

# COLLOCATIONAL COMPETENCE AMONG MALAYSIAN UNIVERSITY STUDENTS AND THEIR SOURCES OF ERRORS

Noranisa Afiqah Mohd Rubani <sup>1</sup>

Farrah Diebaa Rashid Ali <sup>2\*</sup>

Zuraidar Badaruddin <sup>3</sup>

<sup>1</sup> International Islamic University Malaysia

<sup>2</sup> International Islamic University Malaysia

Email: farrahdiebaa@iium.edu.my

\*Corresponding author

<sup>3</sup> Universiti Teknologi MARA

## Article history

**Received date** : 11-6-2023

**Revised date** : 12-6-2023

**Accepted date** : 25-7-2023

**Published date** : 15-8-2023

## To cite this document:

Mohd Rubani, N. A., Rashid Ali, F. D., & Badaruddin, Z. (2023). Collocational competence among Malaysian university students and their sources of errors. *Journal of Islamic, Social, Economics and Development (JISED)*, 8 (55), 349 – 361.

---

**Abstract:** *This study investigated Malaysian ESL learners' level of collocational competence and the source of their collocational errors. The subjects were 30 students from a local university in Malaysia. The research instrument was a Self-Completion Test (SCT) adopted from Huang (2001) that deliberates students' knowledge of four types of lexical collocations: free combinations, restricted collocations, and figurative and pure idioms. The result showed that those participants could answer the questions from free combination category better than other categories because that category is the easiest category among all. On the other hand, the most difficult and challenging category is found to be the pure idioms category. For restricted collocations and figurative idioms, the respondents performed quite equally well in these both categories. Overall, the deviant answers given by respondents signified that Malaysian ESL learners are having inadequate knowledge of English collocations.*

**Keywords:** *Collocation Competence, Malaysian University Students.*

---

## Introduction

Collocation is a kind of multiword expression that includes fillers, functional expressions, idioms, proverbs, and standardized phrases (Boers, Eyckmans, Kappel, Stengers & Demecheleer, 2006). It is built of recurrent alliances of words that often happen to occur habitually and repeatedly such as *take a bath*, *take a picture* and they can be both flexible and fixed. First introduced by Palmer in 1925, it is known as formulaic language competence and has been proclaimed as one of the methods to differentiate between native speakers and non-native speakers) as the automation of collocations formation actually assists the native speakers to deliver their messages better (Farrokh, 2012). Since non-native English language learners are not as exposed to this automation, they often make recognizable non-native errors in their utterance thus making this multiword unit is an important part of both language use and language learning (Ellis, Simpson-Vlach & Maynard, 2008; Yamashita & Jiang, 2010).

In line with incremental learning theory, Webb, Newton and Chang (2013) and Chen and Truscott (2010) reported that both the native and non-native speakers learned new words or structures gradually or progressively as they encountered them and the same is how collocation is learned. The more the collocation is encountered, the more learners will be familiar with it. When they are familiar enough with the collocation, they will have the confidence to use it in their daily conversation as it is easier for them to speak more fluently and naturally by uttering those fixed expressions as one prosodic unit.

In foreign language learning, especially the English language, collocation has been considered one of the most problematic areas to be learned and mastered. Collocation is believed to be seen as arbitrary which makes it complicated and hard to be identified. It seems to be one of the insurmountable obstacles for second-language learners to attain native-like fluency (Jeensuk & Sukying, 2021; Farrokh, 2012).

Even though English is a second language in Malaysia, Musa, Koo and Azman (2012) reported that teachers in their teachings tend to give more attention to the grammar aspects of a language rather than the communicative aspects because of the prioritization of national examinations. This has drawn some attention to the collocational competence among Malaysian ESL learners as collocation is more synonymous with the communicative aspects of a language rather than the grammatical aspects.

As collocation is claimed to be important for native-like fluency and language competency and decreases foreignness (Al-Shammari, 2022; Jeensuk & Sukying, 2021; Farrokh, 2012; Goudarzi & Moini, 2012), this paper investigated the Malaysian ESL learners' level of collocational competence and sources of the collocational error made by them by adopting the Howarth' and Phoocharoensil's (2011) models.

## Literature Review

### Categorizations of Collocations

The categorization of collocation has been systematically sorted out by Benson, Benson and Ilson (1986) who divided it into two major groups which are lexical collocations and grammatical collocations. Lexical collocations consist of the types of word class such as nouns, adjectives, verbs and adverbs, whereas, grammatical collocations are made up of dominant words including all the word class mentioned and combined with grammatical elements. The grammatical elements include preposition or grammatical structure such as infinitive or clause (Bahns & Eldaw, 1993).

To further categorize lexical collocations, Howarth (1996) introduced a collocational continuum. Howarth's (1996) model categorizes collocations into (1) free combinations, (2) restricted collocations, (3) figurative idioms and (4) pure idioms. For free combination, it is interpreted from the literal meaning of individual components contained in the combination (e.g. *watch the television*). For restricted collocation, the words used are more confined in the compositional elements exception and often it has one segment used as a part of a particular context (e.g. *pay a visit*). For figurative idioms, overall, the phrases have metaphorical meaning that people can somehow get the actual or literal interpretation (e.g. *paper tiger*). Whereas, for pure idioms, every sentence of pure idiom is considered as a single component which the meaning is completely unforeseeable and cannot be predicted by the represented components (e.g. *spill the beans*). The Howarth's model is commonly and widely used by researchers in investigating the level of collocational competence among English language learners (Huang, 2001; Wang & Shaw, 2008; and Farrokh, 2012).

Lewis (1997) classified those types of combined words into four more groups which are strong, weak, frequent, and infrequent. Strong and weak collocations are classified based on their persistency and limitation, whereas frequent and infrequent collocations are classified based on their oftenness of co-occurrence in a corpus. Some examples of strong collocations are *drug addict* and *drink beer*. As their level of persistency and limitation is high, they are perceived as firmly connected expressions which work like single word. While for the weak ones, such as *good job* or *great work*, they are combined with two common words. Their level of fixedness and restriction is low because they are very much likely to be occurred with other set of words as well. Lewis (1997) added that those collocations can be in any combination of the category; *strong and frequent*, *strong and infrequent*, *weak and frequent* and *weak and infrequent*.

### Sources of Collocational Errors

Researchers have found that the errors ESL learners made mostly resulted from analogy, overgeneralization, paraphrase, interlingual transfer, intralingual transfer, and shortage of collocational knowledge (Bahns & Eldaw, 1993; Farrokh, 2012). Bahns and Eldaw (1993) and Huang (2001) discovered that many collocational errors of ESL learners are sourced from their first language intervention. Biskup (1992) said that the broader the semantic area of an item, the more it might spark off the first language interference errors. Huang (2001) also reported that the more synonyms that an item has, the harder it is for learners to bring out restrained collocation. It has also been identified that cultural stereotypes also affect learners' production of second language collocation. Learners who do not know a particular collocation tend to overcome it with certain strategies such as avoidance, transfer, use of synonym, and paraphrasing the collocational phrase (Farghal & Obiedat, 1995). These strategies, however, are believed to be the factors that lead English as a second language learners to make collocational errors in producing English collocations (Phoocharoensil, 2011). Among these strategies, first language transfer, synonymy, and repetition and overgeneralization have been identified as the most common strategies that lead the learners to make collocational errors.

ESL learners often produce collocations modelled upon their first language structures. They believe that their native language collocation can be translated word-for-word in the target language. This makes them rely more on their first language collocational knowledge rather than learning and mastering the ones in the second language. Heavy reliance on first language collocational structure leads the learners to make collocational errors in the second language.

The second strategy, synonymy or also known as adopting the analogy is often used by learners of lower second language proficiency level. At this stage, the learners tend to group few words together believing that they are synonymous. Since the learners perceived the words as synonymous, they erroneously believe that these words share the same grammatical collocations (Nation, 2001). As for repetition and generalization, some learners tend to extend the use of the few collocations that they have learned in all possible contexts. Ironically, the learners reported that they avoided using unfamiliar new collocations to avoid making mistakes (Zughoul & Abdul-Fattah, 2001; Phoocharoensil, 2011).

### **Measuring Collocational Competence**

According to Pei (2008), there are two types of knowledge in collocation which are receptive and productive. Both kind of knowledge require different types of collocational test to see whether a learner has acquired the knowledge or not. Receptive knowledge is tested by giving the learners multiple choice questions, whereas for productive knowledge, cloze test and translation tasks are used to see if the learners have acquired sufficient collocational knowledge as expected (Pei, 2008). Usually in a cloze test, participants are expected to fill in the blanks with collocations, meanwhile for the translation task, participants are required to translate collocations from their first language into the target language. Context of the sentences will be provided for both tasks. By running these two tests, through the result, researchers can see if learners are able to produce an acceptable and valid collocation or not, and to see if they have really acquired the target collocation or not.

## **Research Methodology**

### **Participant**

30 third and fourth year undergraduate Malaysian university students of various programmes participated in this study. Their first language is Malay and they learn English formally from primary one to tertiary level. All of them agreed that English is their second language with majority of them claimed that they do communicate in English at home and with their immediate friends and lecturers. They are very much exposed to English songs, television shows, movies, as well as English reading materials. They are aware of what collocation is and have heard or read about it.

### **Measuring the Collocational Competence: The Test and The Analysis**

The research tool was a self-patterned Simple Completion Test (SCT) which was adopted from Huang (2001). It assessed the participants' knowledge on the four types of collocations which are (1) free combinations, (2) restricted collocations, (3) figurative idioms, and (4) pure idioms. In the test, there are 40 items in the form of free-response. Each collocation category consists of ten items. There are two to three sentences in each item. In those items, there are a certain context or theme provided on the particular collocation or idiom. The theme includes verb, adjectives, noun with food and noun with animal. Participants were requested to fill in the blanks with a suitable collocation or idiom. The sentences in the SCT were mostly adapted from Booker's Longman Active American Idioms (1994). The examples of questions asked for each type of lexical collocations can be found in the Table 1 below.

**Table 1: Samples of question asked in the SCT**

Types of Collocation	Samples of question
Free combinations	A2. Tina and her parents don't like lemons because they are too sour. So they _____ apples instead. (fill in the blank with an appropriate verb)
Restricted collocations	B6. Remember to put cream into Jenny's coffee. Otherwise, she'll get angry because she never drinks _____ coffee. (fill in the blank with an appropriate adjective).
Figurative idioms	C4. A lazy person always gives the excuse that working is not his cup of _____. (fill in the blank with an appropriate noun about food).
Pure idioms	D8. Ten years ago, the streets in Chicago were dirty and public services were awful. The city had really gone to the _____. (fill in the blank with an appropriate noun about animals).

The participants were also encouraged to take educated guesses if they were uncertain of the answer and to never leave any question unattended. The first research question was answered by identifying the frequency of correct answers given by respondents in the SCT. The second research question was answered by categorising the wrong answers given by respondents into three types of sources of collocational errors adapted from Phoocharoensil (2011) which are native language transfer, synonymy, and overgeneralization.

After the test was completed by the respondents, all the sheets were collected and the answers are analysed using both qualitative and quantitative methods. Both correct and wrong answers are marked simultaneously. During the marking, special consideration was given for answers of the items under the categories of free combination and restricted collocation. Answers that showed correct choice of words but had wrong inflections were considered as correct. For example, *These boys and girls don't \_\_\_\_\_ orange juice. They prefer something special, like pineapple juice or punch.* For this question, answers like *drink, drinks* and *drinking* were all considered as correct because the main focus of this category is to see whether the respondents can pick the correct word to be collocated in that sentence. The chosen verb *drink* can collocate well with the noun *orange juice* in the sentence. Thus, the inflectional errors like numbers of nouns or verbs were neglected.

However, this kind of consideration had a slightly different application to the items under the categories of figurative idioms and pure idioms. In *In this poor country, many children take small jobs to help their families \_\_\_\_\_ the wolf from the door, although* the item is under the category of idiom, the answers *keep* or *kept* are all considered as correct. This is because the choice of verb *keep* was correct despite the error in the verbal inflection as the error does not affect the meaning of the idiom. Nevertheless, the principal cannot be applied to the following example - *Ten years ago, the streets in Chicago were dirty and public services were awful. The city had really gone to the \_\_\_\_\_. But now it's much better.* For this kind of situation, the only correct answer for the item is *dogs*. Other alternative of the word like *dog* does not fit the idiom. This is because, the collocation for this item is completely frozen. The plurality or singularity in the idiom cannot be changed.

For quantitative analysis, the number of correct answers, wrong answers and blank responses were counted. Then, descriptive statistics were generated to compare the respondents' performance in each category and to observe the degrees of difficulty of those different categories. Blank responses left by the respondents were also counted as it indicated the level of hardness perceived by the respondent. The number of variations to the respondents' wrong answers were also an indicator to the items' level of hardness. This suggested that the more challenging the item was, the more variety of answers were given by respondents.

For qualitative analysis, the wrong answers given by respondents were analysed and the sources of collocational errors made by them were recognized and put into categories. The categories of sources of collocational errors were adapted from a listing made by Phoocharoensil (2011). This categorization of the sources of their collocational errors discovered the reason why they did those errors. In addition, after they completed the test, the researchers asked a few questions to the respondents verbally about the level of difficulty of the test and why are they difficult. This is to know directly and personally the reason why they cannot answer those questions correctly.

## Findings and Discussions

### Levels of Collocational competence

Below are the charts representing the results prevailed from the collocational test. They were divided into three different groups which are correct responses, wrong responses, and blank responses. Blue columns represent correct responses, orange columns represent wrong responses, and grey columns represent blank responses. The letter 'Q' on the x-axis means 'Question' which represents the questions in each section. Q1 for Question 1, Q2 for Question 2, Q3 for Question 3 and so on. The total of responses for each question is 30 responses by 30 respondents.

### Free Combination Collocations

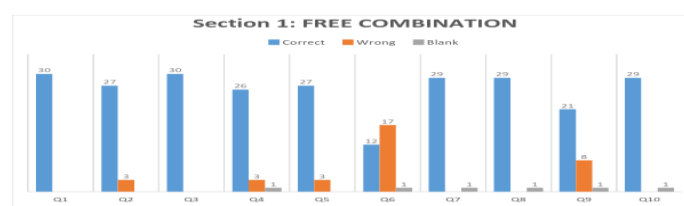


Figure 1: Respondent's responses to Free Combination section

From the Figure 1 above, it can be seen that the respondents had problems answering two questions which are Question A6 and Question A9. The factor is discussed in the later section of this paper. There are two questions that the respondents managed to answer perfectly without any mistake which are Question A1: *Those boys and girls don't \_\_\_\_\_ orange juice. They prefer something special like pineapple juice or punch*, and Question A3: *Today is Sunday. Do you want to \_\_\_\_\_ there to see some rare animals?* For QA1, all 30 of the respondents gave suitable answers to be collocated with the word orange juice. The answers given were *drink*, *like* and *prefer*. For QA3, all of the respondents answered the question with the same word which was *go*.

From the findings above, it shows that the respondents were able to answer most of the questions asked correctly. Not only that, this category also recorded the highest number of

correct answers given by respondents compared to other categories. This is similar to the result presented in the previous study by Huang (2001) which found that this category is the easiest type in the collocational continuum. Respondents were able to answer most of the questions correctly because for this category, the degree of flexibility in the lexical combination is high. They were also not really restricted to choose the suitable words to be put in the blank due to this characteristic.

### Restricted Collocations

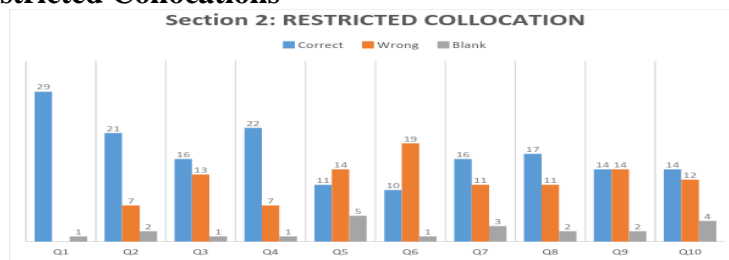


Figure 2: Respondent's responses to Restricted Collocation section

For this section, compared to the first section, respondents were seen to start to give more wrong answers and blank responses in every question. It can also be seen that QB5: *This peach is sweet and delicious. Would you like to \_\_\_\_\_ a bite?* and QB8: *If our hen could \_\_\_\_\_ gold eggs like the one in the fairy tale, we would become rich.* are the two questions that recorded the highest number of correct answers given by respondents. The questions are as follow;

For QB5, 29 of the respondents answered the question with the word *have* and *take*. Both answers are correct and acceptable for the lexical combination. For QB8, there are only two acceptable answers for this question which are *lay* and *produce*. 22 of respondents answered the questions correctly and the rest of them did not. Some of the deviant answers were *hatch* and *give*. At first glance, the word *hatch* seemed suitable to be the answer for this question but it actually does not fit the context of the sentence. Thus, the word cannot be accepted.

From the findings above, it shows that the number of correct answers given are decreasing and the number of wrong answers and blank responses are increasing. This is similar to Huang's (2001) findings. Since the degree of flexibility of lexical combination in this category is lower than the flexibility degree of the previous category, the word choice is more restricted and limited. Thus, many researchers believe that this category is the most prominent category to be taught and learned by English practitioners (Farghal & Obiedat, 1995; Howarth, 1998).

### Figurative Idiom

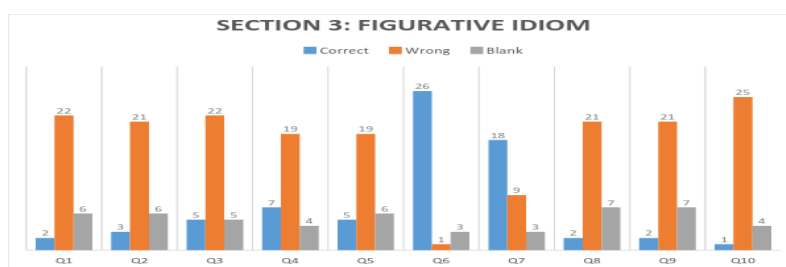
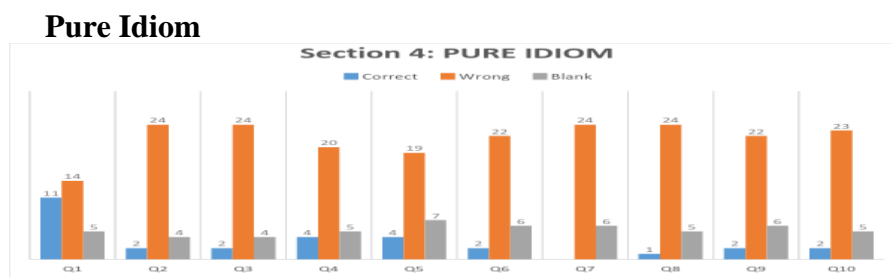


Figure 3: Respondent's responses to Figurative Idiom section

For this section, the answers are not flexible. From the chart, it can be seen clearly that the correct answers given are extremely decreasing and the wrong answers as well as blank responses given are extremely increasing. This was where the respondents felt that the level of difficulty of the collocational continuum was increasing while completing the test. There were only two questions that the respondents managed to give correct answers with an encouraging number. The questions are Q6: A lazy person always gives the excuse that working is not his cup of \_\_\_\_\_ and Q7: D2. He owns the biggest factory in this small town. Though he is well-known in the town, he's just a big \_\_\_\_\_ in a small pond.

For Q6, the answer is *tea*. Almost all of the respondents were able to answer that question correctly because most of them are familiar with that idiom. For Q7, the answer is *fish*. The reason why the respondents can answer this question correctly is because of the clue word 'pond' in the sentence of the question. It can be seen that the word pond is somehow synonymous with the word fish. However, there were also other answers that cannot be accepted such as *carp*, *frog* and *tadpole*. This happened because the students are not familiar with the idiom. In fact, for that one question that respondents were able to answer correctly, it was because they are familiar with that lexical combination.



**Figure 4: Respondent's responses to Pure Idiom section**

For this section, the characteristic of the idiom is more frozen and fixed than the figurative idiom. That is because for the pure idiom, the lexical combination is irregular and unpredictable. It must be remembered and memorized by learners. That is why, it can be seen from the column chart that almost all of the respondents were not able to give correct answers to the questions provided. In average, respondents can only answer three questions correctly for this section. Judging from the result obtained, this section is believed to be the most difficult section of all collocational continuum. This supports the statement by Sipayung and Saragih, (2023), Huang (2001) and Durrant and Schmitt (2010) that pure idiom is the most difficult category in the collocational continuum.

### Sources of Collocational Errors

There are four sources of collocational errors identified in this study which are native language transfer, synonymy, overgeneralization and culture.



## Native Language Transfer

**Table 2: Native Language Transfer**

Types of Collocational Errors	Example
Restricted Collocation	<i>B4. I don't like <b>bitter</b> tea because it upsets my stomach. Also, its dark colour looks terrible.</i>
Figurative Idiom	<i>D4. To become a doctor, you have to work hard at medical school. It's a <b>owl</b>'s life because you'll have to study every day for many years.</i>
Pure Idiom	<i>A12. In this poor country, many children take small jobs to help their families <b>throw</b> the wolf from the door.</i>

For native language transfer, it is the collocational errors that respondents made when they evidently relied on their first language vocabulary or collocational knowledge. In this case, the first language of all respondents is Malay language. From the collocational test conducted for this research, the number of wrong answers caused from the native language transfer can be found in every section. This showed that this strategy is indeed used by most Malaysian ESL learners and they tend to make the collocational errors sourced from this strategy. This is in line with Sipayung and Saragih (2023), Boonraksa and Naisena's (2022) findings on Indonesian and Thai EFL learners.

For B4, some students answered it with the word 'bitter' and this is believed to be very much influenced from their first language which 'bitter' means *pahit* in Malay. The Malays use the word *pahit* for coffee and tea quite frequently. In addition to the above, the word does connote an unpleasing feeling or taste about the drink. Because the word 'bitter' is equivalent to their first language word *pahit*, some respondents came up with this word as their answer.

For the D4 question, although the article before the blank space was the word 'a', and grammatically wrong to be partnered with the word 'owl', still some of the respondents chose this word as their answer to this question. In Malay, 'owl' means *burung hantu*. *Burung hantu* is believed to stay awake at night without having any sleep. It connotes the common habit of students who have to burn their midnight oil on studying. This word is equivalent to the context of the sentence that medical students will always have to work hard and maybe cannot get proper sleep at night because they have to study. That is why some of the respondents chose this word as their answer to this question.

For the A12 question, this is the most difficult part in the collocational test where if respondents did not know the answer to that question, they totally cannot guess for the answer. However, some of them used this native language transfer strategy as their alternative to come up with an answer that they think suitable for this question. Literally, 'throw' means *buang* or *baling* in Malay. They chose the word 'throw' as they assumed that the sentence wanted to say 'to toss the dog away from the door'. It sounds right and makes sense. However, that answer is completely wrong. The meaning of that idiom has nothing to do with throwing or tossing away any wolf or anything. Thus, the other lexical combination that is needed in order to complete the idiom cannot be guessed at all.

## Synonymy

**Table 3: Synonymy**

Types of Collocational Errors	Example
Restricted Collocation	A6. Not all trees <b>breed</b> fruit. For example, an apple tree does, but a pine tree doesn't.
Figurative Idiom	C4. A lazy person always gives the excuse that working is not his cup of <b>coffee</b> .
Pure Idiom	C5. I told all my friends I'd help them with their English, but I got a lot of answers wrong! Now I've got <b>cake</b> on my face.

Looking at the above examples, it is observed that the respondents tried to find a synonym for the collocation that they were not to guess correctly. For question A6, there are three answers that can be marked as correct which are *bear*, *produce*, and *have*. However, 19% of the respondents answered *breed* because they cannot think of the correct word and came up with another word that carries similar or almost similar meaning as the three words mentioned. Nevertheless, the word *breed* cannot be accepted as the correct answer because *breed* is never used to represent tree or fruit.

For question C4, the correct answer was *tea*. However, there was 0.33% of the respondents who answered *coffee*. Although *coffee* is not literally the synonym for the word *tea*, but it carries the same meaning as the word *tea* in terms of the context. *Tea* is a type of a drink as well as *coffee*. Respondents who were unable to think of an answer for that question were choosing another type of drink to be put in that sentence. They got the clue that the sentence required a type of drinks to be put in the blank space in the sentence from the word 'a cup of'. Although *tea* and *coffee* are just the same which are types of drinks, and are in the same context, the word *coffee* cannot be considered as correct as it is not the right collocation to that idiom and it sounds so unnatural for English speakers to use that word in that sentence. The same goes for egg and cake in C5.

## Overgeneralization

**Table 4: Overgeneralization**

Types of Collocational Errors	Example
Restricted Collocation	A8. If our hen could <b>hatch</b> gold eggs like the one in the fairy tale, we would become rich.
Figurative Idiom	D3. Mr Klein seems to have a lot of power in his company, but actually his wife is the real boss. He is just a <b>paper horse</b> .
Pure Idiom	D6. We got the good news from Helen. Since she is the manager's daughter, we all believed that she got it from the <b>tiger's</b> mouth.

For the category of overgeneralization, it is the strategy that the respondents attempted by generalizing the answer of those questions when they were unable to give correct words for those lexical combinations. This is solely based on their assumption referring to the context of the sentence and here we can see whether they can make educated guesses or not.

For question A8, the word 'hatch' was chosen by 17% of the respondents as their final answer. Whereas, there were only two acceptable answers for this question which are 'lay' and

‘produce’. It cannot be denied that in English, the word ‘hatch’ is very synonymous to the word ‘egg’ when the word ‘hen’ is present. Because of that, those respondents thought that that word is applicable to all sentences that have the word ‘egg’ with ‘hen’. However, the answer is wrong. It is a clear example of an overgeneralization.

For the D3 question, the correct answer was ‘tiger’. However, 36.67% of respondents answered that question with the word ‘horse’. In English idiom, ‘horse’ is a common animal to be used as a subject or an object in a metaphor. They overgeneralized the usage of the word ‘horse’ and assumed that that word is applicable to any idiom.

## Culture

**Table 5: Culture**

Types of Collocational Errors	Example
Restricted Collocation	<i>A7. Let's <u>give</u> a toast to the host and wish him good health and many years of happiness.</i>
Figurative Idiom	<i>D5. Mother won't let Bill do the dishes because he may break the plates. He's like a <u>rat</u> in a China shop.</i>
Pure Idiom	

The researcher of this paper found out one more source of collocational error made by respondents in the collocational test. It is culture. This category was not included in the list made by Phoocharoensil (2011), however the same was reported by Al-Shammari (2022). Because this category cannot be put in any sources of collocational error listed above, it has to be listed as a new category separated from other categories. In this category, respondents made collocational error influenced by cultural factor. They are whether unfamiliar with the lexical combinations or they found those collocations to be uncommon because they have never encountered it or they rarely see it in their daily lives.

The result showed that there was no cultural factor of collocational error found in the pure idiom category. Hence, only two examples are listed – from restricted collocation and figurative idiom category. For question A7, the only acceptable and correct answer is ‘make’. There is 12% of respondents answered that question with other words than ‘make’. It has been discovered that the wrong answers given by them was influenced by culture because ‘toasting’ culture is not applicable in the Malay Muslim community.

For the question D5, the correct answer is ‘bull’. According to the result obtained from the test, only one respondent was able to answer this question correctly. The rest, they just guessed and wrote any animal that they think suitable for the sentence context. This unfamiliarity of the lexical combination was also discovered to be caused by cultural factor. Most of them were unable to answer the question because technically, they are not Chinese and this makes them unfamiliar with China shop.

## Conclusion

As a conclusion, it can be concluded that the students’ levels of collocational competence varied depending on the types of collocations. It was observed that higher level of competence can be associated to the free combination collocations, while most of the students are still struggling with the restricted collocation, figurative idiom and pure idiom types of collocational

competence. It is also found that native language transfer, synonymy, overgeneralization and culture are the main source of collocational errors for these participants in this study.

## References

- Al-Shammari, A.H. (2022). The importance of Teaching Collocations in the development of Kuwait ESL/EFL learners' reading comprehension. *Annals of the Faculty of Arts, 50(Oct-Dec)*. 1-59
- Bahns, J. & Eldaw, M. (1993). Should we teach EFL students collocation? *System, 21(1)*, 104-114.
- Benson, M., Benson, E. and Ilson, R. (1986). *The BBI combinatory dictionary of English: A guide to word combinations*, Amsterdam: John Benjamins
- Biskup, D. (1992). L1 influence on learners' renderings of English collocations: A Polish/German empirical study. In P.J.L. Arnaud & H. Bejoint (Eds.), *Vocabulary and applied Linguistics* (pp.85-93). London: MacMillan Academic and Professional Ltd.
- Boers, F., Eyckmans, J., Kappel, H., Stengers & Demecheleer, M. (2006). Formulaic sequences and perceived oral proficiency: Putting a lexical approach to the test. *Language Teaching Research, 10*, 245-261
- Boonroksa, T. & Naisena, S. (2022). A study on English collocation error of Thai EFL students, *English Language Teaching, 15(1)*.164-177
- Durrant, P.& Schmitt, N. (2010). Adult learners' retention of collocations from exposure. *Second Language Research, 26(2)*, 163-188
- Ellis, N.C.,Simpson-Vlach, R. & Maynard, C. (2008). Formulaic language in native and second language speakers: Psycholinguistics, corpus linguistics and TESOL. *TESOL Quarterly, 42 (3)*, 375-396
- Farghal, M. & Obiedant, H. (1995). Collocations: a neglected variable in EFL. *IRAL 33(4)*, 315-333
- Farrokh, P. (2012). Raising awareness of collocation in ESL/EFL classrooms. *Journal of Studies in Education, 2(3)*, 55-74
- Goudarzi, Z. & Moini, M.R. (2012). The effects of input enhancement of collocations in reading on collocation learning retention of EFL learners. *International Education Studies, 5(3)*, 247-258
- Howarth, P. (1996). Phraseology and second language proficiency. *Applied linguistics, 19*. 45-72
- Huang, L.S. (2001). Knowledge of English collocations: An analysis of Taiwanese EFL Learners. In: Luke, C. and B. Rubrecht, Eds. *Texas Papers in Foreign Language Education: Selected Proceedings from the Texas Foreign Language Education Conference, 2001*. Volume 6(1, Fall 2001).
- Jeensuk, S. & Sukying, A. (2021). An investigation of high school EFL learners' knowledge of English Collocations. *Journal of Applied Linguistics and Language Research, 8(1)*, 90-106
- Lewis, M. (1997). *Implementing the lexical approach: Putting theory into practice*, Hove, UK: Language Teaching Publications
- Musa,N.C., Koo, Y.L., & Azman, H. (2012). Exploring English language learning and teaching in Malaysia. *GEMA: Online Journal of Language Studies, 12(1)*, 35-51
- Nation,I.S.P (2001). *Learning vocabulary in another language*. Cambridge: Cambridge University Press
- Pei, C. (2008). Review of empirical studies on collocation in the field of SLA. *CELEA Journal, 31(6)*.

- Phoocharoensil, S. (2011). Collocational errors in EFL learners' interlanguage. *Journal of Education and Practice*, 2(3), 103-120
- Sipayung, R.W. & Saragih, E. (2023). Contextualizing EFL learner's proficiency in using English collocations. *Journal of English Education and Linguistics Studies*, 10(1). 1-25
- Wang, Y. & Shaw, P. (2008). Transfer and universality: Collocation use in advanced Chinese and Swedish learner English. *ICAME Journal*, 32. 201-232
- Webb, S