Neo Scientific Peer Reviewed Journal
Volume 10, May, 2023 ISSN (E): 2949-7752
www.neojournals.com

CRITERIA FOR CREATING PARALLEL TEXTS

Begimqulova Marhabo Ismatulla qizi
Termez State University

Abstract:

Looking at the history of parallel texts, although parallel texts is considered as a new branch of science within corpus linguistics, its history goes back to BC. In 196 BC, texts honoring King Ptolemy V were inscribed on stone by Egyptian temples in two languages (Greek and Egyptian). This is a perfect example of what we now call "parallel texts," that is, the translation of one text into two or more languages. This stone is known as the Rosetta stone. In 1822, Jean-Francois Champollion's study of the Rosetta Stone revealed that hieroglyphs were the key to deciphering the text, the study of which led to huge changes in science and put an end to many controversies and myths. Although the Rosetta Stone is relatively new to stone inscriptions, it is perfect compared to other stone monuments because of the completeness of the text and the fact that it is given in two languages in parallel. It was only in the eighties of the nineteenth century that parallel texts began to be used in automatic language processing systems. Several attempts to use them were tested in machine translation in the late 50s.

Keyword

INTRODUCTION

However, limited data storage and computer capacity, the storage and computing capabilities of computers of those years, and the difficulty of entering a large amount of textual data limited the possibilities of using the corpus at that time. In the late 1970s, the first automatic method of parallel automatic text alternation was developed by Martin Kay and Martin Roscheisen in 1987.

They proposed a number of methods to alternate texts of large size and different levels. That is, they developed units that translate each other. These units include paragraphs, sentences, words and phrases. The use of alternative parallel texts was proposed by Harris . The theory he proposed involves the construction of translation memories, dictionaries, and lists of bilingual terms. Theorized that cross-linguistics can serve as a research resource for computer-assisted instruction or comparative linguistics and translation studies.

The importance of multilingualism has increased in the fields of global markets and worldwide information sharing, the use of parallel text, the need to research parallel corpora and refine the results obtained from them. Parallel texts are sometimes called bitexts, or if the number of languages is more than two, multitext is called multilingual. Today, the number and scope of parallel texts is increasing, thanks to the improvement of electronic document storage and archiving tools, various companies and the World Wide Web(www) are using parallel corpora containing multilingual documents.

Many scientists have conducted scientific research on improving operations performed in parallel corpus. Among them, Martin Kay and Martin Roscheisen, Gale and Church,

This work is published under **CC BY-NC-ND 4.0**

Resnick, Brown, Lai and Mercer have expressed their opinions about the steps and ways of creating a parallel corpus. According to Key and Roscheisen, in order for sentences to match each other in translation, it is necessary to achieve the equivalence of words, that is, for lexical mapping (lexical mapping in English), it is necessary to match the level of alternation of texts in two languages.

According to them, it is not easy to achieve the alternation of text words in two languages, and this alternation may not be adequate, but by providing the meanings of the searched word in several texts in the parallel corpus, the research learning process is facilitated. Another problem in creating a parallel corpus is that one word in one language is represented by one word. , can be represented by two or more words in the second language. If it is expressed in a short sentence in the original text, it can be expressed in a long sentence in the translated text.

An example of a parallel corpus on the Sketch Engine platform A parallel corpus consists of the same text translated into one or more languages. Texts are alternated (corresponding segments, sentences are connected). A corpus allows you to search in one or both languages to search or compare translations. A parallel corpus is a collection of texts, each of which is a collection of translated versions of the original into one or more other languages. Parallel corpora differ according to their structure: 1) Simple parallel corpora involve only two languages: one of the corpora is an exact translation of the other. 2) In multilingual parallel corpora, texts are given together with their translations in several languages. For the correct use of parallel corpora, it is necessary to match the source text and its translation. This means identifying pairs or sets of sentences, phrases, and words in the original text and their translations in other languages.

Parallel text alternation is important because during the translation process, the translator can separate, combine, delete, insert, or change the order of sentences to create a natural translation in the target language. The degree of correspondence between parallel corpus texts varies by text type. For example, a literary text may give the translator more freedom than a legal one. When compiling parallel corpora, the texts of the corpus are selected according to specific criteria, depending on the purpose of its construction.

In particular, a decision must be made whether to include a static or dynamic set of texts, full texts or text samples. Author, volume, subject, genre and style requirements should be considered. Any type of corpora should meet the following requirements: 1) texts should contain naturally occurring language data; 2) it should contain data from various types of speech. CORTRAD translation parallel corpus, created in 2009, is a necessary guide for translators and pedagogues. In it, the texts are presented in Portuguese and English as a linguistic pair. COMPARA Bilateral Portuguese English Parallel Corpus was created in 2011. The Opus corpus (2012) is an open source parallel corpus. Among the parallel cases, the CORTRAD parallel case stands out for its multi-functionality. This corpus was created on the basis of the COMET project of the University of São Paulo, which has two innovative functions: 1) The possibility of comparing different variants of the same text. It consists of the original text, a modified version and a published translation form. 2)

This work is published under **CC BY-NC-ND 4.0**

Availability of a search engine for each searched text. This function is provided by sorting the texts according to their type. CORTRAD's search systemCORTRAD consists of 3 subcorporations: journalistic, scientific and artistic texts. CORTRAD uses DISPARA software and the system provides an easy to use interface.

This case is considered an open case. Created on the basis of the COMPARA Linguatec project. This corpus consists of Portuguese-English fiction texts. All texts are alternated in two languages at the sentence level. It can be used equally easily by experienced corpus researchers and the lay user. COMPARA Corpus Search Engine COMPARA Corpus has a user-inconvenient feature, where each word to be searched must be enclosed in quotation marks. For example, "help" is like "yourself". The next parallel corpus is the Opus corpus, which is broad in terms of domains and languages, and contains several subcorpora. This corpus has many subcorpora and the size of the database is large, and the searched texts are automatically parallelized in the interface. In addition to the subcorpora files, Opus Corpus also provides word matching and phrase tables, bilingual dictionaries, frequency counts, and Opus High-Level Web. - makes it possible to find resources on the site through the search form.

The results of the Opus corpus can be downloaded to any computer provider. Literary translation is a process in which the translator is given greater opportunities than the scientific-technical translator. The main reason for this is that the text of the literary translation is the original and to give the reader aesthetic pleasure in the process of reading the translated text. In affecting emotions, a writer or poet uses various art forms, exaggerated, colorful words. The translator uses various translation transformations to translate such specific processes encountered in works of art. 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 use various translation transformations productively and in the appropriate place, to give the content of the original text as fully as possible in the translated text based on all the features of the language, style, and genre in it.

Lexical transformations are used when there is a discrepancy in the lexical meaning of words. The main reason for the use of lexical transformations in most cases is the result of the inconsistency of the meaning of the word units in the two languages, because the word used in the original language can have one, two, and many meanings, and in the translated language, the meaning of the lexical unit can include a word with only one meaning. . Or, on the contrary, the word in the original text and in the translated language may have the same meaning. This opens the way for literal translation. Lexical transformations are applied when the style and use of word meanings and the features of forming word combinations do not match. Grammatical transformation involves making changes in sentence structures in accordance with the norms of the translation language during the translation process. Morphological transformations do not change almost anything in the content of the text. The syntactic changes made have very little effect on the content of the original text. Syntactic transformations mean changing the syntactic function of a word or

This work is published under **CC BY-NC-ND 4.0**

phrase. Changing the syntactic function of words and word units requires grammatical transformation, i.e. replacing syntactic constructions. In this case, in most cases, it involves converting the passive to the subject and vice versa. Semantic transformation is carried out on the basis of the various cause and effect relationships between the elements describing the situation.

It should be mentioned that translation transformations are rarely found in pure form in translation. they usually come in a mix. This indicates that translation transformations are used as a whole system. Meanwhile, translation transformations consist of the step-by-step combination of complex changes made in the target language.

Semantic transformations differ from other transformations in that they have a greater impact on content. A comparison of English and Uzbek languages shows that these phenomena in each language do not correspond to each other. By phonetic transformation we mean transcription and transliteration used in translation. Transcription and transliteration refers to the transcription or transliteration of famous nouns (names), geographical names, names of organizations, offices, institutions, and names of ships and hotels, that is, giving the sounds in the above-mentioned names with appropriate sounds or letters in the translated language. It is known that in English there are 44 sounds that can be transcribed and transliterated, 20 of which are vowels and 24 are consonants, and 26 letters for transliteration (of which 20 are consonants and 6 are vowels). As an example of this, we cite from parallel texts: Shaibani Khan gazed with pleasure at the young men who were beating each other in the race, flying like golden eagles, and said to himself, "All right, let Babur be proud of his victory in the fall, name his daughter Fakhrinisa, let him go to sleep and write his poem!"

(P. Kadyrov. Starry nights) Shaibani Khan watching his horseman like golden eagles enthusiastically and censoriously thought: "Babur has become proud. Well, let him feast his eyes upon the autumn victory, let him compose poems, let Fakhrinisso be Fakhrinisso!"(P. Kadirov. Starry nights Babur) separated from its composition. The mighty, big river also stood still without waves, as if it had fallen asleep. (M. Twain. Tom Saer's experiences) The mighty river laylike an ocean at rest. (M. Twain. T. Sawyer) This work was translated from English to Uzbek, and the translator used lexical transformation of the translation during the translation process.

That is, in English, the simile like an ocean at rest is literally translated as the ocean is resting (calm), and in the translation it is as if it has fallen asleep. He used the lexical transformation method of translation. We studied the similarities and differences between the Uzbek and English languages in a bilingual comparative analysis of similes.

References:

1. Груздев, Д.Ю. Электронный корпус текстов как инструмент переводчика. Вестник Московского университета. Сер. 22. Теория перевода. 2011. № 2
2. Захаров В.П. Корпусный менеджер как поисковая система. СПб., 2002. С.31.

This work is published under **CC BY-NC-ND 4.0**

3. Захаров В. П., БогдановаС. Ю.Корпусная лингвистика: учебник. 3-е изд., перераб. — СПб.: Изд-во С.-Петерб. ун-та, 2020. — 234 с.
4. Шевчук В.Н. Электронные ресурсы переводчика. М., 2009. С. 52.
5. Laviosa, S. Corpora and Translation: the Methods and Theories of Corpus Work in Translation. Manchester, 2000. P. 3.
6. РУЗИЕВ, Х. Б. (2018). Употребление вопросного и вопросительного предложения в устной речи в английском языке. Развитие и актуальные вопросы современной науки, (2), 69-73.
7. РУЗИЕВ, Х. Б. (2018). Основное значение билингвализма при изучении языка. Развитие и актуальные вопросы современной науки, (4), 101-104.
8. РЎЗИЕВ, Х. Б. (2018). Некоторые советы по использованию диалогов в устной речи. Наука среди нас, (6), 275-278.
9. Рузиев, Х. Б. (2018). Semantic analysis of the words of value and respect in English and Uzbek. In Человекознание (pp. 40-41).
10. Ruziev, K. B. (2020). Proverbs and corpus linguistics. Актуальные проблемы гуманитарных и естественных наук, (6), 64-67.

This work is published under **CC BY-NC-ND 4.0**