

## INVESTIGATING ENGLISH PHRASAL VERBS USING PARALLEL CORPUS

**DSc Nilufar Abdurakhmonova**

Professor at National university of Uzbekistan, Head of the department

Computational and applied linguistics

[n.abduraxmonova@nuu.uz](mailto:n.abduraxmonova@nuu.uz)

**Salomat Otamurodova**

Tashkent State Agrarian University Teacher at the department "Languages"

[soloahmedovna@gmail.com](mailto:soloahmedovna@gmail.com)

**Abstract:** The use of parallel corpora in the process of artistic translation helps to determine the semantic properties of complex structural combinations in the language and to analyze their meaning in depth. Especially, the research of English phrasal verbs in Uzbek language parallel corpus texts serves to reveal its linguistic nature.

In our study, it was noted that when input examples of English verb phrases in Uzbek literary translation into a parallel corpus, special attention should be paid to their linguistic characteristics. As a result of our observations, the representation of phrasal verbs in the Uzbek language in different shades of meaning in the context was analyzed from a semantic-structural point of view.

**Keywords:** corpus linguistics, phrasal verbs, parallel corpus, artistic translation

### INTRODUCTION

We should emphasize that bringing the Uzbek language to the level of world languages today is an urgent need in the process of integration and globalization. The need for computer technologies, especially in computational linguistics, occupies a very significant place in the implementation of this responsible, important goal. The third millennium has become the age of information for mankind, and now the demand for convenient ways to search, store, process and transmit information to others is increasing day by day. Computer technologies have entered all branches of science, economics, sports, arts, medicine, and all spheres of public life. As mentioned above, the receipt, storage and convenient delivery of information to others without harm to oneself are directly related to the language. Thus was born the science of computational linguistics. The main tasks of the science of computational linguistics are language teaching, knowledge testing, complex text editing and development of programs for machine translation. In the 60s and 80s of the last century, a lot of scientific research was carried out in the field of machine translation.

When creating and developing electronic corpora of world languages, consistent research on NLP and language technologies has become an urgent task. In practice, a lot of research is being carried out in this direction. Thus, computational linguistics is a source of both spiritual and material wealth for linguists [1, 2, 3, 4, 5, 6, 7]. In Russian corpora studies V. P. Zakharov, A. K. Kutuzov, E. V. Nedoshivina, V. Rykov, V. Plungyan, M. Kopotiev studied the corpus, its types, features, social significance of the corpus, principles of constructing corpus. Special studies on the

use of idioms in the Chekhov's works are available by O. Kukushkina, A. Polikarpov and E. Surovtseva.

With the help of various grammatical means, a person expresses his opinion orally, in writing or in his speech, using phraseological units, various expressions, that is, the speech of each person is unique and inimitable. But if the information is transmitted using computer technology, it should be clear, concise and capacious.

The English phrasal verb is a peculiar union of a verb and a particle (prepositional or adverbial) that often produces a unique meaning, uninferable from the meanings of its constituents [8].

The parallel corpus is important because it can contain a lot of necessary information devoted to semantic attitudes of the texts respectively. In the field of machine translation, there are specially formatted multilingual corpora for parallel comparison called structured parallel corpora.

Parallel based studies on phrasal verbs have been conducted in a number researches [9].

Information about the structure, composition and capabilities of the parallel corpus it can be found in the works of D.O. Dobrovolsky, Yu.Tao, V.Zakharov, A.A. Kokoreva, E.P. Sosnina.

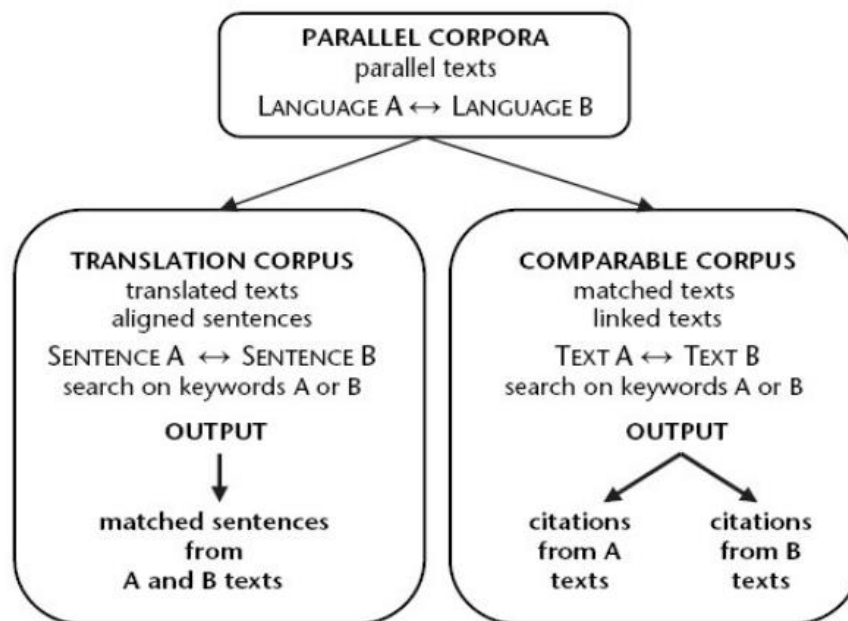
D.O. Dobrovolsky studies can be characterized as corpus-based serial matching of individual pairs correlating phraseological units of different languages in order to identify differences between them, including semantic, syntactic, pragmatic and compatibility ones.

## **METHODOLOGY**

A parallel corpus is a bilingual corpus, that is, the text of the original and its translation into some other language.

A parallel corpus (or parallel corpora) is an electronic analogue of parallel translated texts, usually consisting of many blocks "original text and one / several of its translations". The experts in English corpus linguistics have been working with collections of electronic texts for decades. Electronic texts in the corpus can be a whole original word work or any part of it.

First of all we have to understand parallel corpora as a term. Atkins and Rundell explained pair of translated texts as translation corpus and comparable corpus [11].



The value of this tool is determined by the fact that in linguistics the stage of collecting material is the most time-consuming and least creative, and the corpus of parallel texts can significantly save time and effort for the creative stage of work. In linguistics, translation studies, bilingual lexicography, the importance of parallel corpora in the areas of linguistic comparison is very high.

The use of monolingual corpora in compiling monolingual dictionaries has become a standard in the lexicographic practice pioneered by the Cobuild project in the 1980s [12]. Expressions as a component give major information about equivalents on the lexical item to be translated in parallel corpus. For that reason, as Teubert “parallel corpora are repositories of translation units and their equivalents in the target language” to be re-used in subsequent translations [13].

It is known that the introduction of processes of changes that appear over the years into dictionaries-books requires much effort and money. However always the input and placement of news in language corpora, which is an electronic resource, facilitates this work, and this task can be solved in a short time, without any costs.

Using language data from a corpus, i.e. a carefully compiled collection of naturally occurring language (Biber et al., 1998) can prove to be useful for teaching both general spoken and written English, as well as specific registers usually referred to as English for Specific Purposes (ESP), for example business, legal, medical, engineering, etc. Teachers are increasingly using parallel corpora in the classroom. In so doing they can determine the most frequent patterns of occurrence, enrich their personal knowledge of the language, design teaching materials and provide authentic data in their teaching. Parallel corpora may also be helpful in the planning of teaching units and the identification of specific, potentially problematic, patterns of use and are thus useful tools for syllabus design. Teachers are increasingly encouraging students to make use of parallel corpora themselves in order to become aware of nuances of usage and subtle differences in meaning. Parallel corpora are a recent development, so translators and linguists have barely begun to exploit their full

potential. A large amount of data in a corpus to handle by relying on computer-assisted techniques.

Corpus-based research on phrasal verbs include major phrasal verb dictionary projects which also helped produce a number of practice books for learners of English. For example, major ELT publishers such as Longman, Cambridge University Press or Oxford University Press have all published several reference and practice books based on corpus findings (for example English Phrasal Verbs in Use by McCarthy and O'Dell (1998) and Longman Phrasal Verbs Dictionary (2000) etc.). The comprehensive Longman Grammar of Spoken and Written English (Biber et al., 1999) and its abridged Student Grammar Version (Biber et al., 2002) both take a discourse perspective on the grammatical issues they address including phrasal verbs, and they are both based on corpus findings which are presented to the learner in detail.

In natural language processing (NLP), verb phrases are an essential part of any sentence, and they play a crucial role in determining the meaning of the sentence. In describing phrasal verbs in the Uzbek form. The analysis showed that the use of parallel corpora helps students to identify semantic features of the phrasal verbs. The creation and use of parallel corpora seems appropriate of the development of corpus linguistics – one of the most promising linguistic areas.

Parallel corpora phrasal verbs translation tasks can be difficult and provide samples of professional translation when studying the techniques of translation. Over the past years, many bilingual / multilingual corpora have been created in the world. A great number of works have been done on PV education. In particular, using corpus in translation enables students to achieve in choosing the right or appropriate equivalents in translation process. However, translation sources that Uzbek to English are not many and is considered to be one of the most urgent issues which linguistics to do much.

Bollinger (1971:153) defines phrasal verbs as a semantic unit consisting of a verb plus a particle. He states that this semantic unit has a special degree of what he calls cohesion. In addition, Murphy (2002) explains that the kind of particles that are used with the phrasal verbs can be restricted to the semantic reference of these verbs. For instance, the particle 'out', 'off', and 'up' are used with verbs of movement such as 'drive', 'get' and 'come'. Many linguists say that learning phrasal verbs promotes the development of students' foreign language communicative competence.

While the research on phrasal verbs in the monolingual setting is quite copious, including studies on the syntactic (e.g. Sroka 1972), semantic (e.g. Campoy-Cubillo 1997), pragmatic (e.g. Hampe 2002) and cognitive (e.g. Lindner 1983) features of phrasal verbs, the bilingual aspect is somewhat neglected as only few studies exist on interlingual equivalence and phrasal verbs. Those based on corpus data were conducted by Claridge (2002) with German equivalents and Dezortová 380 (2010) with Czech equivalents. Other studies, much smaller in scope, which examined the equivalents of PVs were done for Polish (Masiulonis 1974, Kretowicz 2005), Russian (Yatskovich 1999) and Arabic (Aldahesh 2009). Given the semantic complexity of phrasal verbs and difficulties with their understanding by non-native speakers of

English, it is worth taking a deeper look into the interlingual differences in the rendition of phrasal verbs into different languages and discover some of the translation strategies used by various translators [8].

In our previous work [14] we estimated as a level machine translation how phrasal verbs translated into Uzbek, moreover we created dictionary [15] (electron version as windows platform) of phrasal verbs that help users to search for all synonyms of verbs not only the main meanings, but also secondary meanings of verbs as well as of that of phrasal verbs. Some separable or inseparable phrasal verbs are taken into account too. Database design is considered to be one of the stages of cycle information system and very crucial task is normalization of each unit in the process of database design. We have gathered more than 12 thousand phrasal verbs and their 3 thousand separate main verbs. But every unit placed in detached cell in a database and overall qualities of them is more than 75 thousand phrasal verbs and main verbs. This database is formed according to following models of phrasal verbs

(V-verb; P-particles (preposition or adverb);

↓ -possible or impossible; smth.-something, smb.-somebody):

– V + oneself +P+smth. => align oneself with smb. or smth.

– V+ oneself +P=> arch (oneself) over

– V + oneself +P+smb. => attach oneself to smb

– V + P+ smb. or smth.+P smth. =>arrange with smb. about smth.

– V+ smb. or smth. + P+ smb. or smth. => associate smb. or smth. with smb.

or smth.

– V+ smth. +P+ smth.=> balance smth. against smth.

– V+ smth. +P=> bail smth. out

– V+ smb. +P+ smth.=> astound smb. with smth.

– V+smb. +P+ smb. or smth.=> bias smb. against smb. or smth. – V+smb.

+P=> beat smb. up

– V +P+smb.=> bet with smb.

– V +P+smb. or smth.=> attend to smb. or smth.

– V +P+ smth.+P+ smth.=>

– V +P+ smb.+P+ smb.or smth.=> book smb. through (to some place) –

V+P+P=> be in for

– V+P+P+it=> be in for it

– V+P+P+smth.=> be off for smth.

– V+P+P+ smb.or smth.=> bound up with smb. or smth.

– V+P+ ↓ P+ ↓ smth=>bear up (against smth.).

On this investigation M.Kh. Boboyeva's works deal with the problems of translating phrasal verbs from English into Uzbek by examples of O.Henry's collection of stories which were translated by Q.Yusupova. While analyzing translation materials the authors observe semantic changes in translation.

There have completed huge quantity of work related to the book Joanna Ketlin Rowling. The book Harry Potter has been translated into more than 80 languages across the world, including Uzbek translator, by Sh.Dolimov. The chief goal of this



work is to classify the collected phrasal verbs in the book “Harry Potter” by J.K. Rowling semantic categories and to find out which categories are possible to be found in the corpus. Such corpora allow learners to select possible equivalents of the studied vocabulary, trace its meanings and functions in certain contexts. In order for the Uzbek language to become one of the world languages, computational linguistics makes it possible to optimize its study and teaching.

The fact that the creation of the national corpus of the Uzbek language has risen to the level of politics is one of the most important issues facing our applied linguistics. Hence in order to increase the prestige of the Uzbek language in society and at the international level, a number of scholars are investigating on creation an electronic corpus of the Uzbek language, including all types of genres of written texts.

**RESULTS:** On the Basis of the novel “Harry Potter” for most striking possibilities of using parallel corpora in the theory and practice of semantic-structural properties of phrasal verbs translation, presenting them in the form of a table of examples have been revealed. The practical part of the thesis consists of the corpus, where an English phrasal verb translation is compared to the Uzbek translator’s equivalents. It consists of the first five chapters, each discussing a different PVs on the side. The focus of the thesis lies in the discussion about translating phrasal verbs, in semantic point of view. The aim of the theoretical part is to justify the translator's choices and to discuss suitable ways of translation of PVs.

From the first chapter of the book Harry Potter common phrasal verb ‘*come in*’ *has* been taken. Its translation version has been interpreted by Sh. Dolimov as the following. Learners have been provided not only literary and figurative meaning but they can explore patterns and search naturally-occurring samples through the use of corpora.

*-Mrs. Dursley was thin and blonde and had nearly twice the usual of neck, which [came in] very useful as she spent so much of her time craning over garden fences, spying on the neighbors.*

*-Qotmadan kelgan sargish sochli missis Durslning bo‘yni esa aksincha, haddan tashqari qiziquvchan tabiati uchun [ayni muddao bo‘lib], oddiy ayolning bo‘yniga nisbatan ikki chandon uzunday go‘yo. Negaki xonim birovning devoridan mo‘ralab qarash, go‘shnilami kuzatib, ulaming sir-asrorlaridan boxabar boiib borishni juda yoqtiradi.*

The author used figurative meaning of the phrasal verb ‘come in’

A translator tried to transfer completely content and form, expressive-emotional impression; the second: even the meaning of idioms is presented correctly, but there are some changes in lexical-grammatical structure of phraseological units. Application of phrasal verbs in everyday speech is so important. Because some of them carry very vague meanings. Even though such phrasal verbs are used in context, the meaning may not be obvious unless we look their meanings up in the dictionary or in any book on idioms. Consequently, translation order can be divided as a fixed phrase which can be met in both languages. There is no need to translate them; it is enough to find suitable and proper phraseological locution from mother tongue’s

<p>They didn't think they could bear it if anyone <b>[found out]</b> about the Potters.</p>	<p>Bu siming oshkor bo'lishi Dursllar uchun chidab bo'lmas andisha bo'lib, bironing ushbu gapdan <b>[voqif bolishi]</b> haqidagi o'zining o'zi er-xotiimi achchiqlantiradi.</p>	<p>High LP In this case the phrasal verb gives the meaning <i>to discover that someone has done something wrong</i></p>
---	---	---

treasure. Sometimes they can match both in meaning and form. The second one is set expression that does not have ready samples in Uzbek language. Undoubtedly, problem of interpretation of phraseological units from one language to another is very significant and complicated task.

**Table 1** shows how some of the many ways in which the English phrasal verb 'come in' is currently used: It carries a number of meanings that depend on the context they are used. They are shown with some examples with Uzbek translation.

<p>Mr. Dursley hummed as he <b>[picked out]</b> his most boring tie for work, and Mrs. Dursley gossiped away happily as she wrestled a screaming Dudley into his high chair.</p>	<p>Mister Dursl allaqanday g'o'ngillagancha, ishga taqib yuradigan ko'rimsiz bo'yinbogiaridan birini <b>[qidirar]</b>, missis Dursl esa baqaloq Dudlini balandgina bolalar kursisiga suqib kiritishga harakat qilib, kechagi mish-mish gaplaming chalasini eriga shosha-pisha oxirigacha so'zlab berishga oshiqar edi.</p>	<p>High LP The phrasal verb 'pick out' is similar to the meaning to choose somebody or something carefully from a group of people or things. Synonym <i>select</i></p>
<p>At half past eight, Mr. Dursley <b>[picked up]</b> his briefcase, pecked Mrs. Dursley on the cheek, and tried to kiss Dudley good-bye but missed, because Dudley was now having a tantrum and throwing his cereal at the walls.</p>	<p>Soat sakkiz yarimda portfelini <b>[ qo'liga olgan ]</b>mister Dursl xotininining yuzidan cho'qidi.</p>	<p>High LP This phrasal verb is a very popular phrasal verb and it means to collect someone from a place or to lift something using hands</p>
<p>"Little tyke," chortled Mr. Dursley as he left the house. He got into his car and <b>[backed out]</b> of number four's drive.</p>	<p>Mashinasiga bir amallab kirib olgan mister Dursl 4-uydan <b>[nari ketdi]</b>.</p>	<p>High LP This phrasal verb with <i>back</i> verb. It means to move or retreat backwards without turning; it is the same a <i>back away</i></p>
<p>But on the edge of town, drills were <b>[ driven out]</b> of his mind by something else.</p>	<p>Ammo shaharga yaqinlashgan mister Durslning parmalar bilan bogliq orzu-havasini yana bir holat miyasidan <b>[quvdi]</b>.</p>	<p>High LP The meaning of drive sb/sth out (of sth ) is to force <i>someone or something to leave or stop doing something</i></p>
<p>He drummed his fingers on the steering wheel and his eyes <b>[fell on]</b> a huddle of these weirdos standing quite close by.</p>	<p>Barmoqlari bilan mashina rulini chertib o'tirar ekan, yonginasida yigilgan, kayfi chogiik-la allanima haqida shivirlashayotgan aqli norasolar to'dasini ko'rib, ular orasida turgan ikki kishi inchalik yosh emasligiga, egniga zumrad zangori rido kiygan anavi kishi esa ancha keksaligiga g'azabla <b>[ e'tibor berdi]!</b></p>	<p>High LP Fall on has been used as <i>steady intent look</i> (of someone's eyes or <i>gaze</i>) be directed towards something.</p>
<p>The traffic <b>[ moved on]</b> and a few minutes later, Mr. Dursley arrived in the Grunnings parking lot, his mind back on drills.</p>	<p>Nihoyat avtomobillar qo'zg'alib bir necha daqiqa o'tgach, mister Dursl yana o'z parmalarini haqida o'ylagancha, &lt;&lt;Grunnings&gt;&gt; idorasining binosi yonidagi mashinalar turadigan maydonchaga <b>[ kirdi]</b>.</p>	<p><i>to leave the place where you are staying and go somewhere else:</i></p>
<p>Mr. Dursley stopped dead. Fear flooded him. He <b>[ looked back ]</b> at the whisperers as if he wanted to say something</p>	<p>Sarosimaga tushgan mister Dursl yashin urganday taqqa to'xtab, shivirlashayotgan anavi odamlarga bir-ikki savol bilan murojaat qilnoqchi bo'ldi-yu, fikridan qaytdi.</p>	<p>to think about something that happened in the past on/to</p>



to them, but thought better of it.		
He dashed back across the road, [ <b>hurried up</b> ] to his office, snapped at his secretary not to disturb him, seized his telephone, and had almost finished dialing his home number when he changed his mind.	Dahshatdan rangi bo'zday oqarib, idoraga gizMladi. yugurgancha xonasiga [ <b>kirar ekan</b> ], kotibaga bo'kirib, uni hechkim bezovta qilmasligini buyurdi.	<i>hurry somebody/something up to make someone do something more quickly, or to make something happen more quickly</i>
[ <b>Come to think of it</b> ], he wasn't even sure his nephew was called Harry.	Familiyasi Potter, o'glining ismi Garri boigan odamlar esa bir talay bolishi mumkin-u, - [ <b>xayoldan o'tkazdi u</b> ], o'sha qarindoshining ismi chindan ham Garri ekanligiga shubha qilib, - ehtimol bolaning ismi Garvi yoki Garolddir.	We use this phrase when we suddenly realize something. This is usually something that is different from what we thought or said earlier. So, we use this when we suddenly remember some piece of information that we had forgotten.
[ <b>There was no point in</b> ] worrying Mrs. Dursley; she always got so upset at any mention of her sister.	Durslni [ <b>bezovta qilishga hech qanday sabab yo'q</b> ], Boyaqish, singlisi to'g'risida gap ochilishi zahoti ta'bi xira bo'ladi.	This expression is <i>used to show that doing something is useless or worthless.</i>
He found it a lot harder to concentrate on drills that afternoon and when he left the building at five o'clock, he was still so worried that he [ <b>walked straight into</b> ] someone just outside the door.	Ishxonani kechki soat beshga yaqin tark etar ekan, o'z harakatini haligacha idora qila olmaganidan bir o'tkinchini [ <b>turtib, qulatib yuborishiga sal qoldi</b> ].	walk into something (informal) to become involved in an unpleasant situation, especially because you were not sensible enough to avoid it.
And the old man hugged Mr. Dursley around the middle and [ <b>walked off</b> ].	Mo'ysafid mister Durslning belidan mahkam quchdida, [ <b>chopgancha nari ketdi</b> ].	to leave a person or place suddenly because you are angry or upset.

It should be noted that interpretation in parallel corpora of phraseological units is the most important and rather difficult task, the culture and national mentality of a define country and nationality should be reflected. The reason is that translators should pay their attention to the coincidence of presentive and logical meanings of phraseological units in the source and target languages during the translation processes and types of contexts in which they are functioning. However, loss and change of stylistic or connotative functions of phraseological units can be the result of misinterpretation of an idiom or the whole expression in the target language.

**CONCLUSION.** Phraseological units are an integral part of any language and knowing these collocations and their adequate interpretation and translation is the proof of proper and adequate translation of the whole expression, fictions and pieces of art.

In Uzbek language it is a single verb which may have different words than the

combined words. Using English-Uzbek parallel corpus in translation classroom enables students to gain access to large information on language usage which will help students in choosing the right or appropriate equivalents in translation process. Based on the reports from the students, they are very positive towards the use of corpus as a tool in translation exercise as they are exposed to a wide range of language use that is beneficial in translating.

Replenishment with texts various genres has created completely new perspectives for lexicographic physical research, including translation lexicography. This a source that allows learners to increase the number of possible transferable equivalents, determine the meaning and frequency of the use of lexical units in certain genres, identify idioms and their variants, verify lexicographic information in already existing dictionaries, etc.

The advantage of using parallel corpora in bilingual lexicography and translation practice is that they provide syntagmatic data about lexical items in both languages but most importantly, they generate ready-made equivalents, or rather, candidates for equivalents to appear in a bilingual dictionary or a translated text [8].

## References

1. Abduraxmonova, N. Z. "Linguistic support of the program for translating English texts into Uzbek (on the example of simple sentences): Doctor of Philosophy (PhD) il dis. aftoref." (2018).
2. Abdurakhmonova N. The bases of automatic morphological analysis for machine translation. *Izvestiya Kyrgyzskogo gosudarstvennogo tekhnicheskogo universiteta*. 2016;2 (38):12-7.
3. Abdurakhmonova N, Tuliyeu U. Morphological analysis by finite state transducer for Uzbek-English machine translation/*Foreign Philology: Language. Literature, Education*. 2018(3):68.
4. Abdurakhmonova N, Urdishev K. Corpus based teaching Uzbek as a foreign language. *Journal of Foreign Language Teaching and Applied Linguistics (J-FLTAL)*. 2019;6(1-2019):131-7.
5. Abdurakhmonov N. Modeling Analytic Forms of Verb in Uzbek as Stage of Morphological Analysis in Machine Translation. *Journal of Social Sciences and Humanities Research*. 2017;5(03):89-100.
6. Kubedinova L, Khusainov A., Suleymanov D., Gilmullin R., Abdurakhmonova N. First Results of the TurkLang-7 Project: Creating Russian-Turkic Parallel Corpora and MT Systems. *Proceedings of the Computational Models in Language and Speech Workshop (CMLS 2020) co-located with 16th International Conference on Computational and Cognitive Linguistics (TEL 2020)*. 2020/11: 90-101
7. Agostini, A., Usmanov, T., Khamdamov, U., Abdurakhmonova, N., and Mamasaidov, M. (2021). Uzwordnet: A lexical-semantic database for the uzbek language. In *Proceedings of the 11th Global Wordnet conference*, pages 8–19.
8. Perdek, M. (2012). Lexicographic potential of corpus equivalents : The case of

English phrasal verbs and their Polish equivalents.

9. Carlos Ramisch and Laurent Besacier and Alexander Kobzar How hard is it to automatically translate

10. Monti, J., Mitkov, R., Corpas Pastor, G. Seretan V. eds. Phrasal verbs from English to French? // Proceedings of the Workshop on Multi-word Units in Machine Translation and Translation Technology, Nice, September 3, 2013.

11. Atkins, B. T. S. and M. Rundell 2008. The Oxford Guide to Practical Lexicography. Oxford: Oxford University Press.

12. Sinclair, John (ed.). 1987. Looking up: An account of the COBUILD project in lexical computing and the development of the Collins COBUILD English Language Dictionary. London: Collins.

13. Teubert, W. 2002. 'The role of parallel corpora in translation and multilingual lexicography.' In B. Altenberg and S. Granger (eds.), Lexis in contrast. Corpus-based approaches, 189–214.

14. Abdurakhmonova N. Database of English-Uzbek verbs and phrasal verbs for machine translation <https://fledu.uz/wp-content/uploads/2019/06/Abdurakhmonova-Nilufar.pdf>

15. Abduraxmonova N. English-Uzbek dictionary of phrasal verbs. – Tashkent: Tronzaminziyo, 2018. – 400 b.