ISSN: 2181-3337 | SCIENTISTS.UZ VOLUME 5 ISSUE 9 | SEPTEMBER 2026

CORPUS-BASED ANALYSIS OF ENGLISH VERB PHRASEMES: COLLOCATION, CONTEXT AND SEMANTIC PATTERNING

Otamurodova Salomat
Tashkent State Agrarian University
Email: soloahmedovna@gmail.com

DOI: 10.67620/sai.59c0005

ABSTRACT

Corpus linguistics has substantially changed the ways in which lexical and grammatical patterns can be investigated. Instead of treating linguistic units primarily as isolated elements, corpus-based research makes it possible to examine their distribution, recurrent combinations and contextual behavior in naturally occurring texts. This article investigates English verb phrasemes from a corpus-oriented perspective, with particular attention to collocation, contextual dependency and semantic patterning. The study draws theoretically on the corpus methodology developed by John Sinclair, especially his emphasis on authentic textual evidence, concordance analysis and recurrent lexical combinations. The analysis demonstrates that the interpretation of verb phrasemes cannot always be adequately explained by considering the meanings of their individual constituents separately. Their functional and semantic characteristics frequently emerge from recurrent relationships between the verbal component, its lexical environment and the wider discourse context. The article argues that corpus evidence provides a more reliable basis for describing such units because it makes their frequency, distribution, co-occurrence and contextual variation observable. Particular attention is also given to the relationship between lexical choice and phraseological organization. The findings suggest that corpus-based investigation can contribute significantly to the classification of English verb phrasemes and to the development of more accurate bilingual dictionaries and translation practices.

ACCEPTED

05.10.26

Keywords: *corpus linguistics, concordance, collocation, verb phraseme, phraseology, semantic patterning.*

CC BY

Introduction

The development of corpus linguistics has brought about a significant methodological transformation in the study of language. Traditional linguistic description has frequently relied on individual examples, introspective judgments and theoretically predetermined categories. Corpus-based research, by contrast, provides an opportunity to investigate language through extensive collections of naturally occurring texts. This methodological shift is particularly important for lexical and phraseological research, where the meaning and function of a linguistic unit are often dependent on the words and structures surrounding it.

John Sinclair's Corpus, Concordance, Collocation occupies an important position in this development. Published by Oxford University Press in 1991, the work examines corpus construction, computational processing, concordance-based investigation and the implications of corpus evidence for

the description of English. Sinclair's approach gives particular importance to actual instances of language use and to recurrent relationships between lexical items. His treatment of collocation is closely connected with his broader argument that linguistic description should be grounded in evidence obtained from texts rather than exclusively in linguistic intuition (Sinclair, 1991).

This perspective is especially productive for the study of verb phrasemes. A verb phraseme is not simply a sequence of independently selected words. In many cases, its semantic interpretation, grammatical behavior and pragmatic function are influenced by the combination of its constituents and by the context in which the combination occurs. Consequently, the analysis of such expressions requires attention not only to their dictionary meanings but also to their patterns of actual use. Corpus evidence is particularly valuable in this respect because it permits researchers to observe repeated linguistic behavior. Concordance lines can reveal which nouns, adjectives, adverbs, prepositions and other elements regularly occur around a particular verb or phrase. Such recurrent combinations provide evidence for identifying phraseological preferences that may remain unnoticed when individual examples are examined in isolation. The present study therefore approaches English verb phrasemes through three interconnected dimensions: collocation, context and semantic patterning.

The main objective is to demonstrate how corpus-oriented analysis can contribute to a more comprehensive description of phraseological units and their meanings.

A corpus can be understood as a structured collection of texts assembled for linguistic investigation. Its principal methodological value lies not merely in the quantity of textual material but in the possibility of systematically observing linguistic behavior across numerous contexts.

Sinclair's work reflects a fundamental methodological principle: linguistic descriptions should be supported by authentic instances of language use. The source emphasizes the importance of examining actual textual evidence and treating individual examples as products of their surrounding contexts. This approach represents an important departure from descriptions based predominantly on invented sentences (Sinclair, 1991).

Such a position is particularly relevant to phraseological research. A phrase may appear perfectly acceptable when constructed by a researcher, while naturally occurring texts may demonstrate that speakers and writers rarely use it in that particular form. Conversely, a relatively unexpected combination may occur repeatedly in authentic discourse and therefore constitute an important feature of actual language use.

Consequently, frequency alone is not sufficient to explain phraseological behavior, but frequency combined with contextual examination provides an important empirical basis for linguistic interpretation.

Concordance analysis provides researchers with a systematic method for examining occurrences of a particular word or phrase in its textual environment. Rather than presenting a lexical item independently, concordance lines place the item within a portion of the original text.

This feature is particularly useful for studying verb phrasemes. Consider the following simplified pattern:

VERB + PARTICLE/PREPOSITION + COMPLEMENT

The same verbal element may occur in several environments, while its semantic and pragmatic characteristics can change according to the accompanying lexical material.

For example, the verb take participates in numerous recurrent combinations:

take care
take part
take place
take advantage
take into account
take responsibility

ISSN: 2181-3337 | SCIENTISTS.UZ

Although take possesses a basic lexical meaning, the semantic contribution of the complete expressions cannot always be predicted by simply adding the dictionary meanings of take and the accompanying word. The phraseological configuration itself contributes substantially to interpretation.

Corpus-based concordance analysis therefore enables the researcher to determine:

which words regularly accompany a verb;

which combinations are particularly frequent;

what grammatical structures occur around the expression;

which semantic environments favor a particular phrase;

whether an apparent phraseological unit is stable or variable;

how the meaning changes according to context.

One of the central concepts in Sinclair's approach is collocation. Later discussions of Sinclair's work describe his treatment of collocation as an important component of his corpus-based theory and relate it directly to the distinction between the open-choice and idiom principles (Barnbrook, 2007; Sinclair, 1991).

From a phraseological perspective, collocation demonstrates that words do not combine with complete freedom. Their co-occurrence is influenced by lexical, grammatical, semantic and contextual preferences.

This observation is crucial for verb phrasemes. A verbal element can have a large number of theoretically possible complements, but actual language use normally reveals a much more restricted set of recurrent combinations.

For example, the following expressions illustrate different degrees of phraseological cohesion:

Table 1

Examples of English Verb Phraseological Combinations and General Meanings

Verb	Phraseological combination	General meaning
make	make a decision	decide
take	take responsibility	accept responsibility
pay	pay attention	attend carefully
reach	reach an agreement	agree
raise	raise a question	ask/introduce a question
bear	bear in mind	remember/consider

These combinations demonstrate that phraseological meaning is frequently associated with the whole configuration rather than with a single lexical element.

The importance of this observation has been widely recognized in subsequent corpus research. Studies discussing Sinclair's model emphasize the relationship between collocation, idiomaticity, meaning and grammatical organization (Barnbrook, 2007). A useful theoretical distinction in corpus-based phraseology is the contrast between the open-choice principle and the idiom principle (Sinclair, 1991).

Under an open-choice interpretation, language can be viewed as a sequence of grammatical choices in which speakers select lexical items to occupy particular structural positions. This perspective is useful for explaining productive grammatical structures.

However, naturally occurring language contains numerous combinations in which lexical choice is considerably more constrained. Speakers do not necessarily construct every expression word by word. Instead, recurrent combinations may be stored and accessed as relatively established patterns.

This second perspective is particularly relevant to verb phrasemes.

ISSN: 2181-3337 | SCIENTISTS.UZ

For example, in:

pay attention

the verb pay is not being used in its most literal financial sense. The semantic interpretation results from the conventional relationship between pay and attention. Similarly:

take into account

cannot be adequately interpreted by treating take, into and account as three completely independent choices.

The corpus-oriented approach consequently encourages researchers to ask not only:

What does this word mean?

but also:

What lexical environments does this word regularly occur in, and what meanings emerge from those environments?

This question represents one of the most productive contributions of corpus linguistics to phraseological studies. A major methodological consequence of corpus research is the recognition that linguistic units are strongly dependent on their contexts.

The meaning of a verb phraseme may be influenced by:

preceding lexical material;

following lexical material;

grammatical structure;

discourse topic;

register;

speaker/writer intention;

semantic field;

surrounding phraseological combinations.

Thus, a dictionary definition alone may not capture the complete functional potential of a phrase.

Consider the expression carry out. In different contexts it may combine with:

carry out an investigation;

carry out an experiment;

carry out a plan;

carry out instructions;

carry out a task.

The semantic core remains relatively stable, but the surrounding lexical environment determines the particular conceptual domain in which the expression operates.

Corpus analysis can therefore reveal that phraseological meaning is not necessarily a fixed property of a sequence independent of context. Instead, meaning may be realized through a relationship between the phrase and its textual environment.

Corpus-Based Analysis of English Verb Phrasemes

For the purposes of phraseological analysis, English verb phrasemes can be investigated through several parameters. The frequency of an expression provides initial evidence concerning its status within the language. A combination repeatedly encountered across independent texts deserves different analytical treatment from a combination appearing only once.

Nevertheless, high frequency should not automatically be equated with phraseological status. Frequency must be considered together with structural stability, semantic cohesion and distribution. The second parameter concerns the words that regularly occur with the verbal expression.

For example:

make + decision

ISSN: 2181-3337 | SCIENTISTS.UZ

is more than an accidental sequence. The repeated association between make and decision represents a conventional lexical preference.

Similarly:

draw + conclusion

shows a strong phraseological relationship in which draw acquires a specialized functional role. Verb phrasemes also display characteristic grammatical structures.

Compare:

take responsibility for something

pay attention to something

make a contribution to something

reach an agreement with someone

The prepositional and syntactic patterns form part of the phraseological behavior of the expressions. Some combinations demonstrate a considerable degree of semantic specialization.

For instance:

give rise to

does not simply denote the physical act of giving something. In many contexts it expresses causation:

X gives rise to Y.

Likewise:

bring about

frequently expresses the idea of causing a particular change or result.

Such expressions demonstrate why phraseological analysis should examine the meaning of the complete construction.

Implications for English–Uzbek Translation

Corpus-based analysis is particularly useful when English verb phrasemes are translated into Uzbek. Structural correspondence between the two languages cannot always be expected.

An English verb phraseme may correspond to:

a single Uzbek verb;

a free Uzbek word combination;

an Uzbek phraseological unit;

a descriptive construction;

a context-dependent lexical equivalent.

For example:

Table 2

English Verb Phrasemes, Possible Uzbek Equivalents, and Translation Strategies

English phraseme	Possible Uzbek equivalent	Translation strategy
take care of	parvarishlamoq / g‘amxo‘rlik qilmoq	one-word or phrase equivalent
make a decision	qaror qilmoq	phraseological/lexical correspondence
pay attention	e’tibor bermoq	functional equivalent
carry out	amalga oshirmoq	phraseological equivalent

ISSN: 2181-3337 | SCIENTISTS.UZ

give up	voz kechmoq	lexical/phraseological equivalent
bring about	sabab bo‘lmoq / yuzaga keltirmoq	contextual equivalent

The table illustrates an important principle: formal similarity is not necessarily the most appropriate criterion for translation. The functional and semantic behavior of the expression in context must also be considered.

This has direct implications for bilingual lexicography. A dictionary that provides only isolated word equivalents may fail to represent the phraseological restrictions of English verbs. Corpus-derived information can instead demonstrate which combinations are conventional and what Uzbek equivalents correspond to them in authentic contexts.

The corpus-based approach to verb phrasemes has several methodological advantages.

First, it provides empirical evidence. Instead of relying exclusively on the researcher's linguistic intuition, the researcher can examine repeated occurrences in actual texts.

Second, corpus analysis makes lexical patterning visible. Recurrent combinations can be identified systematically, allowing researchers to distinguish productive combinations from more established phraseological structures.

Third, concordance analysis allows the researcher to retain contextual information. This is essential because the meaning of a phraseological unit can depend on its surrounding discourse.

Fourth, corpus data can contribute to lexicographic accuracy. Dictionaries can provide more realistic definitions and examples when their descriptions are based on actual language use. Sinclair's corpus-oriented research was closely associated with lexicography and the development of corpus-based reference materials.

Finally, this approach has significant implications for foreign-language teaching. Learners need to know not only what a verb means but also what words it naturally combines with, which grammatical structures it selects and in what contexts particular combinations are appropriate.

The significance of collocation in this respect has been extensively discussed in subsequent evaluations of Sinclair's work, which emphasize the relationship among collocation, phraseology, meaning and grammar (Barnbrook, 2007).

Conclusion

Corpus-based investigation offers a productive methodological framework for the study of English verb phrasemes. The analysis demonstrates that these expressions should not be examined solely as combinations of independently meaningful words. Their linguistic behavior is determined to a considerable extent by recurrent lexical, grammatical and contextual relationships.

Concordance analysis makes such relationships observable by presenting lexical units in authentic textual environments. Collocational patterns, frequency, grammatical structure and contextual variation can consequently be examined as interconnected aspects of phraseological organization.

The corpus-oriented perspective associated with Sinclair provides an important theoretical foundation for this type of investigation. Rather than treating language as an abstract collection of isolated lexical and grammatical units, it directs attention toward the recurrent patterns through which language is actually used (Sinclair, 1991).

For English–Uzbek comparative phraseology, this methodology is particularly valuable. It can help identify functional equivalents, distinguish literal from phraseological meanings, establish translation strategies and improve the representation of verb phrasemes in bilingual dictionaries. Further research based on a balanced English corpus and a comparable Uzbek corpus could provide quantitative evidence concerning the frequency, collocational strength and semantic distribution of individual verb phrasemes.

ISSN: 2181-3337 | SCIENTISTS.UZ

REFERENCES

1. Barnbrook, G. (2007). Sinclair on collocation. International Journal of Corpus Linguistics, 12(2), 183–199.
2. Biber, D., & Reppen, R. (Eds.). (2015). The Cambridge handbook of English corpus linguistics. Cambridge University Press.
3. Gardner, D., & Davies, M. (2007). Pointing out frequent phrasal verbs: A corpus-based analysis. TESOL Quarterly, 41(2), 339–359.
4. Granger, S., & Meunier, F. (Eds.). (2008). Phraseology: An interdisciplinary perspective. John Benjamins.
5. Gries, S. T. (2008). Dispersions and adjusted frequencies in corpora. International Journal of Corpus Linguistics, 13(4), 403–437.
6. Gries, S. T. (2009). Quantitative corpus linguistics with R: A practical introduction. Routledge.
7. Liu, D. (2011). The most frequently used English phrasal verbs in American and British English: A multicorpus examination. TESOL Quarterly, 45(4), 661–688.
8. McEnery, T., & Hardie, A. (2012). Corpus linguistics: Method, theory and practice. Cambridge University Press.
9. McEnery, T., Xiao, R., & Tono, Y. (2006). Corpus-based language studies: An advanced resource book. Routledge.
10. Sinclair, J. (Ed.). (1987). Looking up: An account of the COBUILD project in lexical computing. Collins.
11. Sinclair, J. (1991). Corpus, concordance, collocation. Oxford University Press.
12. Sinclair, J. (1996). The search for units of meaning. Textus, 9, 75–106.
13. Sinclair, J. (2004). Trust the text: Language, corpus and discourse. Routledge.
14. Stubbs, M. (2001). Words and phrases: Corpus studies of lexical semantics. Blackwell.
15. Wray, A. (2002). Formulaic language and the lexicon. Cambridge University Press.

ISSN: 2181-3337 | SCIENTISTS.UZ