CORPUS APPROACH APPLICATION TO MILITARY INTERPRETATION SOLUTION

Mikhaylova A.

Nakhimov Black Sea Higher Naval School, Sevastopol

Abstract: the article considers the possibility of using corpora in the process of translating military texts in a foreign language; electronic parallel corpora can give possibility to translate difficult texts more quickly and provide samples of military translation by studying interpreting techniques and methods.

Keywords: corpus, parallel corpora, military interpretation, context, military text, corpus approach.

The corpus approach is considered to be a promising method of linguistic research aimed at the applied study of language and its functioning in real texts.

Interpretation is not only a text transformation; it is a complex process of transforming the phenomenon of a skill set, analyzing possible inaccuracies and errors by using existing sources of linguistic information. On the one hand, different dictionaries solve lexical and grammatical problems, but on the other hand, they have a limited scope and require a certain amount of time to find a lexical and grammatical unit [1, 2].

Thus, there is a problem of searching resources of linguistic information and approaches to solving translation problems that can reduce the translation time and ensure the quality of text interpretation. To solve this problem it is necessary to apply a corpus approach in translation problems decision, which is based on the use of electronic text corpora – huge arrays of natural texts in a foreign language, collected on a magnetic medium and properly ordered and marked for faster search of linguistic information.

The main role of corpora is to provide the most correct examples of the use of language units that reflect both the complexity and features of the natural

language. The main purpose of the corpus of texts is to show the functioning of linguistic units in the context of a large material and in their natural environment – contextual one.

The corpus is defined by scientists as a set of authentic texts selected as representative of a particular language or as a set of written or oral texts that are stored on a computer. A special feature of the corpus is a fundamental collection of texts available for qualitative and quantitative analysis of the translated material.

Today, Wikipedia or Word Net has been replaced with a new approach based on linguistic corpora. There are many bilingual or multilingual corpora, including OPUS (an open source parallel corpus) – a collection of translations on the Internet, the placement of corpora, and the addition of linguistic information.

OPUS is a growing collection of translated texts from the web. In the OPUS project we try to convert and align free online data, to add linguistic annotation, and to provide the community with a publicly available parallel corpus [3, 4]. OPUS is based on open source products and the corpus is also delivered as an open content package. Data files for all language pairs in different formats and with different kind of annotation (if available) can be downloaded. OPUS also provides pre-compiled word alignments and phrase tables, bilingual dictionaries, frequency counts, and these files can be found through the resources search form on the top-level website of OPUS. OpenSubtitles also includes intra-lingual sentence alignments between alternative subtitle uploads in the same language.

A parallel corpus consists of the same text translated into one or more languages. Texts are aligned (normal sentences are linked).

There are two types of parallel corpora in modern corpus linguistics:

1. Comparable (Multilingual) Corpora;

2. Translation Corpora.

In the corpus of English and Russian languages, as well as using the concordance system (for example, the Sketch Engine program), search of

requested word and get the answer in a new window in the form of fragments of sentences from different texts. The meaning of a given word is considered in the context of its use in the language, which gives more bases for confidence in its use.

Conclusion. Corpora can be used in the process of translating English military texts. OPUS an open source parallel corpus allows users to search bilingual and multilingual data in many languages, find concordances, collocations, word list and more. When searching for a word in the corpus, the program-concordancer or corpus-manager offers a choice of word variants in contexts in the original language with possible meanings in the translated language.

REFERENCES

1. Biber D., Conrad S., Reppen R. Corpus Linguistics: Investigating Language Structure and Use**.** Cambridge: Cambridge University Press. 1998.

2. Chul-Kyu Kim. Personal Pronouns in English and Korean Texts: A corpus-Based Study in Terms of Textual Interaction. Journal of Pragmatics, V. 41, Issue 10, 2009, Pp. 2086-2099 https://doi.org/10.1016/j.pragma.2009.03.004

3. Gilmore A., Millar N. The Language of Civil Engineering Research Articles: A Corpus-Based Approach. English for Specific Purposes. V. 51, 2018, Pp. 1-17 https://doi.org/10.1016/j.esp.2018.02.002.

4. OPUS. Open Parallel Corpus Available at: http://opus.nlpl.eu (accessed 04 April 2020).

5. OPUS OpenSubtitles Available at: http://opus.nlpl.eu/OpenSubtitles.php (accessed 04 April 2020).