IDENTIFICATION OF SEMANTIC
FEATURES OF PHRASAL VERBS IN A PARALLEL CORPUS

Otamurodova Salomat
Toshkent davlat agrar universiteti
Tillar kafedrasi oʻqituvchisi.
E-mail: soloahmedovna@gmail.com

Annotation: The scientific, technological and information revolutions of the 20th century, which covered all areas of human knowledge and practice, could not but affect linguistics. We are talking not only about new forms of recording, such as audio or video recording and electronic data storage, but about phenomena that fundamentally change our views and approaches to the functioning of language. In recent years, corpus linguistics has become widespread - a branch of applied linguistics whose goal is the creation and maintenance (using machine media and appropriate access tools) of large arrays of language data.

Keywords: corpus linguistics, parallel corpora, phrasal verbs, semantic structure, phraseological unit.

Аннотация: Научно-техническая и информационная революции XX века, охватившие все области человеческого знания и практики, не могли не затронуть языкознание. Речь идет не только о новых формах записи, таких как аудио- или видеозапись и электронное хранение данных, но и о явлениях, которые коренным образом меняют наши взгляды и подходы к функционированию языка. В последние годы широкое распространение получила корпусная лингвистика — раздел прикладной лингвистики, целью которого является создание и поддержание (с использованием машинных носителей и соответствующих средств доступа) больших массивов языковых данных.

Ключевые слова: корпусная лингвистика, параллельные корпуса, фразовые глаголы, семантическая структура, фразеологизм.

Introduction

Today, corpus linguistics is actively and successfully used in various fields, such as applied linguistics, philology, stylistics, methods of teaching foreign languages, theory and practice of translation, criminal linguistics, content analysis, discourse analysis, genre analysis.

The role of text linguistics1 is determined by the desire to deeply study the connections of language with various aspects of human activity realized through text. Research conducted both in our country and abroad makes it possible to carry out detailed developments of the essential characteristics of the general linguistics of the text. The question of the boundaries of text linguistics is debatable, namely: does it include pragmatics, functional semantics and syntax, the currently actively developing rhetoric, or does it intersect with them as it intersects with poetics, psycholinguistics and communication theory, entering as an integral part of the theory of text. When creating and developing electronic corpora of languages of the world, the consistent research of NLP and language technologies has become an urgent task [Abdurakhmonova N.2021:202].

Parallel corpora

The phrase "multilingual corpora" also refers to "parallel corpora," which are collections of materials that have been translated into many languages and are aligned across those languages. The primary applications of parallel corpora in research are in translation studies and contrastive linguistics. They also have a significant influence on the creation of translation resources for language learners and on translation training. Beginning in 2016, the Multilingual Student Translation project seeks to integrate learner corpus research with corpus-based translation studies by creating a new, sizable multilingual student translation corpus.

Parallel corpora are divided into monolingual, bilingual and multilingual. In monolingual corpora, dialects and language variants are contrasted. For example, such varieties of English as English as a native language and English as a foreign language remained outside the scope of scientific interest until the advent of new technologies, which made it possible to involve a significantly larger number of compared speech products in contrastive analysis [Abdurakhmonova, N. 2021,2022].

Bilingual and multilingual corpora can be divided into two main types:

1) corpora representing a set of original texts written in any source language, and texts-translations of these source texts into one or more other languages;

2) corpora that combine texts from the same thematic area, independently written in two or more languages; Such corpora help in working with terminology and are often used by translators.

These kinds of corpora are used for comparative language studies (lexicology, grammar, stylistics, translation studies, etc.) and for the development of efficient translation techniques, such as machine translation and the creation of bilingual and multilingual terminological dictionaries.

Electronic dictionaries are excellent resources for researching words and phraseology. Although they often don't offer the same contexts for individual search items as full-text corpora do, the information retrieved can be utilized for additional searches inside historical corpora. Long-term efforts that span a great deal of historical time are large-scale dictionaries, which attempt to cover the vocabulary's evolution in a language. The Oxford English Dictionary Online (OED Online) for English, Der digitale Grimm for German, and Svenska Akademiens wordbook for Swedish are a few examples of these kinds of undertakings.

Large-scale text collections have showed potential, especially for researchers studying literature, although smaller corpora can also be helpful study tools. For literary and socio-historical study, whole text collections like the Innsbruck Computer Archive of Machine-Readable English Texts and the Corpus of Middle English Prose and Verse are invaluable resources. Legal and social historians as well as historical linguists might benefit from electronic editions such as the An Electronic Text Edition of Depositions 1560-1760 (ETED), which include transcriptions of early official documents.

The current stage of research and development in the field of machine translation and natural language processing language requires the use of new methods of representing linguistic reality, reflecting the complex the nature of natural languages and, at the same time, combining the capabilities of already existing approaches. The formation of corpora of parallel texts makes it possible to use machine learning in addition to traditional methods based on heuristic rules compiled by a human expert [Abdurakhmonova, N. 2021,2022].

However, in order to be able to automatically extract rules from parallel tests, it is necessary develop fairly subtle methods for comparing linguistic objects, with the main basis there must be

their semantic characteristics. Until now, semantic alignment of parallel texts has been considered mainly from the point of view terms of constructing lexical-semantic models, in our case we are talking about the semantics of structural components, i.e. ways of configuring linguistic objects (in semiotic linguistics they received name of structural signs [1]). These studies are motivated by the fact that the applied methods for automatically aligning parallel texts word by word or sentence by sentence lead to the emergence of a large number of noises that can only be eliminated manually at the stage post-editing. Phrasal verbs for researchers in the field linguistics skills are attractive element. This linguistic unit is dedicated to quite a large number of works of various linguistic whist. The increased interest in these units explains is determined not only by their prevalence in the linguistic environment (12 thousand according to dictionary entries only), but an abnormal fullness in the field of polysemy, synonymy, antonymy and in other lexical-semantic ski and thematic groups. It should also be taken into account that, since the 1960s, the frequency of use phrasal verbs increase significantly in scientific literature, journalism and academic texts the style. Subsequently, according to the law of analogy, in this productive language model, more and more new ones are included elements. In particular, verbal derivatives, image- bathrooms from nouns or adjectives in terms of conversion, are one of the leading word forms vocal trends at the present stage of development language. Determining the essence and importance of phrasal verbs is one of the modern tasks of the modern English phraseology.

According to Shansky, phraseological units are linguistic units consisting of two or more accentuated components, repeating, of a speech nature, having content and structure in a fixed (i.e., constant) sense. Considering phraseology as a branch of linguistics, A.F. Alefirenko considers its main task to be the mechanisms for coordinating phraseological components (syntactic, lexical or morphological) with non-free units of words, and the phraseological whole is one of these components, will lead to unification [Abdurakhmonova, N. 2021,2022]. He speaks.

A distinctive feature of phraseological units as part of the entire lexical-semantic system of any language is the stability and integrity of their meanings, which allows them to be repeated in speech in the form of semantically indivisible units. Many phrases serve as the basis for new formations due to a revision of the meaning of the words that make up the phraseological unit (A.B. Kunin. 2005).

In her guide to phrasal verbs, Jane Povey devotes an entire chapter to the use of particles and their semantic meaning, takes into account grammatical structure and properties of the verb, but semantic relationships between the components of the phrasal verb there remain inconsistencies followed. We selected for our study phrasal verbs with particles showing direction of movement and related to compositional(compositional). The division of phrasal verbs into compositional and non-compositional (non compositional), in which the meaning of the verb and postposition completely or partially changes, and therefore deduce the meaning of the verb from the meaning of its components does not work out, is traditional.

Along with the three main verbs of state, the microsystem of broad-meaning verbs includes words with a complicated semantic structure, which is formed by adding a dynamic, active component to the stative semantics of *be* or *have*. Thus, the dynamic correlates of *have* are the verbs *take* and *give*, denoting the action of including an object in a sphere, the center of which is the subject intake, and in give a recipient (participant in the situation; perceiver) that is different

from the subject; in the latter case, the transition of an object into the sphere of the recipient may mean the exclusion of the object from the sphere of the subject [Abdurakhmonova, N. 2021,2022].

Both verbs, having become broad-meaning, base their semantics on a significantly expanded representation of the subject, under which actions are also included (*take a look, give a look)*. Although both of these analytical lexemes are denotatively close to each other, which is reflected in their common Uzbek translation of *look,* the broad-meaning verbs in their composition retain the semantic opposite - *give a look* characterizes a *look* as a way of transmitting information about the state of the subject, while *take a look* means the search for visual information by the subject.

Object of study

The object of the study are literally characteristics of English phrasal verbs. At present time, a lot of academics focus on phraseology as a science. One of the concepts is related to this notion that phraseological units are cultural phenomena. As a result, phraseology is now a subject of sociolinguists. One of the best definitions of phraseological units is as follows: it is a stable, coherent combination of words with partially or fully figurative meaning [Anna Jansone].

The term "phraseology" is derived from the Greek words phrases-speech and logos-teaching. One-third part of colloquial speech cover phraseological units. Although phraseological units have different names (phraseologisms, phrases, stable combinations of words, phraseological expressions), they still have one main feature, namely constancy. For this reason, phraseological units are stable combinations of words that cannot be understood from the point of view of their grammatical structure in the vocabulary and, most importantly, from the point of view of their understandable component in spoken language [Abdurakhmonova, N. 2021,2022].

Information was selected using documents to collect sentences containing phrasal verbs during data collection. During the data collection process, the novel was read carefully and attention was paid to its sentences. After this, the types and meanings of phrasal verbs were analyzed and reflected in a parallel corpus.

The study of the semantics of phrasal verbs in a parallel corpus serves as a resource for the future creation of linguistic resources of synonymy, antonymy, hyperanonymy or machine translation. If the authors worked with the literature in this direction, now this information is stored in the form of a database and is used as a linguistic resource.

Conclusion

Some of these aspectual meanings are inherent in the basic semantics of certain subgroups of verbs. So, for example, the meanings of incompleteness or repetition can be expressed using lexical derivation (word production), in particular prefixation (oversimplify, outdo, underestimate).

Thus, from the point of view of the most general division of verbs into subclasses, while relying on the grammatical structure of the English language, we can distinguish two aspectual subclasses of phrasal verbs in the English language

The basis of the above division is the relationship of verbal semantics to the idea of procedural limitation, i.e. to any boundary, beyond which the process expressed by the verb or its implied semantics cannot objectively exist. For example, the verb arrive expresses an action that obviously can develop before arriving at the destination. Having reached it, the action stops.

The verb start denotes a transition from some initial state to a subsequent action, thus denoting the boundaries between the two. In contrast to these two verbal units, the verb move

expresses a process that is in itself alien to any idea of limitation (initiative, discontinuous or ultimate).

References:

1. Abdurakhmonova N. Computer models of electronic corpus of the Uzbek language(monograph)/Tashkent: Muharrir, 2021.

2. Alefirenko N.F., Linguoculturology: the value-semantic space of language, 2009.

3. Abdurakhmonova, N., Tuliyev, U., Ismailov, A., & Abduvahobo, G. (2022). Uzbek electronic corpus as a tool for linguistic analysis. In *Компьютерная обработка тюркских языков. TURKLANG 2022* (pp. 231-240).

4. Abduraxmonova, N. Z. Q., & Urazaliyeva, M. Y. (2022). O 'zbek tili elektron korpusida (http://uzbekcorpus. uz/) og 'zaki matnlar korpusini yaratishning nazariy va amaliy masalalari. *Academic research in educational sciences*, *3*(3), 644-650.

5. Mengliev, D., Barakhnin, V., & Abdurakhmonova, N. (2021). Development of intellectual web system for morph analyzing of uzbek words. *Applied Sciences*, *11*(19), 9117.

6. Abdurakhmonova, N. (2019). Dependency parsing based on Uzbek Corpus. In *of the International Conference on Language Technologies for All (LT4All)*.

7. Agostini, A., Usmanov, T., Khamdamov, U., Abdurakhmonova, N., & Mamasaidov, M. (2021, January). Uzwordnet: A lexical-semantic database for the uzbek language. In *Proceedings of the 11th Global Wordnet conference* (pp. 8-19).

8. Shaskiy N.M., Phrazeology of Modern Russian language, 1996.

9. Англо-русский фразеологический словарь: Ок. 20000 фразеол. единиц. – 4 –е изд., перераб. и доп... –М.: Рус. яз., 1984

10. Джейн Поуви Английские фразовые глаголы и их употребление, 1990.