

THE TYPES AND MEANING OF COLLOCATIONS

(On the basis of examples taken from “Windmills of the Gods” by Sydney Sheldon)

SARVAR SOGLOMOV

Doctoral Researcher at the Department of Linguistics, Faculty of Foreign Languages, Karshi State University. Specialization: 10.00.06 – Comparative Literature, Contrastive Linguistics, and Translation Studies.

E-mail: sarvarsoglomov@gmail.com

Abstract. The article explores English collocations, their types, and spheres of usage, based on the views of prominent scholars in the field. A number of collocations from the novel *Windmills of the Gods* by the renowned American writer Sidney Sheldon, which serves as the primary material for our research, have been analyzed in detail.

Keywords. phraseological units, collocation, idiom, free combinations, semi-phrasemes, sentence, vocabulary, meaning, explanation, word combination.

Collocations, i.e. arbitrarily restricted lexeme combinations such as *make a decision* or *fully aware*, are one type of a group of expressions whose importance in language has been increasingly recognized in recent years. This group of expressions has been variously called prefabricated units, prefabs, phraseological units, (lexical) chunks, multi-word units, or formulaic sequences. They are made up of more than one word and are lexically and/or syntactically fixed to a certain degree. Following a period in which, largely due to the wide influence of generative grammar, prefabricated units were considered peripheral in language, it is today widely assumed that their number is vast and that they play a major role in language processing and use. Bolinger was among the first linguists to point out that a generativist view, which relegates prefabricated units to the periphery of language, fails to account for a considerable part of observable language data. On the basis of numerous examples, he claims that our language does not expect us to build everything starting with lumber, nails, and blueprint. Instead it provides us with an incredibly large number of prefabs.

The term ‘collocation’ is used in widely different and often rather vague senses in linguistics and language teaching. The only common denominator is that the term is (at least mostly)

used to refer to some kind of syntagmatic relation of words. Among the many diverse uses of the term, two main views can be identified. In one of these two views, a collocation is considered the co-occurrence of words at a certain distance, and a distinction is usually made between co-occurrences that are frequent (or more precisely, more frequent than could be expected if words combined randomly in a language) and those that are not. This view has therefore been called the 'statistically oriented approach' or the 'frequency-based approach'. In the other view, collocations are seen as a type of word combination, most commonly as one that is fixed to some degree but not completely. This view has been referred to as the 'significance oriented approach' or the 'phraseological approach'. The frequency-based approach goes back to J. R. Firth and has been developed further in particular by M. A. K. Halliday and J. Sinclair. It is often adopted by researchers who are involved in the computational analysis of syntagmatic relations. The phraseological approach has been strongly influenced by Russian phraseology. Typically, researchers adopting this approach work in the fields of lexicography and/or pedagogy; among the main representatives are A. P. Cowie and I. Mel'cuk. In what follows, I will describe the view of collocations propounded by one of the major representatives of each of the two approaches and briefly outline how other representatives of the two approaches differ from them.³ For the frequency-based approach, Sinclair's view of collocations will be discussed, for the phraseological approach, that of Cowie.

Sinclair defines collocations as "the occurrence of two or more words within a short space of each other in a text". A short space, or 'span', is usually defined as a distance of around four words to the right and left of the word under investigation, which is called the 'node'. If, for example, in a given amount of text, the word *house* is analysed, and the word occurs in an environment such as *He went back to the house. When he opened the door, the dog barked*, the words *went, back, to, the, when, he, opened, the* are all considered to form collocations with the node *house*; these words are then called 'collocates'. Sinclair distinguishes two types of collocations, namely 'significant' and 'casual' collocations, and sometimes reserves the term 'collocation' for the former type. Significant collocations are co-occurrences of words "such that they co-occur more often than their respective frequencies and the length of text in which they appear would predict". In the example above, *the* and *house* would probably not be significant collocations, as, although these two words can be assumed to co-occur frequently, *the* is itself a frequent word in virtually every kind of text. The words *dog* and

barked would, however, very likely constitute a significant collocation, as *barked* is not usually very frequent and, if it occurs, is likely to be found near the word *dog*. Exact formulae of how to determine exactly whether co-occurring words constitute a significant collocation have also been developed by Sinclair and others.

Given that even Sinclair sometimes varies in how he defines collocations, it is not surprising that some researchers adopting a frequency-based approach to collocations consider co-occurrences of all frequencies to be collocations, while others reserve the term for frequent co-occurrences. Some use recurrence, i.e. co-occurrence more than once in a given corpus, as the defining criterion. Other points of variation in the definition of collocations in the frequency-based approach are also mirrored by variation in Sinclair's writings. Whereas he uses 'word' in the sense of 'lexeme' in the above definition, and thus sees collocation as a relationship between lexemes, he previously regarded it as a relationship between 'lexical items'. This latter view is also shared by Halliday, who exemplifies 'lexical item' with the group of derivationally related lexemes *STRENGTH*, *STRENGTHEN*, *STRONG*. According to this view, a *strong argument*, *he argued strongly*, *the strength of the argument*, *his argument was strengthened* would all be considered instances of the same collocation. A third view on this question is that collocation is a relationship between word forms, i.e. that combinations such as *hold tight* and *holds tight* are two different collocations. A more fundamental aspect in which definitions vary is the question of the nature of the collocation as such. Sometimes 'collocation' seems to be used purely to describe a phenomenon in a given amount of text (as in the above definition by Sinclair); more commonly, it also seems to be considered a more abstract tendency in a language. Further points that are viewed differently by authors adopting a frequency based approach are the number of words involved in a collocation and whether or not these have to be consecutive. Occasionally, as in the above definition, "two or more words" are considered to constitute a collocation; often only two words are allowed. The fact that the words are consecutive is, for example, required by Kjellmer; Firth at times considers whole sequences such as *[i]s all the world drowned in blood and sunk in cruelty* as collocations.⁵ A final aspect in which definitions vary is the syntactic relationship of the elements involved in a collocation. In the frequency-based approach, the syntactic relationship between the elements does not usually play a role in deciding whether they form a collocation or not. Among the few exceptions are Kjellmer and Greenbaum. Kjellmer excludes from his definition sequences that have no or only a very

distant grammatical relationship: night he, for example, in a sentence such as *At night, he suddenly remembered what had happened*, would not be considered a collocation according to his definition, even if the criterion of (relative) frequency is met. Greenbaum's definition of collocations only includes words that stand in a close grammatical relationship (such as *adverb + adjective*; 1970). However, as he at the same time completely dismisses the criterion of co-occurrence in a certain span (although he retains the criterion of frequency of co-occurrence), Greenbaum is among the less typical representatives of the frequency-based approach, and his definition approaches the phraseological view of collocations. A. P. Cowie is a typical representative of the phraseological approach: he considers collocations a type of word combination, i.e. an abstract combination with instantiations in actual texts, and defines them by delimiting them from other types of word combinations, most importantly from idioms on the one side and from what he sometimes calls 'free combinations' on the other. At the same time, he is one of the most important representatives of the phraseological approach, as his attempts to define collocations and to delimit different kinds of word combinations are among the most precise. Cowie divides word combinations into two main types, 'composites' and 'formulae'. Formulae are combinations with a primarily pragmatic function such as *How are you?* Or *Good morning*. Collocations belong to the group of 'composites', which are described as having a primarily syntactic function. The distinctions in the group of composites are made on the basis of two criteria, which Cowie assumes to interact closely: the criterion of transparency and the criterion of commutability (or substitutability). Transparency refers to whether the elements of the combination and the combination itself have a literal or a non-literal meaning, and commutability refers to whether and to what degree the substitution of the elements of the combination is restricted. On this basis, he distinguishes the following four types of combinations, stressing, however, that these types are not clearly delimitable, but should rather be seen as forming a continuum:

Free combinations:

- the restriction on substitution can be specified on semantic grounds
- all elements of the word combination are used in a literal sense

Free combinations refer to word groupings in which the elements can be freely combined and replaced without violating grammatical rules. These combinations contrast with set expressions like idioms or collocations, where word choice is fixed or semi-fixed. Here are some core ideas and perspectives on free combinations from a linguistic point of view:

- From the **definitional** viewpoint, free combinations are syntactically and semantically compositional – the meaning of the whole is the sum of its parts.

Example: *buy a book, read a letter, open a door*

- From the **syntactic** viewpoint, these combinations obey general syntactic rules rather than idiomatic constraints and here word classes or parts of speech (nouns, adverbs, verbs) can be freely combined as long as they are grammatically compatible.

Adjective + noun: *beautiful garden, cold water*

Verb + noun: *write a letter, build a house*

- From the semantic viewpoint, the meaning of free combinations is literal and transparent, so there is no hidden or figurative sense.

Example: *drink tea*

(*Drink tea* is a free combination and it means *the act of consuming tea*; on the contrary, *spill the tea* is an idiom that means *reveal gossip*.)

- When contrasting free combinations with collocations and idioms, one should focus on the fact that collocations involve habitual co-occurrence (*start a family* but not *begin a family*) when free combinations have no lexical restrictions (*buy a book – read a book – give a book*). Idioms are fixed and require cultural or figurative interpretation while free combinations are not fixed and do not require cultural and figurative interpretation.

Comparison table:

Type	Example	Replaceable	Meaning derived
Free combination	<i>drink tea</i>	Yes	Yes
Collocation	<i>strong tea</i>	Limited	Yes
Idiom	<i>kick the bucket</i>	No	No (figurative)

Restricted collocations:

- some substitution is possible, but there are arbitrary limitations on substitution
- at least one element has a non-literal meaning, and at least one element is used in its literal sense; the whole combination is transparent

Restricted collocations are a type of word combination in which the words co-occur more often than chance would predict, but they are not completely fixed like idioms. In other words, there is some restriction on word choice, often due to convention, meaning constraints, or usage norms in the language. They are partially fixed and partially flexible.

- From the viewpoint of **partial fixity**, in restricted collocations, one element (a verb or an adjective) is flexible whilst the other (a noun) is relatively fixed.

Example: *strong tea* → *strong coffee*

- From the viewpoint of **limited substitutability**, in restricted collocations, substituting a different word can sound **unnatural** or **non-native**, even though it's grammatically correct.

Example: *heavy rain* (but not *strong rain*)

- From the viewpoint of **frequent co-occurrence**, in restricted collocations, word pairs are **frequent and predictable** combinations in the language, often learned as a unit.

Example: *pay attention*, *catch a cold*, *take a risk*

- From the **semantic** viewpoint, restricted collocations are **semantically transparent** to some extent, whilst idioms are not.

Example: *break the news* means *to tell news*. Here, the collocation *to break the news* is literal or slightly metaphorical.

- From the cultural viewpoint, some restricted collocations are based on cultural norms or lexical preference.

Example: in English they say *blond hair*, but not *yellow hair*.

Figurative idioms:

- substitution of the elements is seldom possible
- the combination has a figurative meaning, but preserves a current literal interpretation

Figurative idioms are fixed or semi-fixed multi-word expressions whose meanings are not deducible from the literal meanings of the individual words. They are often metaphorical and require cultural or contextual knowledge to understand.

- From the viewpoint of **meaning**, the meaning of figurative idioms cannot be predicted from its components.

Example: *spill the beans* means *to reveal a secret*, but here *the beans* are not literally dropping legumes.

- From the viewpoint of **form**, figurative idioms have fixed structure, though some variation is possible.

Example: *kick the bucket* which means *to die* cannot become *kick a bucket* in the same sense. In the same sense we cannot replace *a* with *the* in the figurative idiom *kick the bucket*.

- From the viewpoint of **figurative language**, figurative idioms often involve metaphors, metonymy, or symbolism.

Example: *an oak in the forest of willows* means that someone stands out from the people surrounding or in the society.

- From the viewpoint of **cultural dependence**, figurative idioms are deeply rooted in cultural and historical contexts.

Example: *the ball is in your court* (from tennis) means *your turn to act*.

Figurative idioms is understood by native speakers without needing explanation while they often pose challenges for second language learners.

Pure idioms (e.g. *blow the gaff*):

- substitution of the elements is impossible
- the combination has a figurative meaning and does not preserve a current literal interpretation

Pure idioms (also called opaque idioms) are a subtype of figurative idioms where the overall meaning cannot be inferred at all from the meanings of the individual words. They are completely non-compositional and often have fixed structure. Unlike semi-idioms or collocations, no part of a pure idiom is interpreted literally.

Pure idioms:

- the meaning cannot be guessed from the parts;
- form is usually very rigid and any changes in the form may break the idiom;
- lexical items in pure idioms lose their original meaning;
- pure idioms are often culture bound and need contextual or historical knowledge.

Idiom	Actual meaning	Literal meaning
<i>Rain cats and dogs</i>	<i>Raining very heavily</i>	<i>Animals falling from the sky</i>
<i>Bite the dust</i>	<i>To die or fall</i>	<i>Chew on dirt</i>
<i>Kick the bucket</i>	<i>To die</i>	<i>Strike a pail with foot</i>

The most important variation in Cowie's use of the term 'collocation' is that while he sometimes applies it only to combinations with an arbitrarily limited substitutability in which

one element is used in a non-literal sense, he sometimes applies it to free combinations as well. In this case, however, he makes a distinction between 'open collocations' (i.e. free combinations) and 'restricted collocations'. He also varies in categorizing combinations of the type *foot the bill*, in which one word in a given specialized meaning (*foot* in this case) can co-occur only with one other word. While he usually subsumes such combinations under the category 'restricted collocations', at least in one paper he regards them as constituting an additional category between idioms and collocations. A third aspect in which his definition varies is that while he usually assumes that the elements of a collocation are lexemes, he assumes in at least one publication that these elements are 'roots', abstract units comprising all inflectional and derivational forms of a word, similar to Halliday's and Sinclair's 'lexical items'.

As in the case of Sinclair, Cowie's variation in the use of the term reflects some of its different uses by different authors adopting a phraseological approach. A number of researchers apply the term 'collocations' to both free combinations and restricted collocations. Some of these do not differentiate further while others, like Cowie, distinguish between 'open collocations' (or 'free collocations') and 'restricted collocations'. More frequently, authors adopting a phraseological approach reserve the term 'collocation' for Cowie's restricted collocations and use different terms, such as 'free combinations' or 'co-creations', for unrestricted combinations. The number of categories towards the more restricted and opaque end of the scale also varies between authors. Cowie's distinction between two types of idioms (figurative idioms and pure idioms) is often not made, and Benson et al., for example, consistently postulate an additional category between collocations and idioms, which they call 'transitional combinations' or 'transitional collocations'.

In addition to the two main approaches and their variations, collocations have been defined in numerous other, more idiosyncratic ways. Benson et al., for example, also use the term to refer to what are more commonly called valency patterns, such as *suggest + -ing*. A few authors include compounds, and van der Wouden even extends the term to cover combinations of morphemes that are to some degree fixed (e.g. *cran-berry* or *ox-en*). Examples of idiosyncratic usage can also frequently be found in the area of language teaching. Taylor, for example, includes paradigmatic relations in her definition of collocations. The frequency-based and the phraseological approach are also sometimes mixed, with some authors who primarily adopt a phraseological approach additionally

considering frequency as a defining criterion. Some authors primarily working in the framework of the frequency-based approach, in turn, have also introduced phraseological distinctions. What can also be found is a double use of the term 'collocation', i.e. its use in both the sense of the frequency-based approach and the phraseological approach in one and the same piece of work. F. R. Palmer, for example, on the one hand reserves the term for free and restricted combinations as opposed to idioms, and on the other refers to "the collocation of kick and the bucket", where 'collocation' apparently means co-occurrence. Finally, a few other terms can be found for the syntagmatic phenomena described above, in particular for collocations in the phraseological sense, such as 'non-idiom phraseological units' or 'idioms of encoding'.

References:

1. Morton Benson, Evelyn Benson & Robert Ilson. *Lexicographic Description of English*. Netherlands. John Benjamins Publishing Company. 1986. (302 pages)
2. Morton Benson, Evelyn Benson & Robert Ilson. *The BBI Dictionary of English Word Combinations*. Amsterdam: John Benjamins Publishing Company. 1987. (320 pages).
3. Cowie, Anthony P. "The treatment of collocations and idioms in learners' dictionaries." *Applied Linguistics*, 2, 223–235. 1981.
4. Halliday, Michael A. K. *The Language of Science*. Collected Works of M. A. K. Halliday. Continuum. London. 2004. (268 pages)
5. Sinclair, John. *Corpus, Concordance, Collocation*. Oxford University Press. 1991. (177 pages)
6. Palmer, Harold E. *A Grammar of English Words*. London: Longmans.
7. Kjellmer, Göran "Aspects of English collocations." (pp. 133–140). Amsterdam: Rodopi. (1987).
8. Greenbaum, Sidney. *Verb-Intensifier Collocations in English. An Experimental Approach*. The Hague: Mouton. 1970. (100 pages)
9. Mel'cuk, Igor. "Collocations and Lexical Functions." In Anthony P. Cowie (Ed.), *Phraseology* (pp. 23–53). (1998)
10. Bolinger, Dwight (1979). "Meaning and memory." In George G. Haydu (Ed.), *Experience Forms: Their Cultural and Individual Place and Function*. The Hague: Mouton. 1979. (95-113 pages)