

The mapping of Indonesian collocations: A contrastive analysis with English collocations

Sri Dewiyanti¹, Sri Nur Yuliyawati^{2*}, Saudin¹, and Lina Meilinda¹

¹English Department, Politeknik Negeri Bandung, Gegerkalong Hilir, Bandung 40559, Indonesia

²Civil Engineering Department, Politeknik Negeri Bandung, Gegerkalong Hilir, Bandung 40559, Indonesia

ABSTRACT

Recent years have witnessed a growing interest in research on English collocations, yet studies focusing on Indonesian collocations remain underexplored. Collocations are crucial for language acquisition as proficiency in collocations is indicative of language competence. This study, therefore, intends to reveal and map the patterns of Indonesian lexical and grammatical collocations through a contrastive analysis with English collocation patterns. The research was primarily qualitative, utilizing the phraseological approach supported by a corpus for collocation identification. The word classes of the collocation patterns were determined using KBBI V (the unabridged Indonesian dictionary). Data were sourced from main reports of *Tempo* magazine and headlines of *Kompas* as well as *Pikiran Rakyat* newspaper, comprising a total of 14,206 words. The data yielded a total of 1,488 instances of Indonesian collocations, divided into lexical collocation with 994 instances (66.8%) and grammatical collocation with 494 instances (33.2%). The investigation reveals nine sub-types of Indonesian lexical collocations, seven of which correspond to English lexical collocation sub-types while the other two are Indonesian-specific patterns. With respect to grammatical collocation, Indonesian possesses 14 sub-types, with eight aligning with English patterns and six unique to Indonesian. Future studies are required to delve more into and shed light on Indonesian collocations.

Keywords: Collocations; contrastive analysis; corpus; Indonesian collocations; phraseological approach

Received:

16 April 2024

Accepted:

6 September 2024

Revised:

13 August 2024

Published:

30 September 2024

How to cite (in APA style):

Dewiyanti, S., Yuliyawati, S. N., Saudin S., & Meilinda, L. (2024). The mapping of Indonesian collocations: A contrastive analysis with English collocations. *Indonesian Journal of Applied Linguistics*, 14(2), 349-360. <https://doi.org/10.17509/ijal.v14i2.74908>

INTRODUCTION

The significance of collocation, defined as frequently co-occurring word combinations with a high strength of association (Garibyan et al., 2022), has been extensively recognized in the literature (Srdanovic, 2014, as cited in Lirong, 2022). As evidence, collocations are claimed to link to fluent and natural production associated with native speakers of the language and to be prominent in the understanding of language learning and use (Gablasova et al., 2017). In a similar vein, Umair (2023) highlighted the critical role of collocations in language acquisition, emphasizing their contribution to rapid and fluent language production, enabling speakers to more effectively articulate ideas and convey meanings. Furthermore, collocations have

been established as a key indicator of language fluency and proficiency (Basal, 2019; Bui, 2021; Lateh et al., 2021). Due to their prominent importance, collocations have been investigated in a myriad of studies with a variety of foci, several of which are discussed below.

One central focus of previous studies is on collocation knowledge, which has been positively linked to the four language skills and translation quality. Hsu (2005, as cited in Bui, 2021) demonstrated that learners' listening performance significantly improves following instruction on collocations. Similarly, students who are made aware of formulaic sequences, such as collocations and idioms, demonstrate enhanced oral proficiency

*Corresponding author
Email: srinuryuli@polban.ac.id

(Boers, 2006, as cited in Estaji & Montazeri, 2022). Collocation knowledge also aids in reading comprehension by allowing learners to grasp the meaning of a passage without needing to focus on every individual word (Hunston & Francis, 2000, as cited in Vaupot, 2021). Finally, it is positively associated with writing performance (Abdi & Ariffin, 2020) and is closely tied to translation quality or naturalness as collocations are quite likely to constitute a valid measure to make a distinction between translational language and native-speaker language (Feng, 2020).

Recent research on collocation knowledge has currently paid particular attention to the productive collocation use in learners' corpus, though such studies differ remarkably in the way they define and identify collocations (Du et al., 2022). Some have defined collocations (e.g. *break a promise*) by delimiting them from free combinations (e.g. *break a vase, break a cup*, etc.) and idioms (e.g. *break a leg*), based on the degree of semantic transparency and commutability. These studies largely employ the phraseological approach, relying on native speakers' judgments, to identify collocations. In contrast, other studies have pertained collocation to the statistical probability of word co-occurrence, using various parameters such as minimal frequency thresholds and measures of association strength (e.g. MI score and t-score) based on the frequency information from a large-scale reference corpus (Du et al., 2022; Xia et al., 2023). Du et al. (2022) further mentioned that a few studies have combined the two approaches to mitigate the inconsistency inherent in human judgment within the phraseological approach and to avoid retrieving commonly occurring word combinations that lack meaningful content (e.g. *and the, by the*) in the frequency-based approach.

Another area of research focuses on identifying the most effective methods for teaching collocations. Such research is geared by the fact that many L2 English learners struggle with the appropriate use of collocations (Nguyen & Webb, 2017, as cited in Dang et al., 2022), particularly with the use of verb-noun (V+N) and adjective-noun (Adj+N), the two most often encountered types of collocations in English texts (Cao & Deignan, 2019). Mahvelati (2019) reviewed the pedagogical potential or effectiveness of implicit versus explicit collocation teaching methods. The author suggested that the reported findings regarding the most effective approach have been inconclusive, partly due to the lack of consideration for factors such as individual learner differences. Nonetheless, the researcher underscored the importance of drawing learners' attention to input (e.g. the target collocations) to potentially transform it into intake. Further, Fazlali and Shahini (2019) postulated that input enhancement techniques, which fall under implicit instruction and involve typographically

highlighting target collocations (e.g., through bolding or underlining), have positive effects on learners' retention of collocations. Meanwhile, Snoder & Reynolds (2019) affirmed that the SEM dictogloss technique, which involves the semantic elaboration of target collocations during the pre-task phase, is more effective in promoting collocation learning than the STRUC dictogloss technique, which focuses on structural elaboration.

Previous Studies on Indonesian Collocations

In contrast with the surge of research on English collocations, studies on Indonesian collocations remain limited, with few publications on the topic. A review of several journals on Indonesian linguistics, language education, literature, and arts (*Aksara, Dialektika, Linguistik Indonesia, Ranah, and Wacana*) revealed that none have published articles on Indonesian collocations in the past five years (2019-2023). Consequently, the distinctive patterns and unique features of Indonesian collocations have not been overtly unveiled. This scarcity is further proven by the very limited number of studies available online. A search for journal articles on Indonesian collocations using the Google search engine yielded only a small number of results, despite spanning a period of 13 years (2009-2022).

Artha's (2018) study on COCA (Corpus of Contemporary American English) and COHA (Corpus of Historical American English) investigated various adjectives frequently attached to the word "Indonesia" across three distinct periods of Indonesian governance: The Old Order, The New Order, and The Reform Era. The study discovered that the adjectives following the word "Indonesia" changed every period, informed by the background of its respective situational contexts and historical happenings. Setyaningsih et al. (2021) conducted a comparative analysis of keywords in collocations related to COVID-19 lexicons in Indonesian and English, using AntConc software to analyze data from cyberspace. The results revealed that both English and Indonesian collocations share similar thematic content, as indicated by the keywords identified. However, the differences lie in the patterns of collocation, with Indonesian exhibiting six distinct patterns compared to four in English.

Furthermore, Lirong (2022) conducted a comparative analysis of the nouns that serve as bases for the synonymous verb collocates "*menyebabkan*" and "*mengakibatkan*", utilizing data from the Indonesian Web Corpus (WaC). The study discovered that the verbs "*menyebabkan*" and "*mengakibatkan*" usually form collocations with similar nouns, and both carry negative semantic prosody. The verb "*menyebabkan*", however, tends to be juxtaposed with nouns related to social and geographical conditions, whereas "*mengakibatkan*" is more commonly associated with those linked to

physical health conditions. In a separate study, Prayogi et al. (2022) compared verb collocates commonly combined with the temporal concept (i.e. *time/waktu*) in both English and Indonesian, utilizing data from two reference corpora: the Corpus of Contemporary American English for English data and Leipzig Corpus Collection for Indonesian. The Mutual Information (MI) measure was applied to identify verb collocates. The study identified 15 most frequent verb collocates of “*time*” and “*waktu*” in both languages, e.g. *have, spend, take, membutuhkan, memanfaatkan, memerlukan*, etc. The study also mapped the similarities and differences between these collocates. Meanwhile, Imran et al. (2009) explored types of Indonesian collocations, categorizing them into lexical and grammatical collocations. The discussion intended to discover commonly co-occurring Indonesian lexical items and describe the patterns and meanings of these collocations. The results revealed nine types of lexical collocations with 354 instances (93%) and eight types of grammatical collocations with 27 instances (7%). The study also confirmed that 75% of the collocations have transparent meanings while 25% have specific, unique meanings.

As highlighted by previous research, the study of Indonesian collocations remains relatively underexplored. This current study, therefore, seeks to address this significant gap by mapping the profile of Indonesian collocation patterns using English collocation patterns as a reference. It specifically intends to demonstrate the similarities and differences of the collocation types in both languages and present the frequency of occurrences of various Indonesian collocation types, successively from the least to the most encountered. By pursuing these objectives, the research is expected to stimulate further research on Indonesian collocations and prompt the inclusion of collocations in the curriculum of Indonesian learning. In addressing the purposes, the theory of English collocations proposed by Benson et al. (2009) was utilized as an underpinning one since it is regarded to offer more comprehensive categories of collocation (Mroczyńska, 2023). This framework, which distinguishes between lexical collocations (combinations of content words such as verbs, nouns, adjectives, and adverbs) and grammatical collocations (combinations of a content word with a function word, typically a preposition), has been widely adopted in a great deal of research.

METHOD

The research employed a descriptive-analytical design, in which documents were analyzed to identify occurrences of collocations as a linguistic phenomenon. This design is part of qualitative research methodology, allowing for in-depth document analysis to obtain the desired data. The

study also incorporated a contrastive analysis as it compared the types of collocations found in Indonesian with those in English. Contrastive analysis is a theory that posits all languages, despite their diversity, share fundamental similarities (Dąbrowska, 2015). To put it differently, probing into a certain feature in the system of one language will result more fruitfully by comparing it with the parallel feature in another language.

The data was obtained from news articles (magazine main reports and newspaper headlines) discussing social, political, legal, and economic issues. The articles were selected from *Tempo* magazine (December 18 – 24, 2022; February 13 – 19, 2023; February 20 – 26, 2023; and February 27 – March 5, 2023), *Kompas* newspaper (January 18, 2023; March 7, 16, 23, 25, and 28, 2023), and *Pikiran Rakyat* newspaper (February 9 and 16, 2023), forming a total of 14,206 words. The selection criteria included the diversity of topics covered and the variety of media sources to ensure a rich collection of collocation patterns. Another was the avoidance of the same article writers to minimize the possibility of individual idiosyncrasies.

The process of extracting Indonesian collocations from the documents involved several steps. The documents were first converted into Word format to facilitate a more efficient extraction process. They were then split into four parts, with each section assigned to one of the four research authors for analysis. Prior to the process of extraction, the authors held four discussions to align their perspectives and establish a guideline for determining whether a word combination qualified as a collocation, based on their judgments as native speakers. Once identified, the instances of Indonesian collocations were highlighted in the documents using different colors to distinguish each major type of collocations. The identified collocations were then tabulated and cross-checked to find any potentially questionable instances. Subsequently, an author was tasked with verifying the validity of these dubious collocations by consulting an online corpus, the Indonesian Web Corpus (Indonesian WaC). Finally, the authors integrated the four tabulations into a single table and conducted a thorough review of all collocations to further reduce the likelihood of inaccurate identifications.

The phraseological approach, which treats collocations as a category situated on a continuum between free combinations at one end and idiomatic expressions at the other (Mroczyńska, 2023), was utilized to extract collocations from the documents. Instead of “*mendirikan*” (*erect*), for example, the verb “*menegakkan*” (*uphold*) paired with the noun “*hukum*” (*laws*) or “*keadilan*” (*justice*) will form a collocation. By contrast, the combination of “*menegakkan kepala*” (*put upright the head*) or

“*menegakkan tiang*” (*put upright a pole*) is considered a free combination, and “*menegakkan benang basah*” (*put upright wet thread*), meaning attempting the impossible, is an Indonesian idiom – neither of which qualify as collocations. This study used the phraseological approach to ensure accurate identification of certain collocation patterns since the combinations of Adj+To Infinitive and N+To Infinitive, for example, can be challenging. The juxtaposition of lexical items in such collocation patterns (e.g. *necessary to work* or *a promise to come*) is often merely structural, with the frequency of occurrence and the strength of association between the words being relatively low.

To address the potential inconsistencies in human judgment inherent in the phraseological approach, the frequency-based approach was also applied, albeit to a lesser extent. For this, Sketch Word of the Indonesian Web Corpus (Indonesian WaC), an online corpus, was utilized to verify whether some pre-selected perplexing word combinations (nearly 200 instances) were well-formed collocations or merely free combinations/malformed collocations. A threshold of two-digit frequency of hits was adopted to determine the collocability of these combinations. Combinations as *melansir data*, *vonis berat*, *ketidakpastian kondisi*, *komposisi kepemimpinan*,

and *kesadaran soal* turned out to be the latter, with hits count being lower than two digits (1, 2, 5, 7, and 6, respectively). Further, certain combinations in Indonesian, such as **di tahap (pada tahap)*, and **di usia (pada usia)*, were also excluded due to the unstandardized uses of prepositions in them. Approximately 50% of the instances were not added to the total number of Indonesian collocations found. Moreover, the classification of word classes in these combinations was determined using the KBBI V (2023), an online Indonesian dictionary, particularly because word classes, especially adverbs, remain a challenging area in Indonesian linguistics, even for scholars (Wijana, 2022). One last point worth informing is a book discussing collocation and uncovering the lists of Indonesian collocation types has been published as the outcome of this research.

FINDINGS

To achieve its objectives and generate the findings, this research relied on Benson et al.'s (2009) theory of collocation. The theory, which has been slightly adapted, encompasses two primary categories of collocations—lexical and grammatical—along with their respective sub-types, as detailed in Table 1 and Table 2.

Table 1
English Lexical Collocation

No.	Pattern/Sub-Type	Example
1.	V(transitive)+N	<i>do damage, draw a conclusion, take a trip</i>
2.	Adj+N	<i>fast food, heavy rain, strong wind</i>
3.	N+of (indicating a unit) +N & *N+N	<i>a bouquet of flowers, a pride of lions comfort zone, human resources, traffic accident</i>
4.	N+V (used predicatively)	<i>a baby cries, a dog barks, doctors diagnose</i>
5.	Adv+Adj	<i>extremely generous, strikingly different, truly mad</i>
6.	V+Adv /Adv+V	<i>argue heatedly, rain heavily, flatly deny, fiercely defend</i>
7.	**V+Adj	<i>go blind, keep clean, stay cool</i>

*This N+N pattern is subsumed under lexical collocation by many scholars (Khdhir et al., 2022; Olimat et al., 2022; Yamagata et al., 2023) though the pattern is not listed in Benson et al. (2009).

**This pattern is also regarded as a type of lexical collocation, known as linking verb collocation, by some scholars (Mahvelati & Mukundan, 2012; McIntosh & Francis, 2012; Saudin et al., 2017).

Table 2
English Grammatical Collocation

No.	Pattern/Sub-Type	Example
1.	N+Prep	<i>an attack on, an error in, a need for</i>
2.	Prep+N	<i>at night, by accident, in despair</i>
3.	V+Prep	<i>depend on, dry up, look after, turn off</i>
4.	Adj+Prep	<i>aware of, close to, dependent on, familiar with</i>
5.	N+To Infinitive	<i>a pleasure to meet you, an attempt to investigate it</i>
6.	Adj+To Infinitive	<i>glad to see her, necessary to work</i>
7.	*N+That Clause	<i>a promise that she would come, an agreement that he would help</i>
8.	*Adj+That Clause	<i>afraid that she would fail, happy that he was able to pay a visit</i>
9.	V+Obj+Prep	<i>blame someone for, send the package to, sing a song to</i>
10.	**Binomials	<i>dead or alive, head or tail, safe and sound</i>

*How these patterns are likely realized in Indonesian is not discussed in the current research, considering the data would be too massive to analyze; more importantly, the study sticks to the definition of collocation as a fixed co-occurrence of words, not beyond words into a clause.

** Binomials – though not listed in Benson et al.'s (2009) theory of collocation – are regarded as a type of collocation (Al-Otaibi, 2023) and defined as two words joined with conjunction and the order is fixed, rarely reversed (Boonnoon, 2020; Du et al., 2021).

The analysis of the documents, comprising 14,206 words, yielded a total of 1,488 instances of Indonesian collocations. Each collocation was counted only once, regardless of how many times it appeared in the data. Of these, 994 instances (66.8%) were identified as lexical collocations and 494 instances (33.2%) as grammatical collocations. The study identified nine patterns or types of Indonesian lexical collocations, seven of which align with English lexical collocation types, while two are unique to Indonesian. The total number of patterns observed in this study matches that found in Imran et al.'s (2009) earlier research. However, this current study did not detect the N+Adv collocation reported in the previous study, but it discovered the Adj+Adj collocation type that was not found in the earlier research. The frequency distribution of the nine types of Indonesian lexical collocations is as follows: N+N (f. 415 / 41.8%), V+N (f. 322 / 32.4%), N+Adj (f. 149 / 15%), V+Adv / Adv+V (f. 52 / 5.2%), V+Adj (f. 22 / 2.2%), N+V (f. 17 / 1.7%), Adj+N (f. 10 / 1%), Adv+Adj / Adj+Adv (f. 4 / 0.4%), and Adj+Adj (f. 3 / 0.3%).

In terms of grammatical collocations, Indonesian exhibits 14 distinct patterns. Eight of these correspond to English collocation patterns, while six are unique to Indonesian. This total surpasses the number identified in Imran et al.'s (2009) study, which reported only nine patterns, including Adj+That Clause and N+That Clause collocations that were not included in the current study. The frequency distribution of Indonesian grammatical collocation patterns, from the most to the least frequent, is as follows: N+Prep (f. 100 / 20.2%), Prep+N (f. 99 / 20%), V+Prep (f. 83 / 16.8%), N+Vused attributively / N+To-Infinitive (f. 79 / 16%), V+Obj+Prep (f. 38 / 7.7%), Adj+Vused attributively / Adj+To-Infinitive (f. 36 / 7.3%), Adj+Prep (f. 20 / 4.1%), Prep+Adj (f. 15 / 3.1%), binomials (f. 10 / 2%), Prep+N+N (f. 5 / 1%), Prep-Prep (f. 4 / 0.8%), Prep+V (f. 2 / 0.4%), Adv-Prep (f. 2 / 0.4%), and Adv-Adv (f. 1 / 0.2%).

To be more apparent, the key findings of this research are demonstrated in Appendix A (lexical collocations) and Appendix B (grammatical collocations).

DISCUSSION

In relation to Indonesian lexical collocation patterns presented in Table 3, this study, to the best of the authors' knowledge, is the first to reveal the frequency profile of these occurrences. As such, the findings cannot be directly compared with those of other studies using Indonesian data. Nevertheless, existing research indicates that the most frequently encountered patterns in English data are V+N and Adj+N collocations, either successively (Cao et al., 2022; Sarjono et al., 2022) or inversely (Estaji & Montazeri, 2022). Meanwhile, in English data, the

N+N collocation pattern is typically the third most common type (Sarjono et al., 2022). Therefore, the current research's finding that the N+N pattern is the most encountered (instead of V+N or Adj+N) is conflicting with the common results of other studies investigating English collocations.

The finding, however, indicates something worth noting, that in Indonesian collocation patterns, nouns play a more central role than verbs or adjectives. According to Hausmann (1984, as cited in Adamcova, 2020), nouns are fundamental as they denote the entities and phenomena about which information is conveyed; verbs and adjectives can only be used as basic words insofar as they can be further determined by adverbs. Likewise, Langacker (1987, as cited in Tran & Waluyo, 2021) emphasized that nouns have historically been a primary word type and are essential in grammatical structures alongside verbs. The prevalence of the N+N collocation pattern in Indonesian suggests that the language emphasizes communication about concepts and phenomena rather than processes and states. Simply put, this current research argues that focusing solely on V+N and Adj+N collocations as major types is insufficient, and the significance of N+N collocations must be recognized. This finding offers a novel perspective on collocation research, though further validation from future studies is needed.

Among the nine Indonesian lexical collocation patterns identified, the last two listed in Table 3 are distinctive to Indonesian. They are Adj+N and Adj+Adj patterns. The Adj+N pattern does not occur in English collocations, where it is typically reversed to N+Adj. In Indonesian, idioms (though not regarded as collocations in the current research) are also known to be frequently formed with the Adj+N pattern, such as “*besar hati*” (*optimistic*) and “*patah semangat*” (*giving up*). In relation to Adj+Adj collocation, this sub-type is often used in daily communication and usually involves two synonyms, such as “*cantik (beautiful)+molek/jelita*” or “*lemah (soft)+lembut/gemulai*”. In this case, the collocate (e.g. *molek* or *jelita*) intensifies or emphasizes the meaning of the base (e.g. *cantik*). This collocation pattern is not restricted to Indonesian, German, for instance, also features a similar pattern, as in *peinlich genau* (Adamcova, 2020). Additionally, two Indonesian languages, Balinese and Javanese, also exhibit this collocation pattern, such as *peteng dedet* and *sepi jampi* in Balinese and *adem ayem* and *teles kebes* in Javanese.

With respect to the findings of Indonesian grammatical collocation sub-types in Table 4, to the best of the authors' knowledge, this study appears to be the first to disclose their frequency profile of occurrences. As such, the results cannot be directly compared with those of previous studies on Indonesian collocations. However, when compared

to research on English collocations, some similarities and differences emerge. The sub-types most frequently encountered in Indonesian, such as N+Prep and Prep+N, align with findings in English texts, as reported by Sarjono et al. (2022). Similarly, Thongvitit and Thumawongsa (2017) posited that N+Prep collocation is the most frequently used, and V+Prep collocation is the second. By contrast, Estaji and Montazeri (2022) argued that V+Prep collocation is mostly used.

Further, the identification of 14 sub-types of Indonesian grammatical collocations (with the first eight aligning with the primary sub-types of English grammatical collocations and the remaining six being unique to Indonesian) does not imply that Indonesian has a greater number of grammatical collocation sub-types than English. Apart from the two sub-types (N+That Clause and Adj+That Clause), this study did not examine more than 15 additional sub-types in English, particularly those involving verb patterns. Still, the Indonesian grammatical collocation patterns displayed in Table 4 exhibit considerable diversity.

Special attention should be given to Indonesian grammatical collocation sub-types #5 and #6 shown in Table 4. Though the word patterns in these sub-types consist of content words, they do not fall under the category of lexical collocation. The reason is the verbs in the patterns function attributively and can be preceded by the preposition *untuk* (e.g. *keputusan bertarung=keputusan untuk bertarung*) or by the relative pronoun *yang* (e.g. *duka mendalam=duka yang mendalam*) to modify the noun or the adjective to which they are attached. An intriguing phenomenon, however, is that some verbs used in this way cannot be preceded by *untuk* or *yang* (e.g. *juru bicara* cannot be **juru untuk bicara* or **juru yang bicara*; and *titik temu* is unlikely to be said **titik untuk temu* or **titik yang temu*). This restriction likely arises because these word combinations are compound words, preventing any insertion between them.

Regarding the unique sub-types of Indonesian grammatical collocations, three patterns—Prep+Verb, Prep+Adj, and Prep+N+N—have been classified as distinct sub-types. This consideration is due to the frequent occurrence of these patterns in Indonesian texts. In addition to the examples presented in Table 4, it is frequent to find such realizations as *dengan berhati-hati*, *secara berterima*, *dengan menahan amarah* (Prep+V); *dengan damai*, *dengan sukacita*, *secara umum* (Prep+Adj); and *di bawah umur*, *di luar nikah*, *di seberang lautan* (Prep+N+N). Generally, the prepositions *secara* and *dengan* can be used interchangeably as they are synonymous. However, in certain contexts, their interchangeability is restricted or may sound unnatural, as in **secara menahan amarah* and **dengan umum*. More importantly, the examples of these three patterns

regarded as collocations in the study exhibit significant frequency, with at least two-digit occurrences. For instance, *secara insidental* appears 19 times and *dengan baik* occurs 17,371 times, confirming their status as well-established collocations. Additionally, word combinations *di bawah kepemimpinan* and *di luar kewajaran* should be undoubtedly included in Indonesian collocation sub-types, as evidenced by their high frequency of co-occurrence in the reference corpus used (Indonesian WaC), with 374 occurrences for the former and 19 for the latter.

Concerning the last three unique patterns of Indonesian grammatical collocation (collocation patterns as the correlative conjunction), their collocability is clearly apparent. It is semantically obvious that the use of one lexical component in correlative conjunction naturally prompts the emergence of its paired counterpart. For this reason, this study proposes that Indonesian correlative conjunctions, which involve combinations of both functional/grammatical words such as prepositions and conjunctions/adverbs, need to be included in Indonesian collocations. As shown in Table 4, patterns #12 to #14 include three such collocation patterns: Adv+Adv, Adv+Prep, and Prep+Prep.

CONCLUSION

The mapping of Indonesian lexical and grammatical collocations revealed in this study should not be accepted without further scrutiny. Though considerable effort has been invested in this research, further studies are required to evaluate (confirm and disconfirm) and/or expand upon the findings. Such efforts will contribute to a more comprehensive understanding of Indonesian collocations and support the ongoing development of the national language. This study also highlights the importance of collocations and advocates for their inclusion in the Indonesian language teaching curriculum, and unquestionably in the teaching of translation as well. Currently, widely used textbooks on Indonesian grammar, such as Moeliono (2017), lack discussions on collocations. Furthermore, KBI (Kamus Besar Bahasa Indonesia) still cannot fully serve as a reliable reference, particularly to Indonesian collocation formations. This is due to its lack of syntactical illustrations to depict how a lexical entry commonly co-occurs with other words, and to its inexhaustive categories of word classes for its entries. For instance, there is no exemplification whether it is with the preposition *bagi*, *pada*, or *terhadap* the verb *berdampak* commonly and appropriately co-occur. It is also questionable if the lexical items *keras* and *relatif* are only registered as an adjective, considering that the items in the combinations like *berusaha keras* and *relatif aman* modify a verb (*berusaha*) and an adjective (*aman*). Thus, they function as an adverb as well. While

Indonesian students, as native speakers, possess an intuitive understanding of collocations, this intuition only partially informs them of commonly co-occurring words that sound natural together. Hence, collocation competence should be explicitly taught and learned as it involves the understanding of standard structures in combining words and subsequently in constructing sentences that accurately convey meaning.

To sum up, the field of Indonesian collocation research offers abundant opportunities for further exploration, with many intriguing and important topics available. Continued efforts to map the profile of Indonesian collocations remain in need, despite the time-consuming nature of such research, as evidenced by the current study. Though it has attempted to encompass all possible patterns, the research has at least one important limitation: certain Indonesian-specific collocation patterns remain unveiled. In particular, Indonesian verb patterns, which are potentially subsumed under collocations, have not been thoroughly investigated. These verb patterns, the structures and/or the semantic relations of how a verb is used with another, present a challenging area of study. Finally, an extremely demanding topic worth exploring is lexical bundles—fixed sequences of three to four words that, like collocations, are a form of formulaic language. Lexical bundles, oftentimes termed differently such as prefabricated language, are fixed sequences of three to four words and are described as building blocks of discourse (Johnston, 2000). To the best of the authors' knowledge, Indonesian lexical bundles have only been discussed in one article by Budiwiyanto and Suhardijanto (2020).

REFERENCES

- Abdi, B. M. & Ariffin, A. (2020). The relationship between collocation competence and writing skills of EFL learners. *The Asian Journal of English Language & Pedagogy*, 8(1), 41–52.
- Adamcova, S. (2020). Empirical research of collocations in foreign language learning. *Advanced Education*, 7(14), 75–83. <https://doi.org/10.20535/2410-8286.197363>
- Al-Otaibi, G. M. (2023). Explicitation in the translation of Qur'ānic binomials: A descriptive study. *Indonesian Journal of Applied Linguistics*, 13(2), 293–305. <https://doi.org/10.17509/ijal.v13i2.63098>
- Artha, A. F. (2018). Revolusi pemerintahan, sudahkah berevolusi? Kolokasi adjektiva kata “Indonesia” dalam COCA dan COHA pada periode pemerintahan Orde Lama, Orde Baru dan Era Reformasi. *Etmolingual*, 2(1), 55–71. <https://doi.org/10.20473/etno.v2i1.8441>
- Basal, A. (2019). Learning collocations: Effects of online tools on teaching English adjective-noun collocations. *British Journal of Educational Technology*, 50(1), 342–356. <https://doi.org/10.1111/bjet.12562>
- Benson, M., Benson, E., & Ilson, R. (2009). *The BBI combinatory dictionary of English* (Rev. 3rd ed.). John Benjamins.
- Boonnoon, S. (2020). Exploring Thai EFL students' knowledge of English binomials. *English Language Teaching*, 13(2), 48. <https://doi.org/10.5539/elt.v13n2p48>
- Budiwiyanto, A., & Suhardijanto, T. (2020). Indonesian lexical bundles in research articles: Frequency, structure, and function. *Indonesian Journal of Applied Linguistics*, 10(2), 292–303. <https://doi.org/10.17509/ijal.v10i2.28592>
- Bui, L. T. (2021). The role of collocations in the English teaching and learning. *International Journal of TESOL & Education*, 1(2), 99–109.
- Cao, D., & Deignan, A. (2019). Using an online collocation dictionary to support learners' L2 writing. In C. Wright, L. Harvey, & J. Simpson (Eds.), *Voices and Practices in Applied Linguistics: Diversifying a Discipline* (pp. 233–249). White Rose University Press. <https://doi.org/10.22599/BAAL1.n>. Licence: CC BYNC 4.0
- Cao, D. T. P., Pho, P. D., & Dang, N. A. C. (2022). The treatment of collocations in English textbooks for Vietnamese students. *TEFLIN Journal - A Publication on the Teaching and Learning of English*, 33(2), 220. <https://doi.org/10.15639/teflinjournal.v33i2/220-239>
- Dąbrowska, E. (2015). What exactly is Universal Grammar, and has anyone seen it? *Frontiers in Psychology*, 6. <https://doi.org/10.3389/fpsyg.2015.00852>
- Dang, T. N. Y., Lu, C., & Webb, S. (2022). Incidental learning of single words and collocations through viewing an academic lecture. *Studies in Second Language Acquisition*, 44(3), 708–736. <https://doi.org/10.1017/S0272263121000474>
- Du, L., Elgort, I., & Siyanova-Chanturia, A. (2021). Cross-language influences in the processing of multiword expressions: From a first language to second and back. *Frontiers in Psychology*, 12, 666520. <https://doi.org/10.3389/fpsyg.2021.666520>
- Du, X., Afzaal, M., & Al Fadda, H. (2022). Collocation use in EFL learners' writing across multiple language proficiencies: A corpus-driven study. *Frontiers in Psychology*, 13, 752134. <https://doi.org/10.3389/fpsyg.2022.752134>
- Estaji, M., & Montazeri, M. R. (2022). The representation of collocational patterns and their differentiating power in the speaking performance of Iranian IELTS test-takers. *Frontiers in Education*, 7, 827927. <https://doi.org/10.3389/educ.2022.827927>

- Fazlali, B. & Shahini, A. (2019). The effect of input enhancement and consciousness-raising techniques on the acquisition of lexical and grammatical collocation of Iranian EFL learners. *The Electronic Journal for English as a Second Language, 24*(2), 1–21. <https://doi.org/10.5539/ies.v5n3p247>
- Feng, H. (2020). *Form, meaning and function in collocation: A corpus study on commercial Chinese-to-English translation*. Routledge, Taylor and Francis Group.
- Gablasova, D., Brezina, V., & McEnery, T. (2017). Collocations in corpus-based language learning research: Identifying, comparing, and interpreting the evidence. *Language Learning, 67*(S1), 155–179. <https://doi.org/10.1111/lang.12225>
- Garibyan, A., Schilling, A., Boehm, C., Zankl, A., & Krauss, P. (2022). Neural correlates of linguistic collocations during continuous speech perception. *Frontiers in Psychology, 13*, 1076339. <https://doi.org/10.3389/fpsyg.2022.1076339>
- Imran, I., Said, M., & Setiari, N. L. P. (2009). Kolokasi Bahasa Indonesia. *Proceeding PESAT Universitas Gunadarma - Depok, 27–33*.
- Johnston, K. (2000). *Lexical bundles in applied linguistics and literature writing: A comparison of intermediate English learners and professionals* [Published master's thesis, Portland State University]. Pdxscholar. https://pdxscholar.library.pdx.edu/open_access_etds/3482/
- Khdhir, S. A., Hussein, S. S., Qader, F. M., Ahmed, A. B., & Mohammed, R. (2022). University level Kurdish EFL student's ability to use collocations in University of Raparin. *Canadian Journal of Language and Literature Studies, 2*(4), 89-101. <https://doi.org/10.53103/cjlls.v2i4.60>
- Lateh, N. H. M., Mohamed, A. F., Mahmud, N., Nasir, N. S. M., Hanapi, N. F., Shamsudin, S., & Raof, A. H. A. (2021). Learners' collocation use in writing: Do proficiency levels matter? *Indonesian Journal of Applied Linguistics, 11*(2), 418–426. <https://doi.org/10.17509/ijal.v11i2.31632>
- Lirong, Z. (2022). Studi berbasis korpus: Perbandingan kolokasi dan prosodi semantik sinonim bahasa Indonesia “menyebabkan” dan “mengkibatkan.” *Mabasan, 16*(1), 153–176. <https://doi.org/10.26499/mab.v16i1.517>
- Mahvelati, E. H. (2019). Explicit and implicit collocation teaching methods: Empirical research and issues. *Advances in Language and Literary Studies, 10*(3), 105–116. <http://doi.org/10.7575//aiac.all.v.10n.3p.105>
- Mahvelati, E. H., & Mukundan, J. (2012). The role of cognitive style in the collocational knowledge development of Iranian EFL learners through input flood treatment. *English Language Teaching, 5*(10), 105-117. <https://doi.org/10.5539/elt.v5n10p105>
- McIntosh, C., & Francis, B. (Eds.). (2012). *Oxford collocations dictionary: For students of English* (2nd ed.). Oxford Univ. Press.
- Moeliono, A. M. (Ed.). (2017). *Tata bahasa baku Bahasa Indonesia* (Edisi keempat). Badan Pengembangan dan Pembinaan Bahasa, Kementerian Pendidikan dan Kebudayaan.
- Mroczyńska, K. (2023). How collocations are represented and taught in selected legal English textbooks. *A Journal of English Studies, 40*(CC BY-NC-SA 4.0), 61–80. <https://doi.org/10.15290/CR.2023.40.1.03>
- Olimat, S. N., Mahadin, D., Al-Khawaldeh, N. N., & Almahasees, Z. (2022). Challenges and strategies of translating COVID-19 adjective-noun and noun-noun collocations from English into Arabic. *Eurasian Journal of Applied Linguistics, 8*(3), 120–133. <http://doi.org/10.32601/ejal.803010>
- Prayogi, I., Kesuma, T. M. J., & Winarti, D. (2022). Collocations of verb + time in English and verb + waktu in Indonesian: A corpus-driven study of time metaphors. *Ahmad Dahlan Journal of English Studies, 9*(2), 91-105. <https://doi.org/10.26555/adjes.v9i2.214>
- Sarjono, R. I. L., Heda, A. K., & Bram, B. (2022). Exploring collocations in bahasa Inggris textbook: A corpus study. *Jurnal Penelitian Humaniora, 23*(2), 84–94. <https://doi.org/10.23917/humaniora.v23i2.14799>
- Saudin, S., Sulyaningsih, I., & Meilinda, L. (2017). The investigation of productive and receptive competence in V+N and ADJ+N collocations among Indonesian EFL learners. *Indonesian Journal of Applied Linguistics, 7*(1), 189-200. <https://doi.org/10.17509/ijal.v7i1.6874>
- Setyaningsih, N., Larassati, A., & Suryaningtyas, V. W. (2021). Kolokasi lexicon COVID-19 dalam bahasa Indonesia dan bahasa Inggris di ranah cyberspace. *Humanika, 28*(2), 186–198. <https://doi.org/10.14710/humanika.v28i2.40600>
- Snoder, P., & Reynolds, B. L. (2019). How dictogloss can facilitate collocation learning in ELT. *ELT Journal, 73*(1), 41–50. <https://doi.org/10.1093/elt/ccy024>
- Thongvittit, S. & Thumawongsa, N. (2017). A corpus-based study of English collocations found in the abstracts of research articles written by Thai EFL writers. *International Journal of Social Science and Humanity, 7*(12), 751–755. <https://doi.org/10.18178/ijssh.2017.7.12.920>
- Tran, H. M., & Waluyo, B. (2021). Receptive knowledge of nouns and collocations and the

- impact on English skill performances. *GEMA Online® Journal of Language Studies*, 21(1), 124–147. <https://doi.org/10.17576/gema-2021-2101-08>
- Umair, M. H. (2023). An analysis of the use of collocations in English essay writing at undergraduate level. *Pakistan Languages and Humanities Review*, 7(1), 43-54. [https://doi.org/10.47205/plhr.2023\(7-I\)05](https://doi.org/10.47205/plhr.2023(7-I)05)
- Vaupot, S. (2021). Creating a bilingual dictionary of collocations: A learner-oriented approach. *Indonesian Journal of Applied Linguistics*, 10(3), 762-770. <https://doi.org/10.17509/ijal.v10i3.31888>
- Wijana, I. D. P. (2022). Adverb in Indonesian. *Ranah: Jurnal Kajian Bahasa*, 11(1), 26-34. <https://doi.org/10.26499/rnh.v11i1.2454>
- Xia, D., Chen, Y., & Pae, H. K. (2023). Lexical and grammatical collocations in beginning and intermediate L2 argumentative essays: A bigram study. *International Review of Applied Linguistics in Language Teaching*, 61(4), 1421–1453. <https://doi.org/10.1515/iral-2021-0188>
- Yamagata, S., Nakata, T., & Rogers, J. (2023). On effective learning of English collocations: From perspectives of distributed practice and semantic restructuring. *TESOL Journal*, e767. <https://doi.org/10.1002/tesj.767>

APPENDIX A

Realizations of Indonesian Lexical Collocations

Indonesian Lexical Collocations Similar to English Lexical Collocations				
No.	Pattern	Example	Realization	f /%
1	V+N	<i>mengatasi persoalan, menggaungkan tagar, mudik lebaran</i>	<ul style="list-style-type: none"> “Figur wakil bisa mengatasi persoalan itu,” ujarnya. Kelompok ini juga menggaungkan tagar #tibatibadipenjara. Jumlah warga yang akan mudik Lebaran tahun ini diperkirakan 123 juta orang. 	322/32.4
2	N+Adj Note: This pattern corresponds to the pattern of Adj+N collocation in English	<i>dampak positif, panen raya, titik terang</i>	<ul style="list-style-type: none"> Dampak positif hilirisasi mineral, khususnya nikel, diharapkan berkontribusi positif pula. Harga beras akan tetap naik sampai tibanya panen raya. Sumber ketidakpastian datang dari belum adanya titik terang penyelesaian konflik. 	149/15
3	N+N	<i>bunga acuan, penegakan hukum, serangkaian diskusi</i>	<ul style="list-style-type: none"> Bank sentral AS diprediksi masih akan menaikkan suku bunga acuannya. Pandangan para akademikus itu tetap dianggap sebagai bentuk kepedulian mereka dalam isu penegakan hukum. Kesepakatan ini dicapai setelah Panja BPIH melakukan serangkaian diskusi. 	415/41.8
4	N+Vused predicatively	<i>aral melintang, jaksa menuntut, rapat membahas</i>	<ul style="list-style-type: none"> Jika tak ada aral melintang, anak muda tersebut diperkirakan menghirup udara bebas pada Mei mendatang. Sebelumnya, jaksa menuntut terdakwa dengan hukuman penjara seumur hidup. Rapat itu juga membahas ketersediaan bahan pokok menghadapi Lebaran. 	17/1.7
5	Adv.+Adj. / Adj+Adv	<i>akurat secara substansial, belum tentu, *relatif aman, tetap lezat</i>	<ul style="list-style-type: none"> Ia mengakui bahwa konten tersebut belum tentu akurat secara substansial. Stok bahan pangan menjelang Lebaran relatif aman. Dengan racikan yang pas, bakso yang dihasilkan tetap lezat. 	4/0.4
6	V+Adv / Adv+V	<i>berlangsung *alot, berusaha *keras, dipersiapkan jauh-jauh hari</i>	<ul style="list-style-type: none"> Maka tidak heran jika pemilihan berlangsung alot. Indonesia berusaha keras menarik lebih banyak wisatawan. Hal lainnya adalah manajemen sengketa pemilu yang harus segera dipersiapkan jauh-jauh hari. 	52/5.2
7	V+Adj	<i>bernada negatif, jatuh miskin, merasa waswas</i>	<ul style="list-style-type: none"> Dari hampir 73 ribu <i>mention</i>, 76 persen bernada negatif tentang pengesahan KUHP. Banyak nasabah jatuh miskin akibat tergiur bunga fantastis. Tapi orang tua tetap saja merasa waswas. 	22/2.2
Unique Patterns of Indonesian Lexical Collocations				
8	Adj+N Note: This pattern corresponds to that of English collocation N+Adj (but unavailable in English)	<i>dini hari, mantan ketua, tepat waktu</i>	<ul style="list-style-type: none"> Istrinya menelepon dini hari untuk memberi tahu peristiwa tersebut. Ujar mantan ketua umum kelompok suporter klub sepak bola itu. Selama pembayaran hingga 2019, koperasi selalu membayar bunga tepat waktu. 	10/1
9	Adj+Adj	<i>keliru besar, ribut, berkepanjangan, rugi besar</i>	<ul style="list-style-type: none"> Maka keliru besar jika hukum pidana mengatur pikiran. Ia tak mau ribut berkepanjangan dengan keluarga. Ibunya tak tahu bahwa ia rugi besar. 	3/0.3
Total Occurrences of Indonesian Lexical Collocations				994/100

*relatif, keras, and alot here are considered adverbs since they modify a verb, while in KBBI they are determined as adjectives.

APPENDIX B

Realizations of Indonesian Grammatical Collocations

Indonesian Grammatical Collocations Similar to English Grammatical Collocations

No.	Pattern	Example	Realization	f / %
1	N+Prep	<i>bagian dari, dukungan pada, kelalaian dalam</i>	<ul style="list-style-type: none"> Langkah-langkah itu bisa menjadi bagian dari manajemen krisis yang harus segera dilakukan. Pemerintah Indonesia pun kembali menegaskan dukungan penuh pada Palestina. Kenaikan harga dan menipisnya stok beras di pasaran pada awal tahun 2023 lebih disebabkan kelalaian Bulog dalam menyiapkan cadangan beras serta menstabilkan harga. 	100/20.2
2	Prep+N	<i>atas nama, di tengah, pada tahun</i>	<ul style="list-style-type: none"> Semua bilyet itu tercatat atas nama ibunya. Pemerintah memberikan subsidi agar daya beli warga terjaga di tengah volatilitas harga pangan dan energi. Pada tahun ini, ada beberapa pemilu yang akan dilaksanakan, yaitu di Ukraina, Nigeria, Myanmar, dan Turki. 	99/20
3	V+Prep	<i>berasal dari, mengaku kepada, bersumber dari</i>	<ul style="list-style-type: none"> Dia menyebutkan pembelian itu berasal dari hasil sendiri. Kepada orang dekatnya, dia mengaku tak menanyakan ke presiden maksud persiapan itu. Ketidakpastian bersumber dari banyak hal. 	83/16.8
4	Adj+Prep	<i>dekat dengan, rentan terhadap, sepadan dengan</i>	<ul style="list-style-type: none"> Dua narasumber yang dekat dengan dua tokoh tersebut bercerita. Penggunaan sepeda motor sangat rentan terhadap fatalitas kecelakaan lalu lintas. Tuntutan terhadapnya terlampau tinggi dan tak sepadan dengan kontribusinya sebagai <i>justice collaborator</i>. 	20/4.1
5	N+V(used attributively) Note: This pattern corresponds to English collocation pattern N+To-Infinitive	<i>jurur bicara, kebebasan berbicara, keputusan bertarung</i>	<ul style="list-style-type: none"> Juru bicara Pengadilan Negeri Jakarta Selatan mengatakan sistem hukum Indonesia tak mengenal <i>amicus curiae</i>. Dikhawatirkan hukum pidana baru akan berdampak menakutkan pada kebebasan berbicara, berkumpul, dan berserikat. Tokoh itu menyebutkan keputusannya bertarung dalam pemilu mendatang bergantung pada perintah presiden. 	79/16
6	Adj+V(used attributively) Note: This pattern corresponds to English collocation pattern Adj+To-Infinitive	<i>batal menggelar, duka mendalam, sepakat mengesahkan</i>	<ul style="list-style-type: none"> Konsekuensi berat menanti Indonesia apabila akhirnya batal menggelar turnamen. Perbuatannya mengakibatkan hilangnya nyawa korban dan menyebabkan duka mendalam bagi keluarga korban. DPR sudah sepakat mengesahkan KUHP meski ada sejumlah pasal bermasalah. 	36/7.3
7	V+Obj+Prep	<i>membedakan-dengan, mencoret-dari, mengibaratkan-seperti</i>	<ul style="list-style-type: none"> Unsur inilah yang membedakannya dengan Pasal 338 KUHP tentang pembunuhan. FIFA seharusnya turut mencoret Israel dari piala dunia U-20 karena penjahajahan mereka kepada Palestina. Ia mengibaratkannya seperti “bisul pecah”. 	38/7.7
8	Binomials	<i>oleh dan untuk, suka atau tidak suka, tanah dan bangunan,</i>	<ul style="list-style-type: none"> Prinsip utama tentang pemerintahan oleh dan untuk orang banyak betul-betul terlaksana. Kepesertaan setiap negara dalam kejuaraan ini tidak ada kaitannya dengan masalah suka atau tidak suka. Peningkatan terbesar datang dari aset tanah dan bangunan seluas 253 meter persegi di Jakarta Pusat. 	10/2
Unique Pattern of Indonesian Grammatical Collocation				
9	Prep+Verb	<i>secara bertahap, secara terpisah</i>	<ul style="list-style-type: none"> Rencananya, 100 ribu ton daging sapi dari Brasil akan didatangkan secara bertahap. Secara terpisah, Direktur Angkutan Jalan Kementerian Perhubungan juga menyampaikan hasil survei. 	2/0.4

10	Prep+Adj	<i>dengan matang, secara sah, secara meyakinkan</i>	<ul style="list-style-type: none"> • Kami perlu membicarakan dengan matang soal sosok calon wakil presiden ini. • Dia terbukti secara sah dan meyakinkan bersalah melakukan tindak pidana pembunuhan. 	15/3.1
11	Prep+N+N	<i>di bawah kepemimpinan, di luar kewajaran</i>	<ul style="list-style-type: none"> • MK saat ini bisa memerankan perannya seperti masa-masa awal di bawah kepemimpinan tokoh tersebut. • Harga beras dalam dua bulan terakhir dinilai mahal dan di luar kewajaran. 	5/1
Patterns as the Correlative Conjunction				
12	Adv+Adv	<i>tidak hanya- tetapi juga</i>	<ul style="list-style-type: none"> • Masalah klasik itu tidak hanya terjadi pada beras tetapi juga komoditas lainnya. 	1/0.2
13	Adv+Prep	<i>bukan-melainkan</i>	<ul style="list-style-type: none"> • Undang-undang bukan daftar keinginan melainkan aturan main bersama berdasarkan konstitusi. 	2/0.4
14	Prep+Prep	<i>dari-hingga, dari-ke</i>	<ul style="list-style-type: none"> • Ia menggantikan ketua sebelumnya yang telah menjabat dari 2015 hingga 2018. • Pembahasan BPIH menunjukkan hasil yang membaik dari tahun ke tahun. 	4/0.8
Total Occurrences of Indonesian Grammatical Collocations				494/100