

Lexical Collocation Analysis in Master's Student Reflective Writings

Riwayat draf artikel
Diserahkan 28-03-2023
Direvisi 04-05-2023
Diterima 10-05-2023

Mia Dreina Antira Pujiningtyas¹, Barli Bram²
English Education Master's Programme, Sanata Dharma University,
Yogyakarta^{1,2}
miatyas35@gmail.com¹, correspondence: barli@usd.ac.id²

ABSTRACT: The study aims to explore lexical collocations in master's student reflective writings by classifying the collocations and figuring out the strengths of the word combinations using the Mutual Information (MI) scores. As a corpus-based study, the Corpus of Contemporary American English (COCA) was employed by the researchers. Therefore, the MI score of each lexical collocation could be carefully gathered and compared. The result of the study showed six types out of seven types of lexical collocations found in students' reflective writings. Among the combinations, the noun and verb combinations were the most commonly found. The summarized data revealed that the number of word combinations with a stronger association was higher than the number of word combinations with a less significant association. Being get used to various English language inputs, the students were able to possess a wide range of lexical collocation repertoire and produce writings containing various word combinations with the significant association.

Keywords: EFL context, lexical collocation, reflective writing

***ABSTRAK:** Makalah ini bertujuan mengeksplorasi kolokasi leksikal di dalam tulisan reflektif yang ditulis oleh mahasiswa magister dengan mengklasifikasikan kolokasi dan mencari tahu kekuatan hubungan kombinasi kata menggunakan nilai Mutual Information (MI). Sebagai studi berbasis korpus, Corpus of Contemporary American English (COCA) digunakan di dalam penelitian ini. Oleh karena itu, nilai MI dari setiap kolokasi leksikal dapat diambil dan dibandingkan secara seksama. Hasil penelitian menunjukkan adanya enam jenis dari tujuh jenis kolokasi leksikal yang ditemukan di dalam tulisan reflektif siswa. Diantara kombinasi kata tersebut, kombinasi kata benda dan kata kerja adalah kombinasi yang paling banyak ditemukan. Dari data yang telah dirangkum, ditemukan bahwa jumlah kombinasi kata dengan asosiasi yang lebih kuat lebih banyak daripada jumlah kombinasi kata dengan asosiasi yang kurang signifikan. Dengan membiasakan diri dengan berbagai input bahasa Inggris, siswa mampu menguasai repertoar kolokasi leksikal yang luas dan menghasilkan tulisan yang mengandung berbagai kombinasi kata dengan asosiasi yang signifikan.*

***Kata kunci:** kolokasi leksikal, konteks bahasa asing, tulisan reflektif*

INTRODUCTION

Learning the English language as a foreign learner is challenging. Understanding vocabulary, grammar, and reading, writing, speaking, and listening skills are crucial in language learning. However, being able to use the language to communicate naturally is also a major aspect of mastering the language. By that, learners could be more aware of the word's meaning, the relationship between sentences, the discourse, and the context of the language output (Alsayed, 2020). Being exposed to the English language on an everyday basis enables native speakers to be more capable of using English expressions communicatively and naturally. Nevertheless, English as a Foreign Language (EFL) learners would find it difficult.

One of the ways to produce the English language more naturally is by using collocations. According to Tanihardjo (2018, p. 64), collocations are pivotal to being learned because "...not only do they enhance accuracy, but they also enhance fluency." Collocation is an arbitrary lexical relationship between words that are more affected by the common usage of native speakers (Benson et al., 1986 as cited in Thao & Trng, 2021). By learning collocation, learners could utilize a pair of words that go well together. It is because some words are not meant to be used together due to the different meanings or the uncommon or unnatural expressions as a result. However, due to the influence of the mother tongue, EFL learners would need more understanding of collocations. The different collocations in the mother tongue and the English language are also the reason for the difficulties in learning it (Thao & Trng, 2021).

Master's students in the English Education department as EFL learners are required to produce English writings both academic and non-academic. Thus, to create a good and natural flow in the writing, correctly used collocations should be included. Therefore, analyzing the master's students' use of collocations in their writing is the aim of this study. In the study, the focus was on lexical collocation. Lexical collocation has to do with content words and meaning (Benson & Ilson, 1986). The Master's students were considered to have advanced grammar skills. However, using correct lexical collocation requires more than proficiency in using grammar. Hence, analyzing lexical collocation in the master's students' writing was the focus of the study.

The use of accurate lexical collocations provides clearer meanings of the delivered messages. Hence, it is pivotal for language users to understand how words could be used in the proper context (Umair et al., 2023). However, as mentioned before, conveying messages with accurate lexical collocations would be more challenging for EFL learners than it is for the native speakers. In addition, misused lexical collocations could result in an awkward expression or even unclear meanings. This then could incline to misunderstanding. Understanding the strengths of lexical collocations that the Master's students used could show how the students put their attention to organizing their words and sentences for clearer messages.

The investigations of collocations have also been conducted by some researchers. Kang (2018) compared word association and collocation which revealed the closeness of both parties with imperfections due to the distinct characteristics. Wongkhan and Thienthong (2021) explored the influence of EFL learners' academic experience on academic collocations and synonymy acquisition. The exploration showed that collocation repertoire could be developed incidentally if students have enough exposure to academic discourse. Thao & Trng (2021) investigated the popular collocations that the students used in their academic writing. They found that Verb-Noun and Adjective-Noun collocations were more frequently used by the students. In the Indonesian context, Tanihardjo (2018) investigated the collocations in undergraduate students' essay writing which discovered the errors the students made from the perspective of their first language, the Indonesian language. Alfiandita and Ardi

(2020) and Sukur (2020) investigated the collocation in Indonesian school textbooks. The studies revealed the low variety of collocations used in the books and the need to provide additions to the collocations and other formulaic competence in the materials. Studies regarding EFL collocations topic has been conducted by researchers in Indonesia. However, the investigation of master's students' collocation usage in the Indonesian context needs more attention. Thus, this study would also enrich the collocation research in the context of master's students in the Indonesian context.

The purpose of the current study is to analyze the lexical collocations in the master's students' reflective writings. Based on the aim of the study two research questions would be answered in the study as follows: 1) What types of lexical collocations are used in master's student reflective writings? 2) What are the strengths of the collocations used by the students?

Collocation

Collocation is identified as words that go together due to the commonality for occurring together e.g. take a chance or the parts of the collocation that should be combined to form a particular meaning e.g. by the way (Nation, 2001). Biber and Reppen (2015, p. 41) agreed on the notion of collocation by elaborating on the Firthian terms that, "collocation refers to the relationship between a word and its surrounding context where frequent co-occurrence with other words or structures helps to define the meaning of the word." From the explanation, the existence of collocation help produces the commonly used group of words. The commonly found group of words then depicts the natural flow of the produced sentences. By understanding collocation, some benefits could be acquired. Bartsch (2004, p. 18) noted that "Collocation plays an important role in disambiguating a polysemous word, and together with semantic prosody, helps the language-user to choose near synonyms appropriately." It means that using the correct collocation help to give a clearer message and reduce ambiguity. It is also added that collocation as a "useful strategy" allows the language user to efficiently convey complicated notions. Moreover, an improvement in communicative competence, language skills, and fluency could be acquired with collocations knowledge (Bui, 2021).

According to Sinclair (1966), some terms are used to address the parts of collocation. The term node is used to refer to the main item in the collocations. The number of lexical items preceding and/or following the node is called span. Collocates refer to the lexical items within the span. For instance, in the phrase run out of, the collocation is run out of, with run as the node, and out and of as the collocates. The span on the right side of the node would be two.

Collocation could be classified into several types. According to Benson and Ilson (1986), there are two types of collocation namely grammatical collocation and lexical collocation. Grammatical collocation comprises combinations of content words and also function words like prepositions, infinitives, and clauses.

Noun + preposition: blockade against
Noun + to-infinitive: the purpose to finish it
Noun + that clause: they make a promise that they will work together
Preposition + noun: by accident
Adjective + preposition: fond of the horse
Adjective + to-infinitive: It's nice to see you.
Adjective + that-clause: I am so happy that you do well.
Verb + to-infinitive: they start to argue

On the other hand, lexical collocation includes only the content words such as verbs, nouns, adverbs, and adjectives without prepositions and clauses.

Verb (that refers to activation or creation verb) + pronoun/noun/prepositional phrase: shot a bullet, come to an agreement
Verb (that refers to eradication or cancelation) + noun: crush resistance, reject an appeal
Adjective + noun; attributive noun + noun: strong tea, paper plane
Noun + verb (that is performed by the noun): bee sting
Noun quantifier + noun: a piece of cake
Adverb + adjective: surprisingly happy
Verb + adverb: argue heatedly

EFL Learners and English Collocation

Errors in collocation commonly happen among EFL learners. One of the causes of the errors has something to do with the action of transferring the English collocation to the ones equivalent to the mother tongue without the knowledge of the co-occurring words that should be used in a certain context (del Pilar Agustín-Llach, 2019; Bui, 2021). For instance, 'medicine' does not co-occur with 'drink' but the words were opted because of the mother tongue collocation. Differently, Reynolds and Teng (2021) noted that:

“Each type of collocation may impose difficulty or confusion on learners of English. This is because there is hardly any clear-cut guideline for non-native learners to decide which combinations are exactly acceptable and whose co-occurrence are highly predictable; or which ones are possibly acceptable for creative purpose, or which ones are simply unacceptable.” (p. 247)

This is in line with Karjo and Metta (2019, p. 291) who revealed that EFL learners “do not have any guide on which words are meant to each other.” For instance, the phrase 'make a conclusion' was directly translated from the Indonesian phrase 'membuat kesimpulan'. The English phrase is wrong because the noun 'conclusion' collocates with 'draw' not 'make'. This error occurs because 'make (membuat)' is more commonly used in Indonesian instead of 'draw (menggambar)' which becomes 'draw a conclusion'. The interference from the source language leads to confusion that results in collocation errors (Lubis., 2020).

RESEARCH METHOD

This study is corpus-based. The data was acquired from 18 reflective writings from the Educational Psychology class written by English Education Master Program (EEMP) students of Sanata Dharma University, Yogyakarta. The reflective writings were opted to be the source of data due to the naturality of the writings. While writing an academic text, students tend to check and recheck the sentence construction which exhibits less of the students' authentic way to convey their ideas. With reflective writings, the students were more spontaneous and this made the writings represent the way the students use the English language. The data were read and re-read carefully while highlighting the parts of speech combinations listed in Benson and Ilson's (1986) lexical collocation classification. Afterward, the listed word combinations were classified according to the seven subtypes of Benson and Ilson's (1986) lexical collocation with the following naming.

Table 1. Lexical collocation

No.	Naming	Lexical Collocation type
1.	L1	Verb + Noun (creation and activation)
2.	L2	Verb + Noun (eradication or cancelation)
3.	L3	Noun + Verb
4.	L4	Noun + Noun
5.	L5	Adjective + Noun
6.	L6	Adverb + Adjective
7.	L7	Verb + Adverb

The frequency and the strength of the words' co-occurrences were checked in the Corpus of Contemporary American English (COCA). The corpus was selected due to the size of the data which could be considered massive. This massive data could increase the reliability of the findings. The strength of the collocation concurrences was analyzed using the Mutual Information (MI) score. The MI score showed "the amount of non-randomness present when two words occurred" (Hunston, 2002, p. 71). Collocation with an MI score above 3 is viewed as word pairs that associate significantly (Szudarski, 2017, p. 28). The MI measurement was applied instead of the frequency because the high frequency of the word combinations does not always exhibit the indication of strong association (Prayogi et al., 2022).

FINDINGS AND DISCUSSION

The gathered data would be presented and discussed in this section. The data is presented in the form of tables containing lists of lexical collocation extracted from master's students' reflective writings. Two research questions were answered in this study. The first one is "What types of lexical collocations are used by the students?" and "How are the strength of the collocations used by the students?"

Types of lexical collocations used by the students

Among the seven lexical collocations types classified by Benson and Ilson (1986), not all were used by the master's students. As could be seen in Cart 1, the researchers found six lexical collocation types out of seven which were L1, L3, L4, L5, L6, and L7.

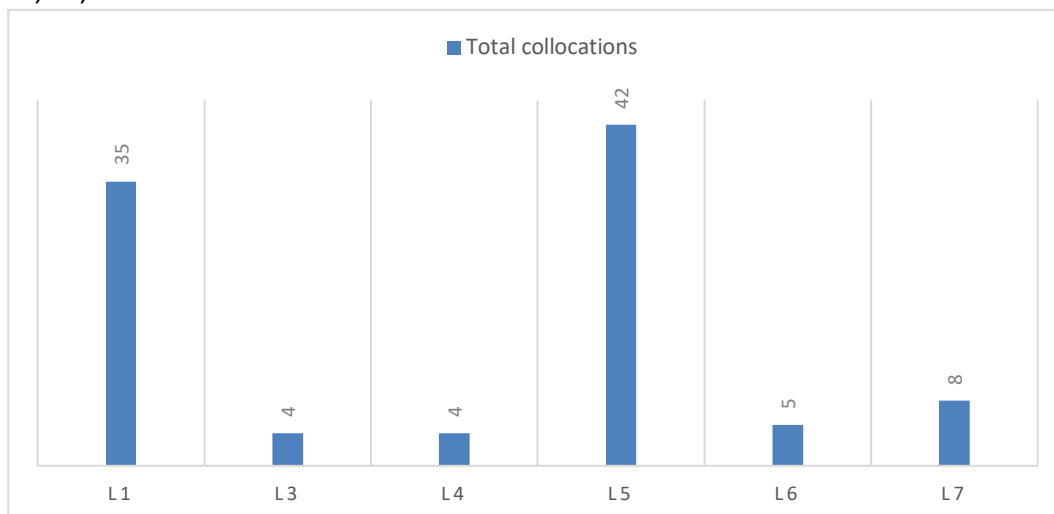


Chart 1. Lexical Collocations in Students' Writings

The first and the second biggest numbers were found in the L1 and L5 categories which were classified the Verb-Noun combination and Adjective-Noun Combination. This finding was in line with Trng and Tao's (2021) findings. Combinations with verbs or with nouns contained in writings could be more commonly found because of the importance of verbs and nouns (as subjects or objects) to construct sentences. On the other hand, the smallest number of collocations could be seen from the L3 and L4 groups which were classified as Noun-Verb and Noun-Noun combinations. The groups contained smaller numbers because most of the combinations used by the students could be considered open combinations. This was due to the high possibility that the nodes had to be combined with various words. For example, the combination of 'teacher' and 'help' (see Table 2). As in the students' writing, 'Teacher' was followed by various verbs like 'help, ask, or scold'. On the other hand, the mentioned verbs could have various subjects either. However, there were some collocations like 'teacher' and 'teach' or 'Students' and 'learns' that occurred together as strong collocations due to the strong connection of the subject and verb (this part would be further explained in the next section).

Another group with an almost similar number was L6 which was Adverb-Adjective combinations. The variations of adverbs used by the students were not varied. The adverbs used by the students were mostly 'more' and 'very'. The last group was L7 as Verb-Adverb combinations. The students rarely modified their verbs using adverbs. There were only several verbs that were particularly modified. For instance, the combinations were 'participate actively, treat equally, look forward, work hard', and 'impact deeply'. Those words could be commonly found together.

The data summarized in Chart 1 portrayed how much students implemented collocation in their writings. Regardless of the significant difference in the number of each classification, the students implemented almost all types of lexical collocation which were six out of seven types. As subjects and predicates were the aspects that constructed most sentences, combinations with verbs and nouns were the most used by the students.

The strength of the collocations used by the students

As explained in the method section, the strength of the collocation was measured using the MI score. The MI score was found by consulting each collocation to COCA. The nodes of the collocations were searched in the collocation search section of the online corpus. The span used was four for each side (left and right) which meant that the collocate was positioned in the span of four words preceding or following the node.

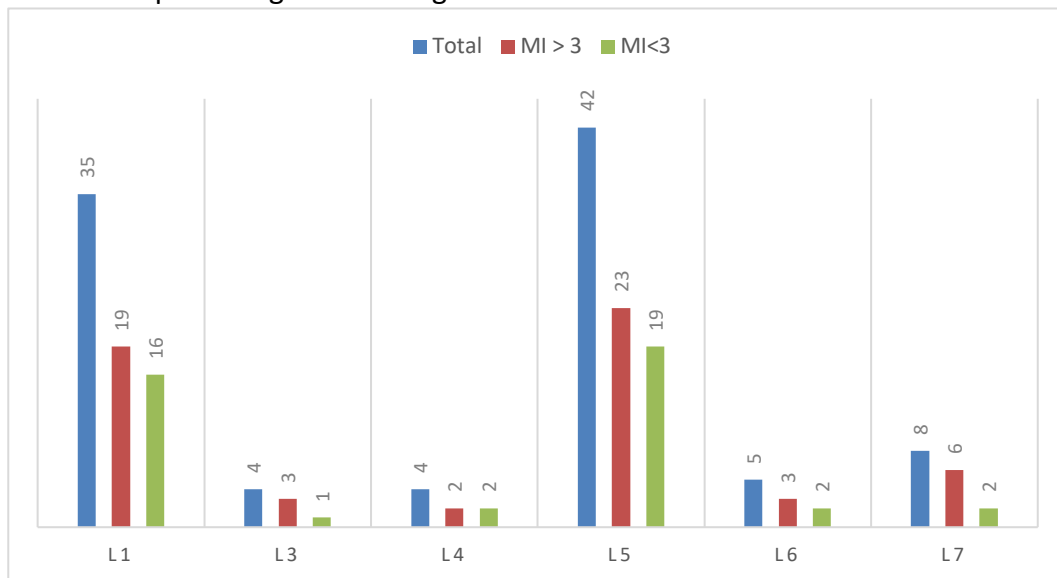


Chart 2. MI Score Classifications of the collocations Used by Students

The collocation was considered strong if it possessed an MI score of three or above (Szudarski, 2017). Chart 2 summarised the MI scores of each type of lexical collocation. It could be inferred that not all collocations in the students’ writings were included in the strong collocation classification. Some collocations were classified as less significant. The following table contains examples of collocations in students’ writings with high and low MI scores.

Table 2. Word Combinations and MI Scores

WORD COMBINATIONS			MI	
VERB + NOUN CREATION AND ACTIVATION (L1)	solve	problems	MI ≥ 3	6.97
	achieve	goals		6.44
	learn	lesson		6.12
	attend	meeting		5.72
	acquire	skills		5.51
	pass	examination	MI < 3	2.97

	understand	concepts		2.87
	feel	urgency		2.78
	understand	need		2.61
	express	ideas		2.58
	feel	need		2.52
NOUN + VERB (L3)	students	learn	MI ≥ 3	3.40
	teacher	teach		3.63
	educator	teach		3.54
	teachers	encourage	MI < 3	2.84
NOUN + NOUN (L4)	learning	process	MI ≥ 3	4.18
	classroom	activity		3.62
	learning	goals	MI < 3	2.67
	learning	perception		2.39
ADJECTIVE + NOUN (L5)	mental	health	MI ≥ 3	7.07
	urgent	need		6.41
	mental	being		5.93
	deeper	understanding		5.74
	controllable	factors		5.30
	future	educators	MI < 3	2.97
	comfortable	environment		2.92
	various	topics		2.79
	inseparable	aspects		2.77
	positive	learning		2.74
ADVERB + ADJECTIVE (L6)	very	exited	MI ≥ 3	3.83
	more	effective		3.50
	very	important		3.27
	very	strong	MI < 3	2.67
	very	motivated		2.41
VERB + ADVERB (L7)	participate	actively	MI ≥ 3	6.90
	treat	equally		4.88
	look	forward		4.37
	work	hard		3.91
	impact	deeply		3.79
	learn	more	MI < 3	2.28
	know	how		2.21

It could be seen that with the same nodes and different collocates, the MI scores could be varied. In the students writings, as exhibited in the chart and the table, there were 19 word combinations included in the strong collocations classification and 16 word combinations in the lower MI scores group from L1. For the L3, three collocations were considered stronger than the other two collocations. In the L4 group, the numbers of collocations were equal which were 2 and 2. For L5, 23 collocations possessed MI scores of 3 and above, and 19 collocations with MI scores of below 3. In the L6 group, three word combinations

were included in the stronger collocations group and 2 were in the other group. Lastly, in the L7 group, there were 6 collocations with higher MI scores and 2 collocations with lower MI scores.

Nodes and collocates of strong collocations are less likely to be random (Hunston, 2002). For instance, the word 'solve' was significantly associated with the word 'problem' due to the meaning of the words which were closely related. In the noun-verb combinations, the verb showed the action of the subject as in 'student learns' or 'teacher teaches'. Those words were more likely to be used together than the combinations with lower MI scores.

On the other hand, collocations with MI scores lower than 3 were more likely to possess less significant association. This means that such combinations were more open with the nodes that could be paired with a wide variety of words. For instance, the word 'learn' could be paired with 'explanation/theory/concept'. Those combinations were considered weaker than 'learn' and 'lesson' since their cooccurrence was low. They were less likely to occur together.

As the data showed, students produced more words with strong collocations than random ones. This could happen due to the nature of master's students who were used to reading both academic and non-academic texts in English. Moreover, they were surrounded by various English language inputs. Being exposed to large English language input gave students the familiarity of word combinations cooccurring in the text. Accordingly, Noreillie and Peters (2023) discovered that language input could help students significantly gain collocational knowledge. Furthermore, the students could take how some words were used more naturally together which then formed stronger collocations.

CONCLUSION

From the study, it could be inferred that students' English lexical collocations repertoire was varied. The students' word combinations included varied nodes and collocates. However, despite the wide variety of lexical collocations used, some of the collocations were not considered strong due to the low MI scores.

In order to increase students' English lexical collocations repertoire, students' exposure to English language inputs is pivotal. Reading and listening to English language content could make students used to the use of words and word combinations in a natural way. Therefore, students could produce more natural word combinations in their writings and even their speech. That way, students could increase their strong collocation repertoire.

For future researchers interested in conducting a study within a similar scope, some adjustments could be made. Larger data from the students' corpus could be used in order to find a larger variety of collocations. Moreover, differences and similarities could be discovered by conducting a comparison of the students' corpus and another English corpus. The mean of collocation measurement could also be added e.g., t-score and log-likelihood.

ACKNOWLEDGMENTS

We express our gratitude to the English Education Master's Programme of Sanata Dharma University, Yogyakarta for providing us with constructive support and facilities.

REFERENCES

- Alfiandita, O. L., & Ardi, P. (2020). Collocations in the English LKS Books for Senior High School Students. *IJELTAL (Indonesian Journal of English Language Teaching and Applied Linguistics)*, 4(2). <https://doi.org/10.21093/ijeltal.v4i2.463>.
- Alsayed, F. (2020). The Effects of Semantics in the Language Development of English as a Foreign Language (EFL) Learners. *International Journal of Learning and Development*, 9(4). <https://doi.org/10.5296/ijld.v9i4.15735>.
- Bartsch, S. (2004). Structural and functional properties of collocations in English: A corpus study of lexical and pragmatic constraints on lexical co-occurrence. Tübingen: Gunter Narr.
- Benson, M., Benson, E., & Ilson, R. (1986). *The BBI combinatory dictionary of English*. Amsterdam and Philadelphia: John Benjamins.
- Biber, D., & Reppen, R. (Eds.). (2015). *The Cambridge handbook of English corpus linguistics*. Cambridge University Press.
- Bui, I.-I.; C. | T. (2021). The Role of Collocations in the English Teaching and Learning. *International Journal of TESOL & Education*, 1(2).
- del Pilar Agustín-Llach, M. (2019). The impact of bilingualism on the acquisition of an additional language: Evidence from lexical knowledge, lexical fluency, and (lexical) cross-linguistic influence. *International Journal of Bilingualism*, 23(5). <https://doi.org/10.1177/1367006917728818>.
- Hunston, S. (2002). *Corpora in Applied Linguistics*. Cambridge: Cambridge University Press.
- Kang, B. M. (2018). Collocation and word association: Comparing collocation measuring methods. In *International Journal of Corpus Linguistics* (Vol. 23, Issue 1).
- Karjo, C. H., & Metta, E. (2019). The Translation of Lexical Collocations in Undergraduate Students' Theses' Abstract: Students Versus Google Translate. *Lingua Cultura*, 13(4). <https://doi.org/10.21512/lc.v13i4.6067>
- Lubis, S. (2020). Critical Study on Collocation as Source of Translation Unacceptability: Indonesian Students' Experiences. *Current Research in Education and Social Studies* Vol. 4, 97-106.
- Nation, I. S. P. (2001). *Learning vocabulary in another language*. Cambridge: Cambridge University Press.
- 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).
- Reynolds, B. L., & Teng, M. F. (2021). Innovating teacher feedback with writing activities aimed at raising secondary school students' awareness of

- collocation errors. *Studies in Second Language Learning and Teaching*, 11(3). <https://doi.org/10.14746/ssllt.2021.11.3.6>
- Sinclair, J. (1966). Beginning the study of lexis. In C. Bazell, J. Catford, M. A. K. Halliday, and R. Robins (eds.), *In memory of J. R. Firth*, 148–162. London: Longman.
- Sukur, S. G. (2020). The Impact of Disregarding Formulaic Competence in Developing Senior High School English Textbooks in Indonesia. *Journal of Languages and Language Teaching*, 8(2). <https://doi.org/10.33394/jollt.v8i2.2521>.
- Szudarski, P. (2017). *Corpus Linguistics for Vocabulary: A Guide for Research* (1st ed.). Routledge. <https://doi.org/10.4324/9781315107769>.
- Tanihardjo, J. (2018). Indonesian EFL Students' Production of English Lexical Collocation in Writing. *Journal of English Language and Culture*, 7(2). <https://doi.org/10.30813/jelc.v7i2.1025>.
- Trng, V. T., & Thao, T. T. (2021). A study on the use of English collocation in writing by students at Thai Nguyen University. *International Journal of Social Science and Human Research*, 4(5), 1044-1049.
- Umair, H. M., Imran, M., & Sarwat, S. (2023). An Analysis of the Use of Collocations in English Essay Writing at Undergraduate Level. *Pakistan Languages and Humanities Review*, 7(1), 43-54.
- Vu, D. V., Noreillie, A.-S., & Peters, E. (2023). Incidental collocation learning from reading-while-listening and captioned TV viewing and predictors of learning gains. *Language Teaching Research*. <https://doi.org/10.1177/13621688221151048>.
- Wongkhan, P., & Thienthong, A. (2021). EFL Learners' Acquisition of Academic Collocation and Synonymy: Does their Academic Experience Matter? *RELC Journal*, 52(3). <https://doi.org/10.1177/0033688219895046>.