

GIVING FORMAL RELATIONS BETWEEN WORDS IN A PARALLEL CORPUS (EXAMPLE OF UZBEK-GERMAN LANGUAGES)

Rakhimova Sh.S.

Termiz State University, Teacher of the Department of Romano-Germanic Languages

<https://doi.org/10.5281/zenodo.11356500>

Abstract. *This article is about the linguistic basis of giving the formal relations between words in the Uzbek-German parallel corpus, and it was found that linguistic corpora are widely used in explaining the practical use of the language and the specific features of the language. Also, in the article, the programs developed through computer technologies for corpus linguistics, their differences, and the opinion of a group of scientists on corpus linguistics are mentioned.*

Keywords: *characteristics of corpus linguistics, parallel corpus, corpus linguistics models, intercultural communication, corpus linguistics.*

Introduction. A linguistic corpus is a collection of texts written in one or more languages. These corpora are data sets that have been analyzed to study the grammatical and lexicological features of the language, the relationships between texts, the relationships between words, and the visible and invisible structural features of words or word units. These corpora are used in linguistics, translation, language learning, automated translation, data interpretation, data analysis, automad correction, word interpretation, word construction, automated word analysis, and more. Linguistic corpora are widely used to explain the practical use of language and its specific features.

Linguistic corpora are widely used for data analysis and strengthening of educational processes with the help of technological algorithms and programs. These corpora are important in the study of language structure, word relations, text types and many other issues in the field of linguistics, and are used in teaching and research.

Literature analysis. A parallel corpus, a collection of originals and their translations, can be used in many ways for the benefit of translation studies, machine translation, linguistics, computational linguistics, or simply the human translator. In computational linguistics, translation corpora have been used for machine translation, as well as for term extraction, word semantics, etc., since the early 1980s. As the first parallel texts, avalanche reports collected in German, French, and Italian languages in Switzerland, and weather information provided by Canadian media in English and French appeared in the late 1980s and early 1990s. One of the earliest electronic resources was Canadian Hansard, which was originally used to perform sentence matching (Gale & Church 1991), a task now a standard feature of applications such as translation memories. In addition, parallel corpora are used as databases for multilingual grammar induction, automatic lexicography, and many other tasks in cross-language information retrieval and language processing. The last goal of all the projects created in these years was to create a machine translation system [1, 1445].

Discussion. Based on the level of study of both disciplines in modern linguistics and the researches and scientific views of world scientists in these fields, the following conclusions were drawn regarding the sharp differences between computer linguistics and corpus linguistics according to the following characteristics:

	Corpus Linguistics	Computational Linguistics
Table of Contents and Overview	Data is collected in between, and texts may depend on other texts.	Often original texts (corpus) are created and they are prepared according to the purpose.
Data usage feature	The information from the texts was obtained statistically and prepared for use.	Operations on texts are used to perform transformations, modeling, translation, analysis, and other operations.
Analysis of practices and preliminary issues	It is used in the analysis of relations between texts, relations between words and visible features of words or word units.	Mathematical models are used for creating mathematical models on data, identifying variables, correcting data and other operations.
Studying primary texts and languages	Primary texts, multiple languages, and language relationships are studied.	Primary texts are studied using systems, algorithms and mathematical models, and these primary texts are used for scientific research and experiments on practices.

Parallel corpora are central to translation studies and contrastive linguistics. Many parallel corpora are accessible through easy-to-use concordancers, making the study of cross-linguistic phenomena much easier. Anarbayev Orzubek Rachmanovich gives the following opinion to the parallel corpus: "The parallel corpus is an important reality for the current era, when intercultural communication is widespread." Through parallel corpora, it will be possible to identify universals in different language environments and cultures, as well as specific mental characteristics of languages, realia and lacunar units. The corpus of parallel texts also serves for the development of automatic translation, ensures the development of computer lexicography [1, 1446].

Parallel corpora have an important reality for the current era of widespread intercultural communication. These corpora contain parallel texts of several languages. They have become an important means of language learning in translation, tercume, language learning and other fields.

In the current era of globalization and the development of information technology, parallel corpora serve as an important tool, as intercultural communication and experiences grow. They help in better preparation for practices in various fields, translation, interpretation of data and strengthening of educational processes.

Also as an important tool in parallel corpora, translation technologies, language learning, automated translation, data interpretation, data analysis, automated correction, word interpretation, word construction, automated word analysis and other fields is used.

The result. Linguistic bases of giving formal relations between words in Uzbek-German languages have the following features:

Phonology:

The sound system in Uzbek and German languages is unique to each language and is of great importance in its explanation and definition.

Relationships between words are explained using phonemes and their separate sounds and letters. Example: In Uzbek, the word "book" is expressed as [ki-tob], and in German it is expressed as "Buch". Here the sound "k" changes from [k] to [b].

Morphology:

The relationship between words in the Uzbek and German languages are determined by the morphological structure of the word, that is, morphemes and their additional elements. Types of words, acquisition of words, additional morphemes (prefix, suffix, infix) and their use, relations between words are important in explanation. Example: In Uzbek, the word "pen" is expressed as [qal-am], and in German it is expressed as "Stift". Here, the word "stay" is formed using the morphemes "kal" (qal) + "am" (-am).

Syntax:

The syntactic relations between the sentences, the acquisition of words and word units in the sentence, variables, adverbs, relative nouns, water, etc. are explained.

The structure of the sentences is of great importance in determining the relationships between words. Example: In the sentence "I read a book" in Uzbek, the word "kitobni" is used as a sign of the word "kitob" added to the word "men". In the German language, in the sentence "Ich lese das Buch" the word "das Buch" is added to the abbreviation "ich lese" and the syntactic structure of the sentence changes.

Semantics:

Semantic concepts and meanings are important in the analysis of relations between words, in determining the meaning between words and word units.

Inter-word relations serve to explain the change of meaning between sentences and the semantic analysis of these changes. Example: In Uzbek, the word "alma" means "fruit", and in German, the word "Apfel" means "apple". Here, the word "apple" has the same meaning in Uzbek and German languages.

Pragmatics:

The relationship between words is related to the practical meaning of the sentence and their practical use.

In communication, the influence and response relations between sentences are analyzed through pragmatic concepts.

Discourse analysis:

Inter-word relations, connections between sentences in the text are important in the detailed analysis of the text.

Logical and content analyzes between sentences are carried out thanks to discourse analysis. Example: In the sentence "I read a book. Then I turn the page" in the Uzbek language, "I read" in the first sentence and "I turn" in the second sentence express the relationship between words. In German, there is a similar relationship between words in the sentence "Ich lese das Buch. Dann blättere ich die Seiten um".

Conclusion. In conclusion, it should be mentioned that it is used by linguists to explain and analyze the formal relations between words in the Uzbek-German languages. They are used as important tools in translation technologies, language learning, automated translation, data

interpretation, data analysis, automated correction, word interpretation, word generation, automated word analysis, and other fields.

REFERENCES

1. Nazira Ganijon Qizi Sobirova. "KORPUS LINGVISTIKASI VA PARALLEL KORPUSLAR TAVSIFI" Academic research in educational sciences, vol. 3, no. 5, 2022, pp. 1442-1450.
2. Jurayeva M.B. "KORPUS LINGVISTIKASINING ASOSIY XUSUSIYATLARI" Экономика и социум, no. 5-2 (92), 2022, pp. 119-123.
3. Abduraxmonova, N. Z. "Linguistic support of the program for translating English texts into Uzbek (on the example of simple sentences): Doctor of Philosophy (PhD) dis. aftoref." (2018).
4. Karshiyeva, B. (2023). Solving complex communicative-knowledge tasks based on integrated bilingual education in English. *Interpretation and researches*, 1(18).
5. Juraev E. S. Foreign experience in conducting financial policies for small business development. *Economy and Society*. – 2018.
6. Razzakov S. J., Xolboev Z. X., Juraev E. S. Investigation of the Stress-Strain State of Single-Story Residential Buildings and an Experimental Theoretical Approach to Determining the Physicomechanical Characteristics of Wall Materials // *Solid State Technology*. – 2020. – T. 63. – №. 4. – С. 523-540.
7. Karshieva, B. F. (2023). A component of professional communicative competence in english is professional engineering knowledge. *Journal of Universal Science Research*, 1(12), 257-261.
8. Karshieva, B. F. (2023, May). **METHODOLOGICAL MODEL OF TEACHING TECHNICAL STUDENTS IN ENGLISH ON THE BASIS OF INTEGRATED BILINGUAL EDUCATION**. In *interdiscipline innovation and scientific research conference* (Vol. 1, No. 9, pp. 58-60).