Conference on Theoretical and Methodological Issues in Machine Translation (p. 20). Chester, UK.
 15. Buranov, J. (1983). *Comparative typology of English and Turkic languages* (267 p.). Moscow: Higher School.
 16. Fung, P.A. (1998). *Statistical view on bilingual lexicon extraction: from parallel corpora to non-parallel corpora*. Third Conference of the Association for Machine Translation in the Americas (p. 50).
 17. Kenny, D. (2001). *Lexis and creativity in translation. A corpus-based study* (pp. 51-52). Manchester: St Jerome.
 18. Marchuk, Yu.N. (1976). *Computational lexicography* (183 p.). Moscow.
 19. Marchuk, Yu.N. (1983). *Problems of machine translation* (232 p.). Moscow: Science.
 20. Motorin, Yu.A., Marchuk, Yu.N. (1970). Implementation of automatic translation on modern serial computers of general purpose. *Problems of Radio Electronics*, 7, 20-39.
 21. Piotrovsky, R.G. (1999). *Linguistic automaton and its verbal and cognitive substantiation* (196 p.). Minsk: MGLU.
 22. Vasilyeva, A.A. (2002). Syntactic transformations in translation from Russian into Yakut: on the example of attributive constructions. *PhD thes. in phil.* (171 p.). Yakutsk.
 23. Vavilova, L.Ya. (2004). *Translation and translation studies. Introduction to the theory of intercultural communication* (p. 40). Krasnoyarsk: SibSTU.
 24. Vereshchagin, E.M., Kostomarov, V.G. (1990). *Language and culture: Linguistic and regional studies in teaching Russian as a foreign language* (4th ed.) (248 p.). Moscow.
 25. Vinogradov, V.S. (1978). *Lexical issues of translation of literary prose*. (58 p.) Moscow.
 26. Vinson L. (2017). *Simultaneous translation from Russian into English techniques, skills, manuals* (320 p.). Moscow: Ryu Valent.
 27. Vlahov, S., Florin, S. (1980). *Untranslatable in translation* (342 p.). Moscow: International Relations.
 28. Zelko, V.M. (1991). Problems of the development of linguistic support for the system of Chinese-Russian informational machine translation. *PhD thes. in phil.* (165 p.). Moscow: Institute of Linguistics of the Academy of Sciences of the USSR.