

# Comparative Analysis of English Collocations Dictionaries: A Study of the Oxford Collocations Dictionary and the Macmillan Collocations Dictionary

Akerke Yessenali  
European Master in Lexicography (EMLex) student  
Université de Lorraine, France  
akerke.yessenali@mail.ru

## Abstract

This research delves into the growing significance of phraseological information in lexicography, specifically through a comparative analysis of two prominent English collocations dictionaries: the Oxford Collocations Dictionary (OCD, 2002) and the Macmillan Collocations Dictionary (MCD, 2002). As collocations play a pivotal role in language acquisition, the study critically evaluates the methodologies used by each dictionary in encoding, classifying, and presenting collocational information. Carefully studied key aspects include the treatment of collocations, entry features, and overall organization.

The OCD adopts a grammatical structure-based categorization and features a dedicated phrase section, while the MCD uses semantic grouping and provides noun collocates for adjective and verb entries. The analysis encompasses dictionary entries, evaluating definitions, collocational categories, explanatory content, and usage examples. Findings underscore the importance of diverse access methods, clear classification, comprehensive descriptive information, and contextual examples for effective language learning.

The comparative study reveals variations in the representation and organization of collocational information, emphasizing the absence of standardized approaches. Consequently, it advocates for further exploration and standardization efforts in phraseological knowledge. This research contributes valuable insights to the discourse on collocations' significance in language learning and informs the representation and structure of collocational information in lexicographic resources.

**Keywords:** collocations; collocations dictionaries; phraseological knowledge; language learning; lexicographic resources

## 1. Introduction

The significance of phraseological knowledge, spanning diverse lexical combinations and linguistic expressions, has become more prominent with the emergence of combinatorial or collocations dictionaries. This paper seeks to explore and comparatively analyze the two most representative English collocations dictionaries designed for general language

use. The primary objective is to assess their potential value for individuals engaged in the process of learning the language.

Adding a temporal dimension to the importance of collocations dictionaries, the dynamic nature of language and its continual evolution over time accentuate their significance. Language is inherently responsive to societal changes, technological advancements, and emerging trends. Collocations, in this context, serve as linguistic markers that reflect these shifts and adaptations. Rundell's perspective (2010) aligns with this notion, positing that collocations are as crucial as grammar, contributing significantly to the attainment of natural and fluent language expression. While grammar provides the structural framework, collocations offer nuanced patterns that enhance the authenticity and effectiveness of communication, enriching both spoken and written language.

Ensuring grammatical accuracy is crucial for producing error-free text, but the selection of appropriate collocations is equally essential for achieving a natural and fluent tone. As Bergenholtz and Tarp (2010: 33) note, the inclusion of more collocations in a dictionary enhances its functionality. Furthermore, collocation plays a vital role in conveying meaning, particularly in contexts where common English words have multiple interpretations. The surrounding context, often guided by collocations, aids in discerning the intended meaning within a given communication, adding layers of nuance to language comprehension.

The English language currently has a number of general collocations dictionaries available for purchase. This paper centers its examination on the *Oxford Collocations Dictionary* (OCD, 2002) and the *Macmillan Collocations Dictionary* (MCD, 2002), focusing on their approaches to presenting collocational information. Despite the prevalence of electronic works of reference in the modern era, this study focuses on traditional paper dictionaries for several reasons. Firstly, paper dictionaries remain widely used and valued by language learners, educators, and researchers, providing tangible and accessible resources. Secondly, the chosen dictionaries, despite being twenty years old, are still regarded as authoritative references in the field, offering insights into historical approaches to lexicography that may inform contemporary practices. Thirdly, by examining these paper dictionaries, this study aims to bridge the gap between traditional and digital lexicographic resources, shedding light on enduring methodologies and practices in collocational analysis.

Each of these resources, as revealed in this paper, adopts distinct methods in encoding, classifying, and presenting collocations. Nuccorini (2003) highlights substantial variations in linguistic and lexicographic aspects, encompassing boundaries, content descriptions, theoretical principles guiding inclusion, selection, classification, and presentation of headwords, as well as the sources and layout utilized in phraseological dictionaries. The variations in these aspects will be explored in detail within this paper. The subsequent sections will delve into a detailed exploration of these variations, offering a comprehensive understanding of the nuances within each dictionary's approach.

## 2. Oxford Collocations Dictionary

The OCD was initially released in 2002. Similar to the MCD, it is a collocations dictionary for English, designed for upper-intermediate to advanced English learners. While sharing a common educational objective, a distinct feature sets the OCD apart.

A distinctive feature of this dictionary is its notable absence of noun collocates for verb and adjective entries. For instance, in the case of the verb "compare" in Figure 1, the OCD presents adverb collocates ("favourably," "well"), verbs ("cannot," "do not"), prepositions

(“with”), and phrases (“be nothing compared to sb/sth”). Nonetheless, no information is provided concerning noun collocates.

### **compare** *verb*

- ADV. **favourably, well** *The city compares favourably with other parts of Brazil.* | **unfavourably**
- VERB - COMPARE **cannot, do not** *These mountains do not compare with* (= are not nearly as high, impressive, etc. as) *the Himalayas.*
- PREP. **with** *Few things compare with* (= are as good as) *the joy of cycling on a bright, sunny spring morning.*
- PHRASES **be nothing compared to sb/sth** *I've had some difficulties but they were nothing compared to yours* (= they were not nearly as bad as yours).

Figure 1: Entry for “compare” in the OCD.

The OCD sets itself apart from the MCD through its exclusive incorporation of a dedicated section for phrases (see Figure 1). In contrast, the MCD excludes a specific section for phrases, a decision explained by its creators. They argue that the OCD already incorporates a “phrases” category, where items are often included due to their lack of transparent meaning (Kilgarrieff et al. 2010: 373). Despite these slight differences in approach, the OCD’s method of delineating phrases proves highly valuable. The segregation of phrases into a distinct section enhances accessibility and understanding for users, especially non-native speakers who might find everyday phrases challenging to comprehend.

In the OCD, a polysemic entry entails the provision of distinct definitions for each sense, whereas a monosemic entry, corresponding to a single meaning, lacks a dedicated definition. Illustrated in Figure 2, subsequent to each definition, relevant collocations are distinctly formatted in bold, facilitating their identification within the text. The deliberate formatting choice serves not only as visual emphasis but also as a practical aid for efficiently identifying and extracting these linguistic associations within the broader textual context. These collocations are systematically organized by their grammatical structures, encompassing categories like Verb + *light*, Adverb, and Preposition.

### **light** *verb*

#### **1** make sth begin to burn

- VERB + LIGHT **attempt to, try to | pause to, stop to**  
*She paused to light another cigarette.*

#### **2** (often **be lit**) give light to sth

- ADV. **well** *a brightly lit room* | **badly, dimly, poorly** *a dimly lit street* | **brightly, brilliantly** | **softly** | **briefly, momentarily** (*figurative*) *A gleam of humour momentarily lit his face.* | **suddenly** | **artificially, electrically** | **up**  
*There was an explosion and the whole sky lit up.*
- PREP. **with** (*figurative*) *Her face lit up with pleasure.*

Figure 2: Entry for “light” in the OCD.

The OCD utilizes a vertical line “|” to delineate meaning distinctions within identical grammatical relations (McGee 2012). McGee contends that while the collocations in the OCD entries exhibit semantic organization, this organization lacks explicit labeling. Contrary to McGee’s perspective, it can be asserted that the semantic organization in this dictionary is somewhat deficient due to its failure to explicitly specify the semantic relationship between the headword and the collocations, leaving it up to the users to deduce.

This deficiency becomes evident upon a thorough examination of the dictionary, particularly when reviewing the entry for “bed” and its associated collocates, as illustrated in Figure 3. Notably, there is a perceptible segregation in the initial sense, where collocates such as “double,” “king-size,” “single,” and so forth are distinctively set apart from other collocates like “feather,” “sofa,” and so on. This segregation poses a challenge for users in comprehending the relationship between disparate terms like “feather bed” and “sofa bed.” The lack of explicit semantic guidance becomes pronounced, particularly in entries where such relationships might be less intuitive.

**bed** *noun*

**1** piece of furniture for sleeping on

- ADJ. **double, king-size, single, twin** | **bunk, camp, feather, folding, four-poster, hospital, sofa** | **warm** | **unmade** *a messy room, with an unmade bed and clothes on the floor* | **marriage**
- VERB + BED **do** (*informal*), **make, make up** | **strip** *Please strip the beds and put the sheets in the washing machine.* | **climb into, crawl into, get into, go to, tumble into** *She crawled into bed exhausted.* | **climb out of, get out of, leap out of** | **lie (down) on, lie in, sit on** *He lay in bed, reading his book. ◇ Elizabeth was sitting on her bed writing a letter.* | **put sb to, tuck sb up in** *It's your turn to put the children to bed.* | **wet** *Don't punish a child who wets the bed.*
- BED + NOUN **clothes, linen**
- PREP. **in ~** *I like to be in bed before 11 o'clock.* **out of ~** *Are you out of bed yet?*
- PHRASES **bed and breakfast, the edge/side of the bed, the foot/head of the bed, get sb into bed** (= have sex with sb), **go to bed with sb** (= have sex with sb), **take to your bed** (= go to bed because you are ill), **time for bed** *Come on, children, it's time for bed.*

**2** piece of ground for growing flowers, vegetables, etc.

- ADJ. **flower, rose, strawberry** | **ornamental** | **raised**
- PREP. **~of** *ornamental beds of roses*

Figure 3: Entry for “bed” in the OCD.

The OCD uses explanatory information enclosed in brackets to offer additional clarity on meaning and usage when required. Additionally, it introduces usage examples in italics to exemplify the application of each collocation within a contextual framework. Recognizing the pivotal role of contextual information in language acquisition, the dictionary acknowledges that the meaning of a word is subject to significant variation based on its context.

It is widely acknowledged that the availability of contextual information is paramount for effective language learning, given the nuanced nature of word meanings. Naturally, due to space constraints, any lexicographic resource can only provide a finite number of micro-

contexts for a given word. Nevertheless, the efficacy of a dictionary for language learners is intrinsically tied to the richness of the contextual information it offers. The extensive and diverse contextual insights play a crucial role in enhancing the learner's proficiency to comprehend and adeptly use words with precision and appropriateness across a spectrum of situations.

### 3. Macmillan Collocations Dictionary

The MCD, similar to the OCD, aims to assist upper-intermediate to advanced English students in enhancing their writing abilities for the purpose of achieving success in the International English Language Testing System (IELTS). Just like the OCD, it is based on a corpus, which is a collection of authentic language samples. Furthermore, despite MCD having approximately half the number of entries as the OCD, the two dictionaries are similar in length.

Figure 4 illustrates the entry for the verb "light" as presented in the MCD. Each headword in this dictionary is accompanied by its respective definition(s). Differing from the OCD, the MCD consistently provides a definition for each headword, even when there is only a single meaning. This attribute enhances the utility of the MCD for both encoding and decoding. From a language acquisition standpoint, it makes easier it for learners to recognize correspondences in both the production phase and early reception stages. Additionally, in keeping with the OCD's methodology, the MCD incorporates italicized usage examples to illustrate the appropriate contextual application of each collocation.

**light** v  
**make a place brighter by giving it light** [usually passive]  
 ✱ **adv+V with not much light** *dimly* *She led me through the dimly lit corridors.*  
 ➤ **well properly, well** *Always park somewhere that is well lit and preferably somewhere with CCTV.*  
 ➤ **badly badly, poorly** *The way to the station is so badly lit, it is unsafe at night.*  
 ➤ **brightly brightly, brilliantly** *We all received a very warm welcome as we entered the brightly lit hall.*

Figure 4: Entry for "light" in the MCD.

Subsequent to the definition of the headword, the MCD systematically introduces collocations, organized by their grammatical category. The grammatical relationship between the headword and its corresponding collocates is indicated through part-of-speech patterns, denoted as (e.g., adv+V), although it lacks the explicit manner of labeling seen in the OCD (e.g., verb+light). Fuertes-Olivera (2011: 59) suggests that explicitly specifying grammatical labels would have been preferable for language students. From an advanced learner's perspective, the lack of explicit grammatical labels might not present a significant challenge. Nevertheless, the incorporation of such labels would contribute to the overall clarity and transparency of lexicographic entries.

The MCD establishes distinctions within the same grammatical category by grouping them into semantic groups, with each group having a definition marked by the symbol ►.

The primary code, denoted by the symbol ●, signifies the grammatical relationship between the headword and its collocates, specifically adv + V in Figure 4. This code indicates that the collocations consist of an adverb followed by a verb. Within this structure, the MCD presents four distinct semantic groupings. The first grouping revolves around situations characterized by limited light, featuring collocations like “dimly light.” The second grouping delves into the concept of well-lit environments, incorporating collocates such as “properly light” and “well light,” alongside additional semantic groupings.

A noteworthy distinction between the MCD and the OCD lies in their treatment of prepositions. In contrast to the OCD’s practice of segregating prepositions into a dedicated section, the MCD seamlessly integrates prepositions within the constructions and examples. An illustrative instance is found in the entry for the noun “light,” where a grammatical code, such as n+of+N, is provided, as illustrated in Figure 5.



- light** N  
 1 brightness from the sun, a fire, a piece of equipment etc
- adj+N bright **bright, brilliant, good, intense, shimmering, strong** *The light became brighter and I could distinctly see three shapes.*
  - not bright **bad, dim, fading, faint, poor, weak** *I couldn't really see his expression in the dim light.*
  - natural **natural** *The dining room has no natural light.*
  - artificial **artificial, electric** *In 1895, Chester's streets were lit for the first time with electric light.*
  - having a particular colour **golden, white, yellow** *The stage was flooded with bright white light.*
  - v+N provide light **cast, emit, give, give out, provide** *An intensely white light was emitted by the object.*
  - reflect light **reflect** *The two sides are white, to reflect the light.*
  - N+v be produced **gleam, glow, shine** *On my way upstairs, I saw light shining through a chink in the door of a room which I knew to be unoccupied.*
  - make something bright **illuminate** *The scene was illuminated by the midday light.*
  - become less bright **fade, fail, go** *By the time I'd got there, the light was already fading fast.*
  - shine from or in a particular direction **come from, emanate from, fall, stream** *There's also a lot of light coming from above.*
  - fill a place **fill, flood** *The door suddenly opened and light flooded the scene.*
  - make someone temporarily blind **blind, dazzle** *She was blinded by the light.*
  - n+of+N line of light **beam, ray, shaft, streak** *Rays of light broke through chinks in the clouds and lit up the hills and lakes.*
  - amount of light that appears **blaze, burst, chink, flash, gleam, glimmer, patch, pool, pulse** *There was a flash of bright light, and thunder rumbled in the distance.*
  - quality **brightness, glare, glow, intensity** *The warm glow of the lights made the room homely and inviting.*

Figure 5: Entry for "light" in the MCD.

The MCD incorporates a dedicated section for usage notes, which offers valuable information including:

(i) Colligation, which refers to the tendency of a word to appear in a specific form (Siepmann 2005). This could include verbs predominantly used in their passive form or nouns primarily used in their plural form, among other examples.

(ii) Alternatives to collocations, which are alternative phrases or expressions commonly used instead of the given collocation.

However, the notable distinction lies in the type of collocational information offered. Specifically, the MCD breaks new ground by including noun collocates for adjectives and

verb entries, which is not done in the OCD. This feature sets the MCD apart as a highly valuable and distinctive English collocations dictionary in comparison to others.

#### **4. Comparative Analysis of the Two Collocations Dictionaries Studied**

In this section, a comparative assessment of the two mentioned resources will be presented, with regard to the types of collocations they encode; the nature of collocational information presented; and the placement of collocations within the micro or macrostructure of the dictionary.

The OCD employs a singular approach for accessing collocations, primarily through the base term. Conversely, a notable advantage of the MCD is its inclusive approach, encompassing nouns and adjectives within verb entries. This unique feature enables users to access collocations through either the base term or the collocate.

In terms of collocation classification within entries, both dictionaries offer classification based on syntactic schema, albeit with variations in organization and formulation. The OCD stands out by explicitly specifying the grammatical category and order of the two components, while the MCD also provides this information, albeit without indicating the headword.

While both the OCD and MCD assert their ability to differentiate meanings within the same grammatical category, their efforts fall somewhat short. The relationship between various components in the group is not clearly expressed, and inconsistencies exist in the members of each group.

Regarding the description of collocations, both resources offer explanations of collocational meanings, supplemented by usage notes and examples when necessary.

Table 1 provides a summary of the analysis conducted in this paper, outlining the strengths and weaknesses of each resource:



	Access	Classification	Description	Advantages	Disadvantages
<b>OCD</b>	Base term	-Syntactic -Semantic	-Explanations when needed -Examples of use	-Distinct section dedicated to prepositions and phrases -Attempt to categorize collocations based on their meaning	-Only access by the base term -Paper format -Not theoretically based
<b>MCD</b>	-Base term -Collocate	-Syntactic -Semantic	-Usage notes -Examples of use	-Designed for both encoding and decoding purposes -It includes noun collocates for adjective and verb entries -Attempt to classify collocations based on their meaning -Detailed and precise description of word combinations	-Only access by the base term -Paper format -Fewer entries

Table 1: Comparative analysis of the OCD and the MCD.

## 5. Conclusion

Thanks to the availability of extensive corpora and lexical analysis tools, a wide array of lexicographic resources have emerged and evolved. Among these resources are collocations dictionaries, which cater to various user groups, including language learners, linguists, teachers, and more. However, as observed in the analysis, these resources differ in their representation and organization of phraseological information within their entries. This variation highlights the absence of a consensus regarding the types of word combinations that should be included in dictionaries and how they should be described and classified.

The comparative analysis conducted in this paper on two notable monolingual collocations dictionaries in English yielded the following conclusions:

- A collocations dictionary should offer various methods for accessing collocations, enabling efficient retrieval of relevant meaning-related information.
- A collocations dictionary should classify collocations within each entry, providing learners and other users with a quick way to find the desired information. This classification can be based on one of the following: (i) part of speech, (ii) syntactic schema between the noun and the verb, or (iii) meaning.
- A collocations dictionary should provide descriptive information about collocations, aiding users in understanding the meaning of specific collocations.

- Dictionary entries should include usage notes and examples to provide contextual information. Such information is highly valuable as it demonstrates how words behave in specific communicative settings and illustrates the practical usage of collocations in real language.

Despite their merits, there are considerations to address regarding the use of these dictionaries by language learners. Firstly, awareness among learners about the existence and utility of such dictionaries may vary. Secondly, the necessity of teaching learners how to effectively utilize these dictionaries warrants attention, particularly regarding accessing, classifying, and interpreting collocational information.

Moreover, while both the Oxford Collocations Dictionary (OCD) and the Macmillan Collocations Dictionary (MCD) have their merits, the choice between them depends on learners' preferences and learning objectives. Therefore, integrating these dictionaries into language learning curricula and providing guidance on their use could enhance learners' proficiency in English collocations.

## 6. References

- Bergenholtz, H., & Tarp, S. (2010). LSP lexicography or terminography? The lexicographer's point of view. *Specialized dictionaries for learners*, 27-37.
- Fuertes-Olivera, P. A. (eds.). (2010). *Specialised dictionaries for learners*. de Gruyter, 33
- Görlach, M. Crowther, J., Dignen, S., & Lea, D.(eds.). (2002). Oxford Collocations Dictionary for Students of English. Oxford: Oxford University Press. ISBN 0-19-431-2437.
- Kilgariff, A., Kovář, V., Krek, S., Srdanović, I., & Tiberius, C. (2010). A Quantitative Evaluation of Word Sketches. In *Proceedings of the XIV Euralex international Congress* (pp. 372-379).
- McGee, I. (2012). Collocations dictionaries as Inductive Learning Resources in Data-Driven Learning—An Analysis and Evaluation. *International Journal of Lexicography*, 25(3), 319–361.
- Nuccorini, S. (2003). Towards an 'Ideal' Dictionary of English Collocations. *A Practical Guide to Lexicography*, 6, 366–387.
- Rundell, M., & Fox, G. (eds.). (2010). *Macmillan Collocations Dictionary*. London: Macmillan Education.
- Siepmann, D. (2005). Collocation, colligation and encoding dictionaries. Part I: Lexicological aspects. *International Journal of Lexicography*, 18(4), 409-443.





