

Research Article

Investigating Verb-Collocate Patterns in Synonymous Self-Compound Nouns : A Corpus-Based Analysis

Ni Made Ayu Widiastuti ^{1*}, Ni Ketut Sri Rahayuni ², Samuel Shaw ³, Yasmine Aryani Dewi ⁴¹ Universitas Udayana, Indonesia; e-mail : ayu_widiastuti@unud.ac.id² Universitas Udayana, Indonesia; e-mail : sri_rahayuni@unud.ac.id³ Universitas Udayana, Indonesia; e-mail : shaw.2201541081@student.unud.co.id⁴ Universitas Udayana, Indonesia; e-mail : yasminedee42@gmail.com

* Corresponding Author: Ni Made Ayu Widiastuti

Abstract: Understanding verb-noun collocations is essential for exploring how abstract personal qualities are represented in language. This study aims to find out the types of verbs that occur before the five self-related compound nouns (self-esteem, self-confidence, self-assurance, self-worth, and self-respect), and investigate their collocational patterns. The data in this study were obtained from the Corpus of Contemporary American English (COCA). The data were collected by applying the documentation method. The technique involved searching for five compound nouns within the corpus and recording instances of their use. These occurrences, along with their collocates, were then copied and organized in an Excel spreadsheet. In the spreadsheet, columns were adjusted to separate and classify key elements such as the compound noun, the collocating verb, frequency data, and context sentences. The collected data were analyzed qualitatively to provide insights into lexical patterns, the types and tokens of frequency. The results show that the verbs that collocate with five self-related compounds in COCA vary in number and frequency, with self-confidence showing the highest types (20) and self-assurance the lowest (9). The verb build appears most frequently with self-confidence and self-esteem (4 tokens each), suggesting both are viewed as qualities that can be constructed or improved. Self-respect predominantly collocates with have (4 tokens), implying it is seen as possessive, while self-worth reflects both positive and negative framing through increase and reduce (2 tokens each). In contrast, self-assurance is frequently associated with lack (4 tokens) and show (2 tokens), indicating its absence or visibility. These patterns reveal that verb collocates offer important insights into how language frames self-constructions.

Keywords: Collocation; Corpus; Lexical Patterns; Synonyms; Verb Collocates.

Received: July 22, 2025

Revised: August 05, 2025

Accepted: August 19, 2025

Published: August 21, 2025

Curr. Ver.: August 21, 2025



Copyright: © 2025 by the authors.

Submitted for possible open

access publication under the

terms and conditions of the

Creative Commons Attribution

(CC BY SA) license

[\(https://creativecommons.org/licenses/by-sa/4.0/\)](https://creativecommons.org/licenses/by-sa/4.0/)

1. Introduction

Collocational analysis has been extensively researched in lexical studies, particularly in relation to improving semantic understanding and vocabulary acquisition by L2 learners or learners who study English as a second language [1], [2], [3]. Although adjective-noun and verb-noun collocations have been the subject of research [4], [5], [6], compound word collocations—especially those involving self-related compounds like self-esteem or self-confidence have received relatively little attention. Those compounds, among others, are the synonyms of the noun pride (Merriam-Webster Dictionary).

Compared with verb-compound noun collocations, adjective-compound noun collocations have been studied more widely. For the case of self-esteem, its collocations are highly recurring and have predictable patterns, such as low self-esteem or high self-esteem.

The studies on self-esteem as the evaluative language often focus on the use of adjectives to describe emotional states, which are common in the psychological field [7], [8]. Since the collocations of adjective and self-related compound nouns tend to have common patterns, this study narrows its scope to investigate verb-collocate patterns in self-compound nouns to uncover their semantic distinctions.

To investigate the collocational patterns of verb and self-related compound nouns, this study focuses on the two-word n-gram, that is, one word that appears right before the compound. This approach prioritizes collocational strength, offering an understanding of how language works in communicative contexts. Additionally, the two-word n-gram is suitable for useful applications like lexical simplification tools and vocabulary training, especially for L2 learners. When learners become familiar with common collocational patterns, they gain insight into native-like usage and meaning and have deeper semantic relationships between words within specific lexical fields.

Corpus linguistics provides empirical proof of word combinations in real texts that are relevant to collocational research to understand how language works in communicative contexts. Biber, Conrad, and Reppen [9] state that rather than just relying on intuition, collocational patterns as part of language structure must be recognized in actual language use. The value of employing corpora to examine semantic behavior across synonyms has been highlighted by recent research. Ly and Jung [10], for example, used corpus data to distinguish between synonymous adjectives and discovered that context and frequency convey a deeper meaning than dictionary definitions. To examine the synonymous self-related compound nouns for verb-collocate patterns, this study uses COCA to gather the data.

This research aims to identify the types of verbs that occur before the five self-related compound nouns (self-esteem, self-confidence, self-assurance, self-worth, and self-respect), and investigate their collocational patterns. Through these focuses, the study investigates patterns in how language is used and structured in natural contexts. The findings are expected to contribute to both linguistic theory and pedagogical practice by offering a clearer understanding of how lexical items are formed and function together.

2. Literature Review

The study of collocations has been explored in the fields of morphology, lexicology, lexicography, and corpus linguistics. However, specific attention to collocations of verbs and compound nouns is still limited. This study reviews previous research on compounding, collocation patterns, linguistic constructions related to self-esteem, and corpus-based methodologies.

According to a number of studies, morphological productivity is largely related to fundamental theories of compounding. Jurida and Pavlović [11] report a comprehensive typology of noun and adjective compounds, differentiating between root and synthetic types, endocentric and exocentric forms, and metaphorical expressions. Their discussion provides a foundation for comprehending compound behavior. Bell et al. [12] examine compounds to determine the influence of morphological structure on the length of consonants at compound-internal boundaries in English. Although the discussion of compounding is well-founded in these studies, they have not discussed the semantic nature of verb-self-related compound patterns.

Basthomi et al. [13] and also Ly and Jung [10] have investigated lexical collocations involving adjectives and nouns. While Ly and Jung compare near-synonyms (cute, pretty, and beautiful) in terms of collocational range and semantic prosody, Basthomi et al. investigate adjective amplification (e.g., absolutely essential) across native and non-native corpora. Both articles illustrate how frequency shapes semantic patterns and provide strong collocational extraction techniques using COCA and BNC as the data sources. These studies only discuss adjectival domains and do not investigate verbal or reflexive collocational behaviors.

The psychological concept of 'self' is framed to offer valuable insight into motivations. Stuppy et al. [14] examine low self-esteem and self-verifying consumption choices. They find that people frequently act in ways that support their negative self-views, which is conceptually consistent with verb-self compounds like punishing or blaming oneself. Similarly, Orth and Robins [15] report a difference in reflexive language use by arguing that higher self-esteem predicts improved emotional and academic functioning. These studies contextualize verb-self compounds as reflective of internal states, linking the target phenomenon conceptually instead of lexically.

Zavala et al. [16] broaden the psychological analysis by establishing a link between collective narcissism and out-group derogation and low self-esteem. This relationship implies that perceived threats to one's identity may give rise to defensive behaviors (such as self-defense and self-justification). Although their research does not specifically address collocation, it suggests that self-directed verbs operate within more general socio-emotional frameworks. This article, along with Stuppy et al. and Orth & Robins, offers essential

background information to investigate how reflexive verbs linguistically encode psychological states; however, none of them incorporates corpus-based linguistic data.

Joharry et al. [17] and also Xin and Rahim [18] both use corpus tools to analyze discourse's collocational networks. By examining the semantic transition of Covid-19 collocates from neutral to negative, Xin and Rahim show how lexical interpretation is influenced by contextual framing. Using CADS and GraphColl, Joharry et al. compare collocation patterns in political reporting. Both studies highlight how word combinations create discursive meaning and exhibit a high degree of technical alignment with possible verb-self collocate analysis. They do not, however, directly address reflexive constructions or internalized agency because of their subject matter, which is politics and health.

Rajeg and Rajeg [19] provide a useful model by using corpus-aligned construction grammar to analyze the English verbs *rob* and *steal* and their Indonesian translations. Their research demonstrates how verbs display systematic collocational behavior and syntactic structure by concentrating on argument roles (THIEF, TARGET, GOODS). Their approach demonstrates how deep structure mapping and collocational roles can uncover meaning. This method is directly applicable to verb-self investigation.

Even though collocations, compounding, and self-referential semantics have received a lot of attention, there is still a significant gap in the literature: no corpus-based, frequency-driven, or morphosyntactic analysis of verb-self compound collocations. To date, no studies have mapped the distribution of these structures across grammatical context, their dominance in certain semantic fields, or their frequency in native corpora (e.g., COCA). Thus, this research provides a crucial contribution by using corpus linguistics to show how reflexive verb constructions are used from collocational patterns.

3. Proposed Method

The data in this study were obtained from the Corpus of Contemporary American English (COCA), a large, balanced, and genre-diverse corpus widely recognized for linguistic research. COCA contains over one billion words drawn from spoken, fiction, magazine, newspaper, academic, blog, and other web-based texts, making it an ideal source for investigating language use in real contexts. The corpus was selected due to its comprehensive coverage of contemporary American English, its advanced search functionalities, and its capacity to provide frequency-based and contextualized collocational data. These features make COCA highly suitable for identifying verb–compound noun collocations and examining patterns of word formation in authentic language use.

The raw data from COCA was downloaded by searching the key words (self-esteem, self-confidence, self-assurance, self-worth, and self-respect) one by one. These compound nouns are the synonyms of the word *pride* (<https://www.merriam-webster.com/thesaurus/pride>). The synonymous compounds were selected based on the minimum availability data in COCA, that are sufficient for gathering 200 samples for each. The rest compounds that have also been searched were self-regard, self pride, self-assuredness, and self trust. The data provided for those four compounds were less than 200, therefore, they were under analysed. The occurrences, along with their collocates, were then copied from the results displayed by COCA and organized in an Excel spreadsheet. In the spreadsheet, columns were adjusted to separate and classify key elements such as the compound noun, the collocating verb, frequency data, and context sentences. This process allowed for a structured dataset that supports detailed analysis of collocational patterns.

The data were analyzed qualitatively [20] based on two main objectives. First, to identify the type of verbs that collocate with each of the five self-related compound nouns within the corpus, the chosen word combinations, referred to as the n-gram or the lexical bundle (McEnery & Hardie, 2012, p. 124), consist of two words, i.e., the node and one preceding word, in relation to the identification of collocations. The two-word n-gram is the combination of two words (a verb and a compound noun). Moreover, the collocational patterns were determined by examining the co-occurrence of verbs with compound nouns. The analysis provided insights into collocational patterns based on the types and tokens of frequency.

4. Results and Discussion

Before investigating the types and tokens of verbs that occur before each self-related compound nouns, it is necessary to report the types of word classes that occur before each compound noun. This aims to know the fact of the number verb-compound noun

collocations in the 200 samples of corpus data. Table 1 shows the types of word class that occur before *self-esteem*, *self-confidence*, *self-respect*, *self-worth*, and *self-assurance*.

Table 1. Types of Word Class Occurring before Each Self-related Compound Noun

Types of Word Class	<i>Self-esteem</i>	<i>Self-confidence</i>	<i>Self-respect</i>	<i>Self-worth</i>	<i>Self-assurance</i>
Adjective	51	34	7	26	45
Preposition	33	27	35	70	42
Possessive Determiner	28	27	45	36	20
Determiner	6	29	41	13	26
Verb	18	20	14	13	9
Conjunction	21	30	26	27	39
Punctuation	23	21	23	12	6
Possessive Noun	8	5	0	0	4
Pronoun	7	0	2	0	1
Noun	3	4	2	3	7
Adverb	2	0	3	0	0
Auxiliary Verb	0	0	2	0	0
Possessive Pronoun	0	1	0	0	0
Hashtag	0	0	0	0	1

Table 1 shows that adjectives, determiners, prepositions, possessive determiners, verbs, and conjunctions are among the word classes that precede each compound noun. However, there are notable differences in the frequency and dominance of each word class among the nouns, indicating distinct collocational behaviors.

There are three notable similarities from the findings. First, the occurrences of prepositions, possessive determiners, and adjectives are consistently high across all five compound nouns. This pattern implies that these compound nouns are frequently found in relational and descriptive structures. Possessive determiners (such as *her self-esteem*, *his self-respect*) highlight ownership or internal experience, whereas prepositions usually mark phrases that indicate relationships of belonging, comparison, or value (e.g., *of self-worth*, *with self-confidence*). Adjectives are often used to describe nouns to indicate various degrees or states, especially when used with *self-esteem* and *self-assurance*. The second similarity lies in the occurrence of function words like conjunctions and determiners that frequently preceded all five compounds, for example, *self-confidence* and *self-respect*. Third, verbs appear before all the compound nouns, but vary in frequency. This implies that these compound nouns may be used in verbal constructions as complements or objects, such as boost *self-respect* and gain *self-respect*. However, verbs appear less frequently before *self-assurance* and *self-worth*. Another finding is that punctuations like comma (,) and full stop (.) also occur before all five compounds. However, it does not affect the lexical or semantic interpretation of the compound noun.

Other word classes, such as possessive nouns, possessive pronouns, pronouns, nouns, adverbs, auxiliary verbs, and even hashtags, occur less frequently and are unevenly distributed before the compound nouns. It indicates that they are not the typical patterns of the compounds but rather context-, genre-, or stylistically marked uses.

4.1. Types of Verbs Occur before Five Self-related Compound Nouns

The verb collocate distribution that appears right before five compound nouns (*self-assurance*, *self-worth*, *self-esteem*, *self-confidence*, and *self-respect*) is shown in Tables 2 to 6. With frequent collocational patterns, these tables give a clear picture of how various verbs semantically frame each self-related concept.

Table 2. Types of Verbs Occurring before *Self-confidence*

No.	Verb	Frequency
1	build (1)/builds (3)	4
2	develop	2
3	improve	2
4	radiated	2
5	deluded	1
6	enhancing	1
7	have	1
8	include	1
9	lacking	1
10	lost	1
11	promoted	1
12	regaining	1
13	requires	1
14	rising	1
	Total	20

There are a total of 20 occurrences of 14 different verbs in Table 2 that come before self-confidence. The most common verbs are *build/builds* (4 occurrences), *develop*, *improve*, and *radiated*, each of which occurs twice. The majority of these indicate that *self-confidence* is frequently presented as something that can be actively created or enhanced. Additionally, the list includes verbs like *deluded* and *lost*, which create a negative semantic contrast by framing *self-confidence* as something that can be misguided or undermined.

Table 3. Types of Verbs Occurring before Self-esteem

No.	Verb	Frequency
1	build	4
2	increase	2
3	improving	1
4	diminish	1
5	enhances	1
6	gain	1
7	have	1
8	limited	1
9	lost	1
10	measure	1
11	nurtured	1
12	promotes	1
13	raises	1
14	boost	1
	Total	18

Table 3 shows the list of 18 different verbs before *self-esteem*. The verb *build* appears four times, followed by *increase* twice. Every other word, such as *diminish*, *gain*, *nurtured*, *boost*, and *measure*, appears once. *Self-esteem* is more frequently defined as something that is fostered, assessed, or changed internally than *self-confidence*, which is frequently externalized or performed. Particularly, the use of verbs like *diminish*, *limited*, and *lost* conveys vulnerability and emphasizes how *self-esteem* can decrease.

Table 4. Types of Verbs Occurring before *Self-respect*

No.	Verb	Frequency
1	have (2)/ having (2)	4
2	develop	1
3	enhanced	1
4	feels	1
5	has	1
6	includes	1
7	lack	1
8	learn	1
9	require	1
10	say	1
11	suggests	1
	Total	14

The verbs have/having appear for times as shown in Table 4, which lists verbs that collocate with self-respect. Other verbs like *develop*, *enhanced*, *feels*, and *lack* appear once. This table shows that rather than being something that is developed or acquired, *self-respect* is frequently portrayed as a state or quality that is possessed or experienced. The verbs *feel*, *say*, and *suggests* also imply an emotional or discursive framing, showing that *self-respect* is more individualized and subjective, influenced by relationships, belief systems, and introspection.

Table 5. Types of Verbs Occurring before *Self-worth*

No.	Verb Collocate	Frequency
1	reduced	2
2	increased (1)/ increasing (1)	2
3	acquire	1
4	buried	1
5	cited	1
6	comparing	1
7	confuse	1
8	diminished	1
9	equate	1
10	finding	1
11	labeled	1
	Total	13

There are thirteen verbs in Table 5, which precede *self-worth*. The most common verbs are *reduced* and *increase/increasing*, which appear twice. Every other verb, such as *diminished*, *labeled*, *buried*, and *equate*, only occurs once. This lexical set implies that discussions of *self-worth* are typically connected to external labeling or judgments and occur in contexts of decline, comparison, or confusion. *Self-worth* is framed as damaged, misinterpreted, or suppressed, in contrast to *self-confidence* and *self-esteem*, which involve growth-oriented verbs.

Table 6. Types of Verbs Occurring before *Self-assurance*

No.	Verb Collocate	Frequency
1	lack (2)/ lacking (2)	4
2	show	2
3	believes	1
4	complete	1
5	exuded	1
	Total	9

Table 6 displays the collocates of five types of verbs and *self-assurance*. The verbs *lack/lacking*, and *show* occur four times and twice, respectively. Besides that, each of the verbs *believes*, *complete*, and *exuded* occurs once. According to this distribution, self-assurance can be seen as either completely absent or demonstrated and emanated. While verbs like *lack* and *believes* describe an internal conviction or absence, verbs like *exuded* and *show* highlight external behavioral expression. It is more in line with *self-confidence* because *self-assurance* seems to be more performative and observable than *self-respect* and *self-worth*.

4.2. Collocational Patterns of Verbs and Five Self-related Compound Nouns

Even though high-frequency word classes (like possessive determiners, adjectives, and prepositions) frequently appear before all compound nouns, this does not imply that every word in those classes is a collocate or a member of a collocational pattern. The basis for collocations is the statistical tendency of particular words to occur together more frequently than would be assumed by chance. For example, although many adjectives can grammatically occur before *self-esteem*, only a few, such as *high* and *low*, are used often enough in everyday speech to be regarded as true collocates. Furthermore, word class frequency does not reflect semantic preferences, but rather syntactic roles. Although the high frequency of a word class (adjectives, for example) indicates that the noun is frequently modified, it does not identify the important words in the class. While prepositions frequently come before *self-worth*, only certain prepositions, such as *of*, *with*, or *in* can form frequent or semantically significant combinations (e.g., *sense of self-worth*). This indicates that lexemes, not categories, are the concern of collocational significance.

In this study, verb collocates of compound nouns are specifically examined to determine how language use is represented through action or process. In order to investigate a distinct collocational behavior, this study purposefully limits its focus to verbs, even though previous research and preliminary corpus findings suggest that adjectives (such as *low self-esteem*, *high self-confidence*) are frequent and semantically salient.

The study of verb collocates that occur before five self-related compound nouns—*self-worth*, *self-assurance*, *self-esteem*, *self-confidence*, and *self-respect*—shows that only a small percentage of verbs show consistent and meaningful collocational patterns. In addition to grammatical compatibility, these patterns are also influenced by lexical conventionality, frequency, and semantic relevance.

The collocational pattern is more varied when it comes to *self-confidence*. Considered as a single lemma, the verbs *build* and *builds* appear four times and are dominant. This supports a powerful metaphorical framework, where confidence is something that can be created or developed, much like *self-esteem*. Verbs that occur twice, such as *develop* and *improve*, also support this developmental schema and are frequently used in texts that are psychological or motivational. Remarkably, the verb *radiated* occurs twice, suggesting that it is linked to observable behavior or emotional expression. It may also have contextual significance in genres that are descriptive or narrative. However, some verbs, such as *deluded*, *include*, or *requires*, are too rare and semantically inconsistent to be considered core collocates.

With four occurrences, the verb *build* is the strongest collocate for *self-esteem* and reflects a common metaphor in psychological discourse: building or strengthening internal traits. With two occurrences, the verb *increase* also exhibits moderate recurrence, which is consistent with the widely held belief that *self-esteem* is a measurable quality that can fluctuate. Other verbs like *boost*, *gain*, *raise*, and *nurture* are semantically relevant, but because they only occur once each, they cannot be firmly established as fixed collocates without additional corpus-based evidence. On the other hand, verbs such as *measure*, *limited*, or *have* are too general or structurally generic to have any collocational meaning.

The pattern lessens with the collocations of the verb and *self-respect*. The fact that the verbs *have/having* appear four times indicates that this compound noun is frequently used in possession-related expressions. These forms, however, lack the lexical specificity needed for collocates because they are incredibly broad and occur with almost all nouns. Since they only appear once each, other verbs like *develop*, *enhance*, *learn*, and *require* are not as good predictors of common usage. Furthermore, verbs like *say* and *suggest* reflect reporting structures.

Self-worth collocates are also quite limited. The only word that appears more than once is *reduced*, and only twice at that. However, the fact that this verb reflects a common evaluative frame that discusses *self-worth* in terms of loss or damage makes it noteworthy. Semantically related to the concept are verbs like *diminished*, *increased*, and *buried*, but they only occur once

each and are not frequent enough to be fixed collocates. Many others, such as *labeled*, *cited*, and *confused*, appear to be context-specific constructions rather than part of lexical patterns.

The verbs *lack/lacking* occur four times, and *show* occurs twice for self-assurance. These verbs are used when social situations either lack or display self-assurance. The verb *show* is especially useful when describing characteristics that you have seen in other people (e.g., “she showed great self-assurance”). Verbs that only occur once, like *believes*, *excluded*, and *complete*, may be used in literary, metaphorical, or expressive contexts rather than following consistent collocational patterns.

Among the five compound nouns, the verb collocates that have the strongest support are *lack* and *show* for *self-assurance*; *build*, *develop*, and *improve* for *self-confidence*; *build* and *increase* for *self-esteem*; and *reduced* for *self-worth*. In addition to being semantically appropriate, these verbs are frequently used enough to imply lexical preference. The frequency and fixedness requirements needed to be regarded as strong collocational patterns are not met by the remaining verbs, even though they are occasionally relevant. This analysis emphasizes how crucial it is to differentiate between statistically supported lexical collocation and grammatically acceptable co-occurrence.

5. Conclusions

In conclusion, the verbs that collocate with self-related compounds among 200 samples (each compound) found in COCA vary from 9 to 20. *Self-confidence* has the highest number of verb types, in which there are 20 types of verbs. The dominant occurrences of verbs and *self-confidence* collocations are *build/builds* (4 tokens), *develop* (2 tokens), *improve* (2 tokens), and *radiated* (2 tokens). In the second position, *self-esteem* collocates with 18 types of verbs, with *build* (4 tokens) and *increase* (2 tokens) as the dominant frequency. *Self-respect* and *self-worth* have a similar number of verb types, they are 14 and 13, respectively. There is only one verb that mostly collocates with *self-respect*, namely *have/having* (4 tokens), but for *self-worth*, there are two verbs – *reduce* and *increase* (each occurs twice) – as collocations. The least total number of collocations is verbs with *self-assurance*, which only has 9 types of verbs. *Lack* (4 tokens) and *show* (2 tokens) are the dominant collocations. The other types of verbs, which only occur once for each self-related compound noun, are considered less productive.

In terms of collocational patterns of verbs and five self-related compound nouns, the results reveal meaningful distinctions in how each concept is linguistically framed. The verb *build* has the strongest association with *self-confidence* and *self-esteem* (4 tokens each), indicating that these are commonly thought of as constructible and improvable. Furthermore, *self-esteem* is associated with growth-oriented processes like *increase*, while the verbs *develop*, *improve*, and *radiated* further show that *self-confidence* is something that can be fostered and expressed externally. On the other hand, *self-respect* primarily collocates with the verb *have* (4 tokens), indicating a possessable quality. Both positive and negative values are present in *self-worth*, which frequently collocates with the words *reduce* and *increase* (2 tokens each). Finally, the verb *lack* appears four times and the verb *show* twice, showing that *self-assurance* is frequently viewed as either a visible quality or an absence. In addition to highlighting the semantic distinctions between self-compounds, these patterns also show how verb collocates exhibit deeper linguistic conceptualizations of identity, emotional well-being, and social behavior.

Author Contributions: Methodology: Ni Made Ayu Widiastuti, Ni Ketut Sri Rahayuni; Software: Ni Made Ayu Widiastuti, Samuel Shaw, and Yasmine Aryani Dewi; Validation: Ni Made Ayu Widiastuti; Formal analysis: Ni Made Ayu Widiastuti, Ni Ketut Sri Rahayuni; Investigation: Ni Made Ayu Widiastuti; Resources: Ni Made Ayu Widiastuti; Data curation: Ni Made Ayu Widiastuti; Writing—original draft preparation: Ni Made Ayu Widiastuti, Ni Ketut Sri Rahayuni; Writing—review and editing: Ni Made Ayu Widiastuti; Visualization: Ni Made Ayu Widiastuti; Supervision: Ni Made Ayu Widiastuti; Project administration: Ni Made Ayu Widiastuti, Ni Ketut Sri Rahayuni; Funding acquisition: Ni Made Ayu Widiastuti”

Funding: This research was funded by the Institute for Research and Community Service, Udayana University, grant number B/229.268/UN14.4.A/PT.01.03/2025.

Data Availability Statement: The collected sample data for collocation analysis taken from COCA can be accessed from this link: https://drive.google.com/drive/folders/12So9xo_Qj2jM3Zy0MVCybIoGW9AC4l48?usp=sharing. This is in the restricted mode; permission from the corresponding author is needed to see the Data Set.

Acknowledgments: We would like to express our sincere gratitude to our advisors, Prof. Dr. Drs. I Made Rajeg, M.Hum., and Gede Primahadi Wijaya R., S.S., M.Hum. Ph.D. for their guidance, support, and valuable insights throughout this research. I also thank the English Department, Faculty of Humanities, Udayana University, for providing a conducive research environment.

Conflicts of Interest: The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript; or in the decision to publish the results.

References

- Ajmal, M., Kumar, T., Ritonga, M., & Nukapangu, V. (2022). A corpus-based analysis of the adjectives and synonyms beautiful, handsome, and pretty. *World Journal of English Language*, 12(2), 159. <https://doi.org/10.5430/wjel.v12n2p159>
- Basthomi, Y., Yannuar, N., Widiati, U., & Martiningtyas, R. (2015). C-Smile, COCA, and BNC: A focus on amplifiers and adjective collocations. *3L: Language, Linguistics, Literature®*, 21(2), 73–87. <https://doi.org/10.17576/3L-2015-2102-06>
- Bell, M. J., Ben Hedia, S., & Plag, I. (2021). How morphological structure affects phonetic realisation in English compound nouns. *Morphology*, 31(2), 87–120. <https://doi.org/10.1007/s11525-020-09346-6>
- Biber, D., Conrad, S., & Reppen, R. (1998). *Corpus linguistics: Investigating language structure and use*. Cambridge University Press.
- Creswell, J. C. (2009). *Research design: Qualitative, quantitative, and mixed methods approach* (3rd ed.). SAGE Publications.
- Fang, N., & Zhang, P. (2021). L1 congruency, word frequency, collocational frequency, L2 proficiency, and their combined effects on Chinese–English bilinguals’ L2 collocational processing. *International Journal of Bilingualism*, 25(5), 1429–1445. <https://doi.org/10.1177/13670069211024747>
- Gablasova, D., & Brezina, V. (2025). Adjective + noun collocations in L2 spoken English: How robust is the role of proficiency? *International Journal of Learner Corpus Research*, 11(1), 79–113. <https://doi.org/10.1075/ijlcr.23034.gab>
- Golec De Zavala, A., et al. (2020). Low self-esteem predicts out-group derogation via collective narcissism, but this relationship is obscured by in-group satisfaction. *Journal of Personality and Social Psychology*, 119(3), 741–764. <https://doi.org/10.1037/pspp0000260>
- Joharry, S. A., Turiman, S., & Mohd Noor, M. (2023). Comparative collocation(al) analysis between two Malaysian online newspapers: A case study of the 2020 Sabah election day. *3L: Language, Linguistics, Literature®*, 29(2), 146–168. <https://doi.org/10.17576/3L-2023-2902-11>
- Jurida, S. H., & Pavlović, T. (2023). Noun compounds and adjective compounds in English. *Science International Journal*, 2(4), 73–80. <https://doi.org/10.35120/sciencej0204073h>
- Kruawong, T., & Phoocharoensil, S. (2022). A genre and collocational analysis of the near-synonyms teach, educate and instruct: A corpus-based approach. *TEFLIN Journal*, 33(1), 75–97. <https://doi.org/10.15639/teflinjournal.v33i1/75-97>
- Ly, T. H., & Jung, C. K. (2015). A corpus investigation: The similarities and differences of cute, pretty, and beautiful. *Journal Title*, 21(3), 125–140.
- Masrai, A. (2023). Lexical knowledge and L2 general language proficiency: Collocational competence and vocabulary size as determinants of lexical knowledge. *Cognitive Processing*, 24(2), 289–300. <https://doi.org/10.1007/s10339-022-01120-2>
- Nguyen, D. T., Wright, E. P., Dedding, C., Pham, T. T., & Bunders, J. (2019). Low self-esteem and its association with anxiety, depression, and suicidal ideation in Vietnamese secondary school students: A cross-sectional study. *Frontiers in Psychiatry*, 10, Article 698. <https://doi.org/10.3389/fpsy.2019.00698>
- Orth, U., & Robins, R. W. (2022). Is high self-esteem beneficial? Revisiting a classic question. *American Psychologist*, 77(1), 5–17. <https://doi.org/10.1037/amp0000922>
- Rajeg, G. P. W., & Rajeg, I. M. (2022). A corpus linguistic study of constructional equivalence for the Indonesian translation of rob and steal based on the Open Subtitles Parallel Corpus. *Parole: Journal of Linguistics and Education*, 12(2), 177–197.
- Strakatova, Y., & Hinrichs, E. (2019). Semantic modelling of adjective-noun collocations using FrameNet. In *Proceedings of the Joint Workshop on Multiword Expressions and WordNet (MWE-WN 2019)* (pp. 104–113). Association for Computational Linguistics. <https://doi.org/10.18653/v1/w19-5112>
- Stuppy, A., Mead, N. L., & Van Osselaer, S. M. J. (2020). I am, therefore I buy: Low self-esteem and the pursuit of self-verifying consumption. *Journal of Consumer Research*, 46(5), 956–973. <https://doi.org/10.1093/jcr/ucz029>
- Xin, C. L., & Abdul Rahim, H. (2021). From positive to death: A corpus-based semantic analysis of COVID-19 representation in Malaysian English news reports. *3L: Language, Linguistics, Literature®*, 27(4), 16–28. <https://doi.org/10.17576/3L-2021-2704-02>