



**CORPUS-BASED ANALYSIS OF THE USAGE PATTERN OF HOMONYMS  
AMONG MALAYSIAN ESL LEARNERS**

**LAW HUI QIN**

**19AAB06848**

**SUPERVISOR: MS. DEEPA A/P VISVANATHAN**

**UALZ 3023 - FYP2 REPORT**

**SUBMITTED IN  
PARTIAL FULFILMENT OF THE REQUIREMENTS  
FOR BACHELOR OF ARTS (HONS) ENGLISH LANGUAGE  
FACULTY OF ARTS AND SOCIAL SCIENCE**

**MAY TRIMESTER 2022**

CORPUS-BASED ANALYSIS OF THE USAGE PATTERN OF HOMONYMS AMONG  
MALAYSIAN ESL LEARNERS

LAW HUI QIN

A RESEARCH PROJECT  
SUBMITTED IN  
PARTIAL FULFILMENT OF THE REQUIREMENTS FOR  
THE BACHELOR OF ARTS (HONOURS) ENGLISH LANGUAGE  
FACULTY OF ARTS & SOCIAL SCIENCES  
UNIVERSITI TUNKU ABDUL RAHMAN

MAY 2022

## ACKNOWLEDGEMENTS

I am grateful to be able to complete this thesis with the guidance and assistance of a few individuals that helps me throughout this research. Therefore, I would like to express my deepest gratitude to my supervisor, Ms. Deepa a/p Visvanathan for her undivided support, guidance, useful information, and advice given that have greatly helped me in conducting the research and producing the thesis writing. Besides, I would also wish to express my sincere gratitude to the university lecturers, Mr. Mohd Amir Izuddin bin Mohamad Ghazali, and Ms. Tan Swee Mee who have gave permission for me to collect students' essays as the data of this research study.

Apart from that, I am deeply grateful for having parents that are supportive, spread their positive vibes to me that has encouraged me to put more effort while doing this research. Meanwhile, I am thankful to have friends that are always by my side giving strong moral support and encouragement that keeps me working continuously with positivity for my thesis. Once again, I am very thankful to have everyone involving and helping me as without your kind assistance, it would be impossible for me to complete and submit this thesis on time.

### Approval Form

This research paper attached hereto, entitled “Corpus-Based Analysis of the Usage Pattern of Homonyms among Malaysian ESL Learners” prepared and submitted by “LAW HUI QIN” in partial fulfilment of the requirements for the Bachelor of Arts (Hons) English Language is hereby accepted.

---

Supervisor

Date

Supervisor's name: Ms. Deepa a/p Visvanathan

## ABSTRACT

Learners usually deal with words that have multiple meanings, and similar pronunciation and spelling when they acquire English vocabulary. These words are homonyms that often confuse learners due to their ambiguous nature, similarity in orthography and pronunciation. Most learners faced difficulties distinguishing several types of homonyms and knowing the usage of homonyms in different contexts. This study analyses the types and usage patterns of the homonyms used by Malaysian undergraduates in their essay writing. The study adopted a corpus-based method to conduct the research, and a total of 145 essays were collected and digitised to generate a learner corpus. In the learner corpus, there are 8098 homonyms found and classified into homophones, homographs, lexical homonyms, and lexico-grammatical homonyms. The usage patterns of homonyms are examined through common usage, misuse pattern, and collocational pattern based on different homonym categories. The result showed that the word classes of most homonyms determined the word sense used in a particular context, except for lexical homonyms. Apart from grammatical class, context also played a significant role in disambiguating multiple senses of homonyms. Based on the usage patterns, learners experienced biases, confusion, L1 interference, and misunderstand senses of homonyms while building sentences. The corpus highlighted grammatical class, meanings, and context are the key elements used to address these issues. This paper concluded that corpus could be a practical tool to learn several homonym categories and their usage in different contexts.

## DECLARATION

I declare that the material contained in this paper is the end result of my own work and that due acknowledgement has been given in the bibliography and references to all sources be they printed, electronic or personal.

Name : LAW HUI QIN

Student ID: 19AAB06848

Signed : 

Date : 7<sup>th</sup> September 2022

## TABLE OF CONTENTS

	Page
ABSTRACT	i
DECLARATION	ii
TABLE OF CONTENTS	iii
LIST OF TABLES	iv
LIST OF FIGURES	v
LIST OF ABBREVIATIONS	vi
CHAPTERS	
1 INTRODUCTION	
1.1 Background of the Study	1
1.2 Problem Statement	5
1.3 Research Objective	9
1.4 Research Question	9
1.5 Significance of the Study	9
1.6 Limitation of the Study	11
2 LITERATURE REVIEW	
2.1 Definition of Terms	13
2.1.1 Ambiguity	13
2.1.2 Sense Relations	13
2.1.3 Lexical Ambiguity	14
2.1.4 Homonym and Polysemy	14
2.2 Past Studies	15
2.2.1 Past Research in Malaysia	15
2.2.2 Past Research in Overseas	18
2.3 Theoretical Foundation of the Study	20
2.3.1 Usage-Based Theory	20
2.3.2 Katz's Semantic Theory	21
2.4 Conceptual Framework	23
3 METHODOLOGY	
3.1 Research Design	26

3.1.1 Corpus-Based Method	26
3.2 Sampling Selection	27
3.3 Instruments	28
3.3.1 Concordance software	28
3.4 Data Collection	28
3.5 Data Analysis	29
3.6 Ethical Considerations	30
4 FINDINGS & ANALYSIS	
4.1 Types of Homonyms	31
4.1.1 Homophones	31
4.1.2 Homographs	33
4.1.3 Lexical Homonyms	34
4.1.4 Lexico-Grammatical Homonyms	36
4.2 Usage Pattern of Homonyms	38
4.2.1 Common Usage of Homonyms	38
4.2.2 Misuse Pattern of Homonyms	48
4.2.3 Collocational Pattern of Homonyms	52
5 DISCUSSION & CONCLUSION	
5.1 Research Objectives	56
5.1.1 Research Objective 1	56
5.1.2 Research Objective 2	59
5.2 Recommendation for Future Research	65
5.3 Conclusion	68
REFERENCES	70
APPENDIX	
A Consent Form	83
B Consent Form	85



## LIST OF TABLES

Tables		Page
4.1	Homophones list	32
4.2	Homographs list	33
4.3	Lexical homonyms list	35
4.4	Lexico-grammatical homonyms list	36
4.5	Common usage of homophones 'affect' and 'effect'	39
4.6	Common usage of homophones 'advise' and 'advice'	40
4.7	Common usage of homograph 'present'	41
4.8	Common usage of homograph 'impact'	42
4.9	Common usage of lexical homonym 'live'	43
4.10	Common usage of lexical homonym 'get'	44
4.11	Common usage of lexico-grammatical homonym 'lead'	46
4.12	Common usage of lexico-grammatical homonym 'need'	47

## LIST OF FIGURES

Figures		Page
2.1	Diagram of the conceptual framework	23
4.1	Chart of the overall usage of all types of homonyms	38
4.2	Misuse lines of homophones ‘affect’ and ‘effect’ taken from the concordance software (AntConc)	48
4.3	Misuse line of homophone ‘palette’ taken from the concordance software (AntConc)	49
4.4	Misuse lines of lexical homonym ‘cover’ taken from the concordance software (AntConc)	49
4.5	Misuse lines of lexical homonym ‘live’ taken from the concordance software (AntConc)	50
4.6	Misuse line of lexico-grammatical homonym ‘lead’ taken from the concordance software (AntConc)	51
4.7	Collocate list of homophone ‘effect’ in the generated learner corpus	52
4.8	Collocate list of homograph ‘conflict’ in the generated learner corpus	53
4.9	Collocate list of lexical homonym ‘live’ in the generated learner corpus	54
4.10	Collocate list of lexico-grammatical homonym ‘lead’ in the generated learner corpus	55

## LIST OF ABBREVIATIONS

DDL	Data-driven learning
ESL	English as second language
KWIC	Key-Word-In-Context
L1	First language
UTAR	Universiti Tunku Abdul Rahman
WSD	Word sense disambiguation

## CHAPTER 1 – INTRODUCTION

### 1.1 - Background of the Study

English language is one of the commonly studied languages worldwide with around 1.5 billion people learning the language and more than half of the learners are non-native speakers (Noack & Gamio, 2015). For second or foreign language speakers to begin with their English language learning journey, they need to acquire four skills including reading, writing, listening, and speaking. Among these core skills, writing is relatively more complex than other language skills as the message conveyed by the writer must be clear and concise for readers' comprehension, and the words used must be appropriate in the context of writing (Gautam, 2019). Meanwhile, learners are required to master other language knowledge such as grammar, vocabulary, spelling, and punctuation that act as supportive roles in assisting them to write precisely (Setiani, 2019). In other words, writing involves a set of language processes that are fundamental for writers to organize their thoughts and transfer it into a coherent piece of work. Since writing plays a crucial role in transforming ideas into text, it is significant for learners to make their writing comprehensible for readers without committing errors that could lead to misunderstanding in the readers.

Although English language is quite popular among second or foreign language learners, it would be quite challenging for them to deal with English words that have multiple meanings while writing a text or an essay. In such circumstances, it is linguistically known as ambiguity when an expression has more than one interpretation in a sentence or context. Generally, words, phrasal verbs, or idioms are called lexemes that carry lexical meanings and keep as concepts in learners' mental lexicon before they interpret the utterance as ambiguous (Loebner & Hooimeijer, 2013).

The ambiguity that appears at the lexical level is known as lexical ambiguity which can be mainly categorised into syntactic lexical ambiguity and semantic lexical ambiguity (Small et al., 1988). Syntactic lexical ambiguity refers to a word that consists of different word classes. For instance, swim refers to a noun when it is described as an activity, it also can be a verb when someone performs the action. On the other hand, semantic lexical ambiguity can be subdivided into polysemy and homonym. Polysemy is defined as words that have multiple related meanings (Small et al., 1988). For example, the word cut in these two sentences, the butlers cut the meat, and the boys cut a hole on the paper conveyed the same meaning of using a sharp tool to perform the action of cutting. While homonyms refer to words that have several unrelated meanings such as the word bat in these sentences, he hits the ball using a bat and the bat eats fruit at night. The first sentence refers to the wooden object used to play ball games, while the second sentence is referring to an animal. Both polysemy and homonyms could be difficult for non-native learners to distinguish their differences, therefore it is vital for them to acquire this semantic knowledge while learning English language.

Alekse (2017) stated that etymology of words could be an approach to differentiate polysemy and homonym because polysemy is recorded as a single entry based on one origin, while homonym is recorded separately according to different origins. However, this way is not effective to identify their differences when the origins of some words are not available, and the present form of the words could change over time (Palmer, 1976, as cited in Alekse, 2017). Apart from checking the historical background of words, the relatedness or unrelatedness of meanings would show the distinction between polysemy and homonym where polysemy refers to related sense, while homonym refers to unrelated sense (Klepousniotou, 2002). Nevertheless, this method is only used to indicate a spectrum from “pure” polysemy to “pure” homonym; there is no clear division between homonym and polysemy (Lyons, 1977, as cited in Klepousniotou, 2002). As homonyms appear to be more diverse than polysemy in terms of

having various spellings and different orthography, hence the present study would like to focus on this type of semantic lexical ambiguity.

Specifically, homonyms can be further categorised into four groups based on the differences in phonetic forms, graphic forms, lexical semantics, and lexico-grammatical meanings (Mamedova, 2019; Khisamova et al., 2020). Firstly, words that have the same or similar written forms, but pronounced differently are known as homographs. For example, tear has two meanings, one refers to crying, another refers to pulling something apart. Secondly, homophones are defined as words that have the same or similar spoken forms but differ in written forms (Parent, 2012). For instance, knew and new are two different words, but they have the same pronunciation. Next, words that are from the same word class, but have different lexical meanings are called lexical homonyms. Taking the adjective fair as an example, this word has two meanings, one refers to treating people equally, and another means light colour. Lastly, lexico-grammatical homonyms are defined as words that have distinct lexical and grammatical meanings and from different word classes (Mamedova, 2019). For instance, bear can be a noun, means an animal, and it can also be a verb, referring to taking responsibility. As homonyms have various forms which require learners not only to notice its lexical structures, but also the use of homonyms corresponds to its meaning in a particular context. Thus, the researcher would like to explore further on the use of homonyms by learners while writing.

In the case of ambiguity, it could possibly occur in both spoken and written contexts. However, Demir (2020) highlighted that if an ambiguous situation happens in a conversation, the speaker could correct it instantly, while this could not be applied to writing. Once learners have decided on the points included in their writings, it will be recorded permanently in a paper or digital format, and the contents are no longer amenable if it is submitted to someone or published on a platform. Therefore, learners need to pay close attention to the words used to construct sentences in different contexts, especially homonyms to avoid causing any confusion

in the readers. Besides, the ambiguous nature of homonyms could be difficult for listeners to decode the intended messages conveyed by the speaker in a spoken context if they do not possess sufficient linguistic knowledge (Demir, 2020). Nevertheless, if listeners are unable to comprehend the utterances, they are allowed to prompt questions to the speaker, and the ambiguity issue could be resolved when the speaker uses speech cues or visual aids to explain further to the listeners. Conversely, in written discourse, readers are not provided with such cues or aids which could hinder them from clearly understanding the intended meaning when they read ambiguous words or sentences (Demir, 2020). Hence, a careful selection of words, which in this case, homonyms are important to ensure the choice of words do not affect the transmission of the exact sense to the readers and maintain the credibility of the writing produced by the writer.

On the other hand, another factor that contributes to the learning difficulties for homonyms could be that over 80% of common English words are homonyms like homophones and homographs (Rodd et al., 2002, as cited in Demir, 2020). This indicates that learners are unaware about some of the meanings of homonyms and would commit errors if their vocabulary knowledge is not on par with their competency level. As a result, learners would randomly select a word that is inappropriate in a context which led to the distortion of the overall sentence meaning that could be misleading for readers (Rohmatillah, 2014). In this case, learning the word usage is prominent for learners to have a clear understanding about the use of homonyms while making sentences. By noticing the word usage, learners could observe the common patterns of homonyms used alongside with other words in a corpus which would help them to distinguish the differences of homonyms because usage pattern of homonyms would reflect different meanings used in various context (Levy & Bullinaria, 2001). Therefore, the present study will focus on exploring the usage pattern of homonyms in the writing of Malaysian ESL learners through corpus.

## 1.2 – Problem Statement

There are two forms of lexical ambiguity which are polysemy and homonym. Both ambiguities pose several challenging issues to learners while using them in their writing. The problem that is struggled by most learners is identifying the differences of polysemy and homonym and classified them into different categories accordingly. Demir (2020) claimed that polysemy is much more complex in nature due to its variety forms of lexemes; it could be difficult to detect when polysemy is presented as a metaphor, metonymy, simile, and other forms.

In fact, polysemous words are quickly processed in the mental lexicon when compared to homonymous words because unlike polysemous words, the unrelated senses of homonyms have triggered competition between meanings which has caused a delay in the process of word recognition (Pustejovsky, 1995, as cited in Yurchenko et al., 2020). This situation could lead to word selection problems in learners when they are unable to decide the correct word (Vethamaiccam & Ganapathy, 2017), they need to be familiar with the concept of homonyms first which will guide them towards accessing multiple meanings of homonyms better. In the process of acquiring semantic knowledge, learners usually do not connect the word forms to its meanings but develop gradually. Learning homonyms is considered a challenging task due to its ambiguous nature and could affect the process of acquiring a new meaning when the learners are more biased towards the primary meaning of words (Ushiro et al., 2013). Learning the notion of homonyms would give learners a clear idea on the categories of homonyms, meanwhile it would acknowledge the importance of knowing their differences while building sentences in different contexts.

Furthermore, the varieties of English like American English and British English could have words that have different meanings in both varieties. When learners are not aware of the differences in both varieties, they could misinterpret the meanings if they are biased towards



one of the varieties, simply because English homonyms could be used interchangeably in both varieties in a sentence depending on the context (Deng, 2011). It deduced that learners are confused with the multiple meanings of homonyms, and this could happen particularly in general English. When learners show confusion in comprehending the meanings of homonyms, they would eventually struggle in selecting appropriate words to be used in a context (Jacobson et al., 2007; Saleh, 2017). If learners are allocated with a time to complete the writing task, due to the time constraint, there is a possibility that learners would randomly select words to fit into the sentence when they are dealing with homonyms. Consequently, the writers or learners could have chosen incorrect words that could obstruct the meanings of the sentences.

On the other hand, the confusion also could appear in both spoken and written context. As homonyms show identical features in written and spoken forms, learners could have written the spelling of a word (homophone) which does not correspond to the context, or they could have struggled in articulating different pronunciations of a word (homograph) accurately (Afzal, 2019). In other words, this phenomenon could be the result of the learners not knowing the word has other meanings and the usage of the word in various contexts, it eventually led to misusing. Since homonyms are mostly common English words, learners are more likely to use it to construct sentences (Rodd et al., 2002, as cited in Demir, 2020), this has emphasized its importance for learners to know the common usage to assist them in learning different homonyms that could pose confusion.

Moreover, learners' confusion over English homonyms not only varies from English varieties to semantics, but also have a close relation to vocabulary. In some situations, learners know the words have multiple meanings, but they could not differentiate the differences of the words which has caused them to commit word choice errors. Vethamaiccam & Ganapathy (2017) reported that high school students are confused with word choice related to homophones due to its similarities in pronunciation, the spoken form of the words could sound same or

similar with another word which has a different meaning. In contrast, Demir (2020) asserted that homographs are often unaware by learners merely because the amount of word meanings are more than what they can remember, or they do not realize about the ambiguous nature of some of the English words. Both circumstances could mean learners have limited vocabulary knowledge about homonyms, and such confusion would follow the learners continuously even after they grow into an adult.

Apart from having insufficient vocabulary, second language or foreign language learners would experience first language interference which would cause difficulties in interpreting homonyms (Ushiro et al., 2013). This also indicates that learners try to translate words from their mother tongue into English language which they thought the translated words match with the context of writing (Vethamaiccam & Ganapathy, 2017). However, learners need to know that the lexical and semantic structures of their mother tongue are very different from the English language, and the intention of translating the words will affect the meaning of the expressions.

Additionally, the unrelated sense of homonyms does not show a semantic connection between primary and target meanings which could be the reason why homonyms are not easy to acquire when learning English. The major issue that often faced by learners would be misinterpret homonyms because they tend to be biased towards the primary meaning and this could lead to learners' misunderstanding in the unfamiliar meanings (Ushiro et al., 2010, as cited in Ushiro et al., 2013). While confronting this problem, learners should know that it is possible to learn homonyms and differentiate them even with many different meanings. When they are exposed more on homonyms used in several contexts and by referring to this information regularly will help them to expand their existing vocabulary knowledge on homonyms through learning the new meanings (Ushiro et al., 2013). It can be concluded that knowing the contextual information will assist learners in comprehending multiple meanings

of homonyms used in different contexts and performing this action of learning will guide them in their writing. Hence, the present study would like to analyse homonyms from several contexts in the writing of ESL learners.

Besides, learning homonyms in different contexts could help learners to develop their vocabulary knowledge because they are more likely to encounter unfamiliar words during the learning process (Nagy, 1995). By comprehending the contexts, learners know how to disambiguate different homonyms and have a better understanding of multiple meanings of homonyms. Next, learners could pose a high possibility of experiencing cross-linguistic differences in learning homonyms with distinct meanings, and this requires learners to take initiative in finding ways or solutions to solve this problem (Nagy, 1995). A great exposure to different contexts could be an effective approach to resolve the issue even for learners who are at their beginning stages of language acquisition.

Similarly, Ovu (2011) stated that when learners are referring to a context while learning a word, they could clarify the meaning by observing other words that co-occur with the target word. In other words, the concordance list of a corpus which highlights all contexts of a word would be a practical tool to use for learning homonyms. By using corpus, learners will not overlook the usage pattern of homonyms and have a full understanding on how contextual information will assist them in selecting the right homonym to use in their writing, as well as, knowing the significance of learning usage pattern to avoid their confusion on homonyms. However, there are limited studies found on analysing the usage pattern of homonyms in the writing of Malaysian ESL learners. Thus, the current study aims to help ESL learners to identify the differences of homonyms by referring to the word meanings used in various contexts and seek to minimize their doubts while choosing homonyms to use in their writings.

The aim of this study is to investigate how usage of words in different contexts could help in disambiguating homonyms. Specifically, the objectives are as follow:

### 1.3 – Research Objective

1. To identify the type of homonyms used in the writing of Malaysian ESL learners.
2. To report the usage pattern of homonyms in the writing of Malaysian ESL learners.

### 1.4 – Research Question

1. What is the type of homonyms used by Malaysian ESL learners in their writing?
2. What is the usage pattern of homonyms in the writing of Malaysian ESL learners?

### 1.5 – Significance of the Study

In the process of learning the English language, ESL learners are required to expand their vocabulary knowledge constantly through adopting distinct ways which would help them in constructing sentences. While learners acquire new vocabulary, they need to learn the lexical and phonetic form of words that correspond with its meanings. Most of the English words are usually ambiguous, and this requires learners to pay extra attention to these words, which are homonyms. Unlike polysemous words have related senses which allow ESL learners to transfer knowledge easily from one meaning to the others, homonyms have unrelated senses which pose difficulties for ESL learners to fully comprehend multiple meanings of homonyms without a context given (Kulkarni et al., 2008). As such the current study would like to emphasize the importance of learning the usage pattern of homonyms through collocation. A collocation refers to other words that co-occur with the homonyms in a sentence, and this feature is normally found in a corpus (Mukoroli, 2011). By using the collocation lists in the corpus, learners will know what word is associated with the homonyms in different contexts,

they can identify the differences of homonyms and understand several distinct meanings better when it is presented in the form of contexts. Meanwhile, when learners have frequent exposures to different usage of homonyms, it will help in developing their word recognition abilities and this eventually improve their language production while writing a text or an essay (Mukoroli, 2011). In other words, when learners have mastered different usage of homonyms, it will help them in resolving the word choice issue, minimize their confusions, as well as reduce their errors made while writing.

Besides, the ambiguous nature of homonyms has caused misinterpretation of multiple meanings of homonyms in ESL learners when it is presented in isolation. Even learners are provided with semantic information about homonyms, they are not fully aware of which homonyms should be used in a specific context, and this phenomenon is resulted from learners' bias towards the primary meaning which is acquired earlier than other meanings. As such learners' interpretation on the secondary meaning of homonyms is interfered with by the primary meaning due to familiarity of using it more frequently, and this indicates that they could not distinguish various meanings of homonyms without the context (Ushiro et al., 2010). When learners are given contextual information to learn homonyms, they can make correct interpretations of homonyms by choosing the right word that matches with the context. Although it could be difficult for novice learners to acquire homonyms at the beginning stage of learning, once they have more exposure to the uses of homonyms in different contexts and familiar with various types of homonyms, they could identify the differences of meanings and assist them in selecting appropriate homonyms to use in a specific context. Therefore, contextual information played a crucial role in helping learners to use homonyms accurately in their writing.

On the other hand, using corpus as an analysing tool could bring advantages to both teachers and learners. While teachers incorporate corpus as a language teaching tool in the

classroom, they could facilitate students to learn the usage and meaning of vocabulary through the concordance lists (Poole, 2016). Since homonyms show similarity in spoken and written form which has caused confusion in the learners, the teachers could utilize the concordance lists in corpus to teach homonyms by initiating discussion with their students about the usage of homonyms. Besides, concordance enabled learners to explore different examples of homonyms and provide an opportunity for them to learn language components like collocations, semantic, and pragmatic along with homonyms (Poole, 2016). As concordance lines included all contexts of the keyword, learners can make a comparison between the usage and meaning of different homonyms. Making comparison also allow learners to observe the usage pattern of homonyms and categorise the homonyms based on different features. By using this approach to learn a new pair of homonyms, it will help in promoting vocabulary learning among learners (Hoshino, 2010, as cited in Safataj & Amiryousefi, 2016). Hence, the current study seeks to promote the use of corpus as a pedagogical tool to teach students on homonyms, meanwhile, students can learn more about the usage of different homonyms are depends on the contexts and helps to reduce their misuse problems.

#### 1.6 – Limitation of the Study

The current study will discuss the topic of ambiguity, specifically about lexical ambiguity that consist of both syntactic lexical ambiguity and semantic lexical ambiguity. In the present study, the researcher will only focus on homonym instead of both polysemy and homonym under the category of semantic lexical ambiguity. As this study only highlighted homonyms, learners could not fully understand the concept of semantic lexical ambiguity and were not fully aware of the categories of polysemous words since polysemy is not included as part of the research.

Next, the present study will select homonyms used in the written context only since the researcher will analyse the usage of homonyms in the writing of ESL learners and discover the types of homonyms used by the learners while writing an essay. As the current study will focus on the writing skills of ESL learners in terms of using homonyms, therefore, the homonyms used in the spoken context will not be emphasised in this study. In other words, the homonyms found in this study cannot represent the spoken context of usage as it does not correspond to the speaking or reading skills.

Moreover, the current study will only analyse the homonyms used by Malaysian ESL learners in their writings as it is designed to help them to identify homonyms that are usually used in the local context by reporting its usage pattern that enabled learners to form a connection between their writings with the local setting. As the present study is only limited to ESL learners at undergraduate level, the findings presented in this research cannot represent the usage pattern of homonyms for all ESL learners in general.

## CHAPTER 2 – LITERATURE REVIEW

### 2.1 – Definition of Terms

#### 2.1.1 – *Ambiguity*

A new meaning often generated by people through forming a new sense to an existing word or combining the existing words which has caused multiple interpretations occurs within an expression (Jackson, 2020). This phenomenon is known as ambiguity which is defined as a word, or an utterance can be interpreted in many ways when it has multiple meanings. It can be found at different levels of meaning which includes expression meaning, utterance meaning, and communicative meaning (Loebner & Hooimeijer, 2013). As an expression or a phrase is comprehended by people in more than one way in distinct contexts, it seems to be common while learning English language (Asher, 1994, as cited in Kristian, 2018).

#### 2.1.2 – *Sense Relations*

Generally, ambiguity is more than interpreting the meaning at lexical level, instead it has close connections with other words. This circumstance is known as sense relations whereby a word's meanings influence the sentence's meanings, as such the meaning also depends on other words that co-occurs within the sentence (Kreidler, 1998). The co-occurrence of words in the same phrases or sentences can be noticed through these two types of sense relations, syntagmatic relations, and paradigmatic relations. When there are two or more words associated together, the meaning of one of the words will affect another, and their meanings will influence the meaning of the phrase or sentence. On the other hand, paradigmatic relation occurs when words with a slightly different meaning are chosen to replace another word with similar meanings. Besides, the word meanings will change when it is used in various contexts, and the contextual information plays a fundamental role in determining the meaning of words (Kreidler, 1998). In other words, a word's meanings do not stand alone, in fact, its meanings contribute to the sentence or utterance meaning and vary in contexts.



### *2.1.3 – Lexical Ambiguity*

Lexical ambiguity is defined as a word or an expression that has multiple different concepts. It is also known as semantic ambiguity because only the meaning of the word is ambiguous, the lexical structure or grammatical properties are not affected (Vitello & Rodd, 2015, as cited in Rodd, 2017). Nevertheless, these two terms are used interchangeably to refer to the ambiguous nature of English words. As lexical ambiguity is omnipresent in most English words and multiple interpretations could be found within a sentence, it is crucial for learners to learn more about different forms of lexical ambiguity. In general, there are two forms of lexical ambiguity which are homonym and polysemy.

### *2.1.4 – Homonym and Polysemy*

The etymology, relatedness and unrelatedness of meanings of words are the factors used to distinguish both homonyms and polysemy. When a word has multiple unrelated meanings, and these meanings have different origins that are usually recorded as separate entries in the dictionaries, it is commonly known as homonym (Rodd, 2017). Homonyms can be further classified into homophones, homographs, lexical homonyms, and lexico-grammatical homonyms. Homophones refer to words that have the same or similar pronunciation but have different written forms. While homographs refer to words that have the same or similar spelling but are pronounced differently. Next, lexical homonyms refer to expressions that are from the same word classes but have distinct meanings. On the other hand, lexico-grammatical homonyms refer to expressions that are from different word classes, lexical and grammatical meanings (Mamedova, 2019).

Moreover, polysemy is another form of lexical ambiguity which is defined as the meaning of words are semantically related to one another, and it is usually recorded as a single entry in the dictionaries (Rodd, 2017). Polysemy can be divided into two main categories, linear

polysemy, and non-linear polysemy. Linear polysemy highlights the relationship of specialisation and generalisation between senses, and it is further classified into four subcategories which are autohyponymy, automeronymy, autosuperordination and autoholonymy. Autohyponymy refers to a word provide a general idea to another sense that refer to it to form a subvariety. While automeronymy refers to the specific sense of a word is a subpart of the general sense of another word. Next, autosuperordination refer to the meaning of a word could be used to represent the similar sense of another word (Cruse, 2000, as cited in Balloqui, 2017). Autoholonymy refer to a word that act as a larger part of sense to other words (Cruse, 2010). Furthermore, non-linear polysemy can be sub-categorised into metaphorical polysemy and metonymous polysemy. Metaphorical polysemy is defined as polysemous words are used figuratively in a sentence, while metonymous polysemy is defined as a word that is used to refer to words that have similar senses are closely associated with each other (Balloqui, 2017).

## 2.2 – Past Studies

### 2.2.1 – *Past Research in Malaysia*

There are several past studies in Malaysia focused on ambiguity themes. Mahadzir et al. (2018) highlighted that ambiguity could happen in both English and Malay languages when words in these languages have the same or similar written form but are semantically different. It is easier to identify the exact sense of the word when it is allocated in a sentence where the meaning of other words presented along with it could provide a hint to know the language used and the right meaning of the word. This has indicated that collocations played a vital role in determining the meaning of ambiguous words used in a context, even in only one language, since ambiguity is ubiquitous in the English language. As ambiguity is commonly present in most English words, learners faced the challenge of disambiguating words or sentences with

multiple meanings. Regarding this issue, Al-Harbi et al. (2011) stated that contextual information could assist learners in selecting the intended meaning from several meanings of a word. This process is known as word sense disambiguation (WSD) and it showed a close connection among words and meanings with its context (Agirre & Edmonds, 2007, as cited in Al-Harbi et al., 2011). Generally, a context covers a wide range of language components which includes word and syntactic relations, orthographic properties, collocation, semantic categories, and selection preferences (Abney, 1991, as cited in Al-Harbi et al., 2011).

Next, Tan (2018) mentioned that learners not only need to acquire the lexical and semantic forms of a word while learning vocabulary, but it is also equally important to learn words in different contexts through frequent exposures. Such an approach could facilitate the process of vocabulary learning because it expands their existing knowledge about words through discovering various meanings of a word used in different contexts, meanwhile it helps them to revise the meanings of learnt words (Tan, 2018). When learners are provided with contextual information to disambiguate words or sentences and it is carried out on a regular basis, this will help learners in minimizing their difficulties of comprehending the intended meaning of words or sentences. Besides, Tan (2018) conducted the research by utilising direct vocabulary instruction to teach students in finding the definitions of homonyms through dictionary and constructing sentences with multiple meanings of the homonyms. Unlike this past research, the present study will focus on using corpus to learn the meanings of homonyms through their usage pattern.

Moreover, there are a few past studies focused on integrating corpus as part of the process of acquiring different linguistic knowledge. Using corpus as a learning and pedagogical tool is known as data-driven learning (DDL) approach which aids in learning vocabulary, teaching of prepositions, and resolving the issue of learning lexical ambiguity (Shamsudin et al., 2013; Yunus & Awab, 2012; Azmi et al., 2021). Corpus is regarded as an effective approach

in learning specialised vocabulary, which in this case jargons used in a specific field. Although a specialised dictionary is available for students to find out the meaning of the jargon, corpus could help in addressing the issue of limited entries found in the dictionary (Shamsudin et al., 2013). It seems to be necessary to generate a specialised word list through the collocation because it assists students in learning and comprehending jargons used in their field of study. On the other hand, Yunus and Awab (2012) conducted a study to test the effectiveness of DDL approach in improving learners' knowledge about the meaning and usage of different prepositions. Similarly, this study also highlighted the benefits of using collocation to learn prepositions where learners can observe the usage pattern of different prepositions through surrounding words, and deduce the meanings of the sentences. Besides, by utilising corpus, it will help to increase learners' awareness about the close connection between grammar and vocabulary, and strong ties between the forms and meanings of words.

Furthermore, Azmi et al. (2021) adopted an indirect corpus-based approach to explore its effectiveness towards solving the problems of learning lexical ambiguity. This study focused on using language modules which consist of concordance lines from corpus to disambiguate near synonyms. As the language modules introduced the concordance features, learners are exposed to more usage of synonyms in different contexts. The concordance function provided an authentic learning experience to the learners with examples given are all in real context, as well as, it acts as a guideline for learners to select the most appropriate synonym used in various contexts and this avoids making errors related to incorrect word choices. Overall, these corpus studies underline the significance of collocation and concordance in facilitating language learning among ESL learners. In other words, a corpus-based approach helps to promote learners' autonomy in learning different language components through discovering the usage patterns of English words which will also be the main focus in the present study.

### 2.2.2 – *Past Research in Overseas*

A few past studies overseas highlighted problems encountered by learners while using different types of homonyms (Adha & Widyaningtyas, 2017; Owu-Ewie & Williams, 2017; Ibrahim, 2018). The similar features of both homophones and homographs in spoken and written form have caused communication barriers between the teacher and students. In a classroom setting, when the teacher mentioned a homograph or homophone in an utterance, students have misinterpreted the homonyms whereby they referred to another similar word from a pair of homographs or homophones (Adha & Widyaningtyas, 2017). This issue not only occurs in spoken context, in fact, it also influences learners' performance in writing.

Next, in another study conducted by Owu-Ewie and Williams (2017), students committed most errors related to homophones and such errors are subject to the problems of learners' misspelling and misunderstanding towards the meaning and usage of homophones. Apart from these two factors, L1 interference is believed to be another factor that led to learners committing homophone errors. This type of error does not limit to primary or secondary school students only, but it also applies to tertiary students partly because the ambiguous nature of homonyms usually causes confusion among learners of different ages. Ibrahim (2018) proved that undergraduate students could not comprehend both homophones and homographs, and they are unable to incorporate it into their writing. This phenomenon could be resulted from insufficient coverage by the teacher on different types of homonyms with its form, meaning, and usage introduced to the learners. All these past studies related to homonyms errors have emphasized the importance of learning the usage pattern of homonyms in different contexts whereby it could possibly help learners to minimize their confusion and errors made in their writing.

Additionally, in a written context, learners usually posed challenges in incorporating ambiguous words to construct sentences, without having sufficient knowledge about lexical

ambiguity, it has led to overlooking the usage pattern of ambiguous words used in different contexts. As students normally incorporated more lexical ambiguity than syntactic ambiguity into their writing (Demir, 2020), they should be more aware on ambiguous words to avoid from misinterpreting the meanings of the sentences, and this will directly influence their interpretations of the message conveyed in the texts or paragraphs. Besides, verbs are frequently used by learners to build sentences, and most verbs have multiple meanings, which in this case verbs are regarded as ambiguous words. This has indicated that verbs are prominent in comprehending ambiguous sentences (Garnsey et al., 1997, as cited in Demir, 2020), if verbs are misused in a sentence, it will affect the overall meaning of the sentences. Hence, verbs are seemed as a disambiguating element in processing language because a study found that learners access various meanings of verbs through activating one underspecified sense of word and contextual information act as a supportive role to attain the most precise meaning within a context (MacDonald et al., 1994, as cited in Demir, 2020).

Since lexical ambiguity is more commonly encountered by learners, it is also more difficult for them to process while writing a text. Yu (2013) stated that learners' lack of lexical knowledge could be the reason for facing this challenge, and this could result in making word choice errors in their writing. Apart from that, learners are more likely to select English words that match with their L1 after translation as they assume that this way could help them to express the intended meaning in English language better, but learners' dependence on their mother tongue may lead to overgeneralization. Dictionaries are always the companion of learners while acquiring vocabulary, but Yu (2013) claimed that overly reliance on online dictionaries would not help learners to decide ambiguous words that are appropriate to be used in different contexts because dictionary usually provided semantic information and less likely presented words in real context. In short, the obstacles faced by learners in using ambiguous words could be resolve when learners are fully aware of the usage pattern as they can observe

the lexical and semantic pattern of words used in various contexts presented through collocation, and this approach will be highlighted in the current study by using corpus.

## 2.3 – Theoretical Foundation of the Study

### 2.3.1 – *Usage-Based Theory*

Usage-based theory is a theoretical perspective that highlights the effect of usage on linguistic structure through cognitive process. As such it entails usage patterns, frequency of occurrence, variation and change as part of the mental representations for linguistic items. The research conducted based on this theory emphasized on the area of grammaticalization as the grammatical meaning and grammatical structure exists through repetition in actual language use. Since the large electronic corpora emerged, more studies have shifted their focus to analysis of natural language use with access to various linguistic items and patterns found in the corpora (Bybee & Beckner, 2015). By using corpus, it allows linguists to explore how language is used in real contexts and study the variations of language that are presented in different forms and structures. As language changes perpetually over a period, it requires individuals to take note of the importance of repetition, chunking, and knowledge of usage while learning a language. When learners constantly repeat their learning activity, it helps to develop procedural knowledge and this reinforces the cognitive representations of the learning process that is stored in long-term memory (Shadmehr & Brashers-Krug, 1997, as cited in Bybee & Beckner, 2015). Once repetition has gradually developed into habits, the cognitive system will act towards enhancing the learning performance through the process of chunking. Through this process, individuals can identify patterns and process co-occurrence of words in a sentence effortlessly in their brain, and this is closely related to the individuals' linguistic knowledge which includes the procedural knowledge, frequency and patterns (Saffran et al.,

1996, as cited in Bybee & Beckner, 2015). These three elements are fundamental for learners to process various linguistic items since language evolves and changes over time.

On the other hand, repetition will only work when there is similar characteristic or property found in the linguistic structure, and this indicates that categorisation played a crucial role in connecting different inputs into same classes in a context (Pierrehumbert, 2001, as cited in Bybee & Beckner, 2015). When inputs are presented in the form of a category, they tend to show the same or similar traits that could facilitate the process of memorising a particular category. By using categorisation, learners not only learn the specific features of linguistic items, but also retain it along with generalisation in their memory. Overall, the usage-based theory allows learners to explore the internal structure of language from the perspective of natural language use.

The theory has highlighted that through exposure to the usage pattern of words could help learners to remember how words are used in constructing sentences. This approach is related to the current study whereby the researcher will apply the categorisation aspect in this theory to classify the types of homonyms into different categories. By applying this theory to the study, it allows the researcher to discover the types of homonyms that are normally used by Malaysian ESL learners.

### *2.3.2 – Katz's Semantic Theory*

Katz' semantic theory is a theory of natural language semantics that emphasized on the semantic relations and differences between word meanings and highlighted the relationship between lexical meanings and sentence meanings (Katz & Fodor, 1963, as cited in Falkum, 2011). The theory consists of two major components, which are dictionary and projection rules. A dictionary provides a list of lexical entries with meanings instead of presenting it sentence



by sentence, and every sense of these entries can be used in any sentence. On the other hand, it has highlighted the role of projection rules whereby it acts as a decision maker to determine the exact or appropriate sense of the lexical items used within a sentence, and this could refer to the individuals' ability in understanding the sentences through its own interpretations (Katz & Fodor, 1963). The projection rules will be activated when learners encounter ambiguous words or sentences, it helps in disambiguating the word meaning by combining different senses of the word with other surrounding words within the sentence to select the most compatible meaning that could fit into the context (Falkum, 2011). This has indicated the semantic relations of words in a sentence are represented in the form of semantic marker, whereby it reflects the semantic relations between multiple senses of a word and other lexical items (Katz & Fodor, 1963). In this case, semantic markers function as a connector that links distinct parts or words of a sentence through semantic relations, which help individuals in making correct interpretations on ambiguous sentences.

The theory has pointed out that a dictionary provides the senses of a word, while projection rules help in disambiguating the meanings of a word. As the present study focuses on homonyms, the researcher will utilise this theory in the process of examining the usage pattern of homonyms to check whether the homonyms included in the sentences match with the context of writing by looking into its surrounding words. Hence, the present study will adopt these two theories as part of the process of analysing the homonyms used in texts written by Malaysian ESL learners.

2.4 – Conceptual Framework

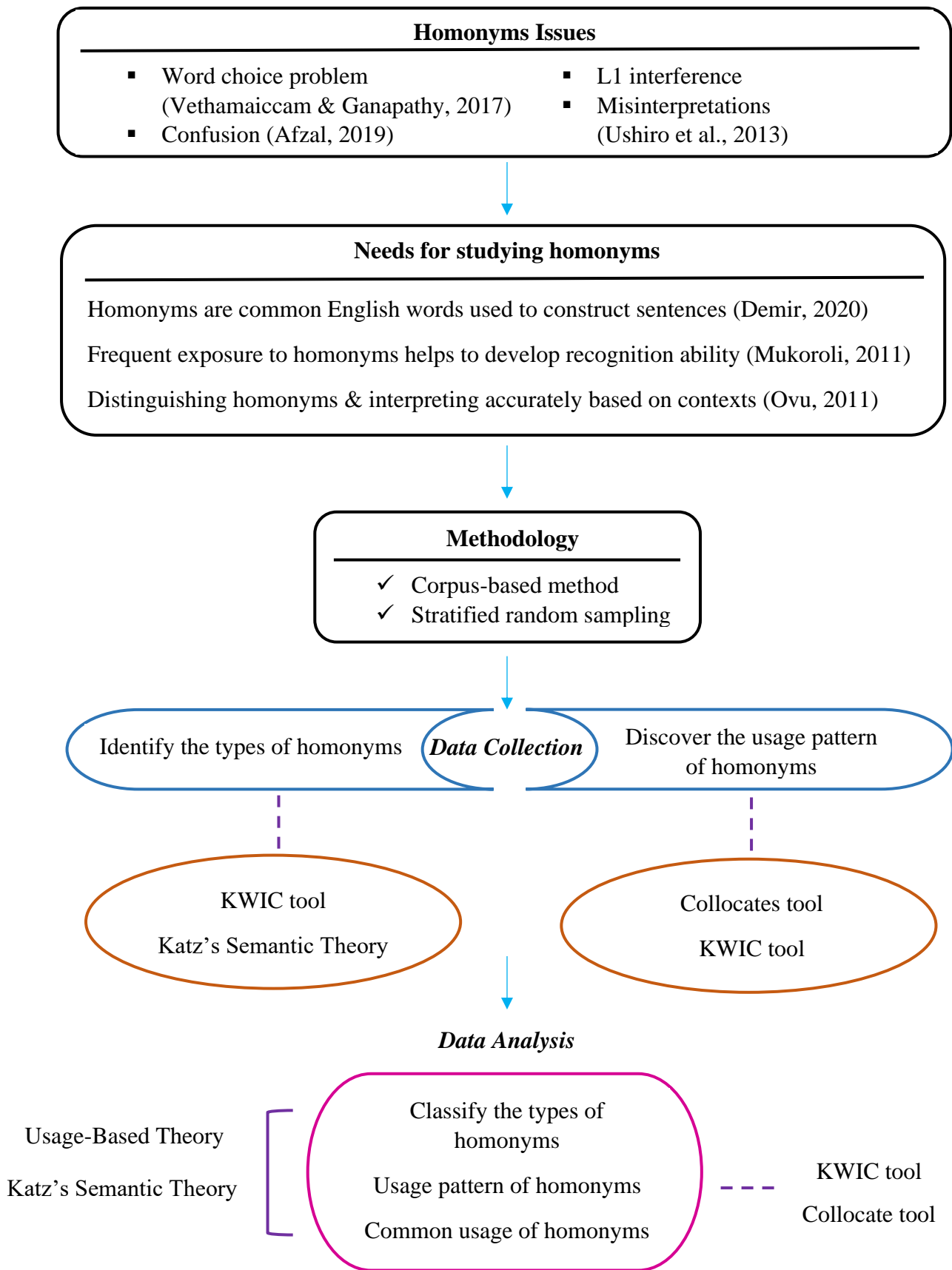


Figure 2.1 – Diagram of the conceptual framework

Diagram 1 above shows the conceptual framework of the present study that focuses on identifying the types and analysing the usage pattern of homonyms by using corpus. As the current research adopts the corpus-based method to conduct the study, there are several tools in the corpus that could assist the researcher in gaining the actual data on how Malaysian ESL learners incorporate and use homonyms in constructing sentences in their writing. By utilising corpus, the researcher will use the Key-Word-In-Context (KWIC) tool to find out homonyms from the content words while referring to the dictionary for the meanings of the homonyms. This approach involves the use of Katz's Semantic Theory as the researcher will look into the spelling, meaning, and word class of the homonyms in order to identify their types based on the characteristics shown in these structures. Once the types of homonyms are identified, the researcher will classify them into different categories which are related to the Usage-Based Theory as homonyms that showed same or similar properties are grouped together through the process of categorisation stated in the theory.

Moreover, the researcher will explore the usage pattern of homonyms in these two aspects including the common usage and the appropriateness of homonyms used in certain contexts. In terms of common usage, the researcher will use the Key-Word-In-Context (KWIC) tool in the corpus to discover how homonyms are commonly used by Malaysian ESL learners in their writing. By knowing the common usage of homonyms, learners' bias towards using single meaning of the homonyms can be seen, and this indicates that exposure to other meanings of the same homonym should be done. Once the common homonyms are identified, a collocate tool in the corpus will be used to collect data for the co-occurrence words that are used along with the homonyms, while the file tool will be utilised when more contextual information is needed for examining purpose. During this process of analysing, the researcher will use online dictionary to comprehend the meanings of the homonyms first, then observe the usage pattern of homonyms used in the context of writing, and lastly based on the context

given, disambiguate the meaning of homonym to identify whether the homonym is appropriately used by Malaysian ESL learners. In other words, the researcher will apply both Usage-Based Theory and Katz's Semantic Theory to examine the how homonyms are used in the writing of local undergraduates.

## CHAPTER 3 – METHODOLOGY

### 3.1 – Research Design

#### *3.1.1 – Corpus-Based Method*

The researcher would like to use the corpus-based method to conduct research about the usage pattern of homonyms in the writing of Malaysian ESL learners. Corpus-based approach is a method that uses existing corpus as a tool to support intuitive knowledge, discover evidence for theories, examine various linguistic phenomena, and collect samples from a rich source of information used in natural language (Storjohann, 2005). As corpora provides a large database that covers natural language use, and such a platform allows the researcher to analyse the actual pattern of homonyms used in real contexts (Biber et al., 1994). By adopting this method, the frequency, regularity, and usage patterns of homonyms could be observed, analysed, and interpreted through applying a few corpus techniques (Gablasova et al., 2017). These techniques basically feature multiple functions of corpora including concordance, collocations, and word frequency lists that will be used in the present study.

A concordance is usually presented in lines that show the occurrences of the search word in the centre of the line along with other words co-occur with it which are located at left or right side of the search word (O'Keeffe et al., 2007). This function allows the researcher to distinguish the homonyms used in different contexts which helps in discovering the meaning of homonyms by noticing the lexical and grammatical patterns of other co-occurrence words. Similarly, collocations also displayed words that exist in the surrounding of a search word. However, unlike concordance, collocations focused more on accompanied words rather than search words (Scott, 2010, as cited in Hussein, 2015). As such collocations provide contextual information to the researcher in identifying semantic restrictions that homonyms impose on different contexts which will be significant to address the word choice issue. Next, the word frequency lists basically feature the total number of individual words (tokens) and the number

of types in a generated corpus, and it will arrange either based on most to least frequency of words or alphabetical order (Hussein, 2015). By referring to the word frequency lists, it enabled the researcher to find out which homonyms are commonly used in the writing of Malaysian ESL learners.

### 3.2 - Sampling Selection

The current study aims to analyse the usage pattern of homonyms with the selected population, which is undergraduate students from different faculties in UTAR, Malaysia. As the researcher will examine written text, different types of essays including causal analysis essay and argumentative essay done by students in the Academic Writing subject will be chosen as the sampling unit. A total of 145 essays from these genres will be collected from the subject lecturer of the same institution as the study is a specialised corpus with a relatively small size, the researcher seeks to find out how the generated learner corpus link to the local context (Koester, 2010, as cited in Ngula, 2018). The researcher will apply the stratified random sampling method. This sampling method divides a whole population into homogeneous groups, which is known as strata, and each stratum is sampled at random (McEnery et al., 2006). Under the stratified sampling, the researcher will follow the proportionate stratified sampling procedure where the number of essays sampled from each category is proportional to the target population of the representative group (Daniel, 2012). As the stratified random sampling method is used in one of the corpus studies where the target population is similar with the present study, the researcher would like to apply the same method (Budiwiyanto & Suhardijanto, 2019). Unlike the past research, the samples of the current study will select from the Faculty of Arts and Social Science and the Faculty of Business and Finance for students who have completed the Academic Writing subject in the year 2021 and 2022. Only the full texts will be kept from all the essays collected (Crosthwaite et al., 2019), and the

rest of the details will be excluded before it is digitised into plain text format to generate a learner corpus.

### 3.3 – Instruments

#### 3.3.1. – *Concordance software*

The current study seeks to identify the types and report the usage pattern of homonyms used in the writing of Malaysian ESL learners through building a learner corpus. Once the texts are digitised into plain text format by using AntConc Converter, a concordance software (AntConc) with the version 4.0.11 will be used to create the learner corpus. Although there are several functions featured in the software, the researcher will only utilise the KWIC tool, collocates tool, and file tool. Firstly, the Key-Word-In-Context (KWIC) tool displays how the homonyms are used in different contexts which is presented sentence-by-sentence that will be used to observe the usage pattern of homonyms used in sentences. On the other hand, the collocates tool highlights the surrounding words used along with search terms (Anthony, 2020). Both concordance and collocates tools assists the researcher in examining homonyms used in different contexts by exploring the co-occurrence words and identifying whether the homonyms are selected appropriately to fit into the context. Lastly, the file tool in the corpus will be used when more contextual information is needed to disambiguate multiple senses of the homonyms (Anthony, 2021).

### 3.4 – Data Collection

As the current study uses corpus to identify the types and usage pattern of homonyms, there are several functions that will be useful in collecting these data. The KWIC tool in the corpus will help the researcher to identify the homonyms from the content words, and referring to a dictionary is necessary to clarify whether the content words are homonyms by looking into

the meanings of content words. While using a dictionary, the researcher is applying Katz's Semantic theory to find out other characteristics from lexical, semantics, and grammatical structure of words that will be crucial to identify the types (Katz & Fodor, 1963). In terms of usage pattern, the researcher would like to discover the common usage of homonyms, how homonyms are misused by students, and what words that are usually used together with these homonyms to construct sentences. Therefore, the usage pattern of homonyms will be collected through the collocates and KWIC tools in the corpus. The KWIC tool will show the total frequency of occurrence of the identified homonyms, and it allows the researcher to report the common usage of homonyms used by Malaysian tertiary learners. While the collocates tool will show the co-occurrence words that are used along with the homonyms in various contexts.

### 3.5 – Data Analysis

The present study seeks to identify the types and analyse the usage pattern of homonyms through the corpus-based method. When the data for types of homonyms is identified, the researcher will classify these homonyms into different categories based on their characteristics, and this process is known as categorisation. As categorisation is one of the important aspects under the usage-based theory, the researcher employed this theory as it could facilitate the process of learning different types of homonyms (Bybee & Beckner, 2015). The categorisation of homonyms will be presented in the form of a table to provide a general overview on the types of homonyms, and further explanations about the categories will also be included. The collected data for overall usage of homonyms will be presented in bar charts to show the homonyms that are usually used by Malaysian ESL learners. Apart from using collocates to view these homonyms and its surrounding words, file tool also will be utilised when the accompanying words are unable to provide sufficient contextual information to disambiguate the meaning of the homonym used in a certain context. The KWIC tool will show the syntactic structure in which homonyms are used, and refer to the dictionary to identify



whether a homonym is used appropriately in the context by applying both Usage-Based Theory and Katz's Semantic Theory in the analysing process. The data for the usage pattern of homonyms will be presented in the form of picture format as it provides a clearer view to observe and comprehend the usage pattern and characteristics of different homonyms.

### 3.6 – Ethical Considerations

As the present study does not involve participants, samples will be collected through two lecturers who are teaching students taking the Academic Writing subject from different faculties. A consent form will be given to the lecturers respectively to gain permission to use undergraduates' essays as the data. While adhering to the confidentiality and protection of the data, all the private details of the students such as name, student ID, and programme will not be enclosed and the data will be presented anonymously.

CHAPTER 4 –  
FINDINGS & ANALYSIS

4.1 – Types of Homonyms

A total of 145 essays written by UTAR undergraduates is collected to generate a learner corpus with 84940 words. This learner corpus helped to analyse the types and usage patterns of homonyms used by Malaysian ESL learners. The identified homonyms are categorised based on their differences in orthography, pronunciation, lexical meaning, and grammatical meaning. Apart from noticing all these features, it is necessary to distinguish different senses of the homonyms while referring to an online dictionary, Oxford Languages. Before classifying the homonyms into different categories, the researcher would examine the meanings of the homonyms used to match with their context to avoid overlapping issues happening during the categorisation process. Altogether, there are four types of homonyms - homophones, homographs, lexical homonyms, and lexico-grammatical homonyms found in this study.

4.1.1 – Homophones

Homophones are words that have the same or similar pronunciation but are different in spelling. This type of homonym is often identified as a pair or in a group of three. There are in total 34 pairs and 2 groups of homophones found in the generated learner corpus. These homophone pairs are arranged sequentially and presented in a table format with a record of their frequency of occurrence.

<i>Homophones</i>	<b>Frequency of Occurrence</b>	<i>Homophones</i>	<b>Frequency of Occurrence</b>	<i>Homophones</i>	<b>Frequency of Occurrence</b>
<i>Accept</i>	21	<i>Except</i>	5		
<i>Advice</i>	8	<i>Advise</i>	1		
<i>Affect</i>	24	<i>Effect</i>	64		
<i>Buy</i>	20	<i>By</i>	320	<i>Bye</i>	1
<i>Close</i>	3	<i>Clothes</i>	5		
<i>Court</i>	1	<i>Caught</i>	1		

<i>Fined</i>	2	<i>Find</i>	25		
<i>For</i>	838	<i>Four</i>	1		
<i>Hear</i>	9	<i>Here</i>	12		
<i>Higher</i>	33	<i>Hire</i>	2		
<i>Holy</i>	1	<i>Wholly</i>	1		
<i>Knew</i>	2	<i>New</i>	91		
<i>Know</i>	77	<i>No</i>	99		
<i>Lessen</i>	1	<i>Lesson</i>	4		
<i>Meat</i>	102	<i>Meet</i>	10		
<i>One</i>	300	<i>Won</i>	14		
<i>Ours</i>	2	<i>Hours</i>	1		
<i>Paced</i>	1	<i>Paste</i>	4		
<i>Peace</i>	7	<i>Piece</i>	7		
<i>Programme</i>	1	<i>Program</i>	1		
<i>Red</i>	6	<i>Read</i>	7		
<i>Right</i>	102	<i>Write</i>	3		
<i>Sea</i>	2	<i>See</i>	18		
<i>So</i>	156	<i>Sow</i>	1		
<i>Some</i>	252	<i>Sum</i>	9		
<i>Sort</i>	2	<i>Sought</i>	1		
<i>Source</i>	16	<i>Sauce</i>	75		
<i>Stake</i>	1	<i>Steak</i>	1		
<i>Their</i>	635	<i>There</i>	285		
<i>To</i>	2636	<i>Too</i>	69	<i>Two</i>	23
<i>Vary</i>	3	<i>Very</i>	75		
<i>Wait</i>	9	<i>Weight</i>	10		
<i>Weak</i>	2	<i>Week</i>	10		
<i>Wear</i>	5	<i>Where</i>	39		
<i>Weather</i>	1	<i>Whether</i>	37		
<i>Wood</i>	2	<i>Would</i>	103		
<b>Subtotal</b>	5283	<b>Subtotal</b>	1411	<b>Subtotal</b>	24
			<b>Total</b>		6718

*Table 4.1 – Homophones list*

Based on table 4.1, the most frequently used homophone is the preposition to (2636), followed by the preposition for (838), pronoun their (635), preposition by (320), cardinal number one (300), adverb there (285), determiner some (252), conjunction so (156), the verb would (103), noun meat (102), and adjective right (102). Most of these homophones are closed class words that act as a marker to introduce the content words like nouns, verbs, adverbs, and adjectives. As function words are usually used as a connector to link all content words together, students used them more frequently because they are familiar with the use of function words compared to the content words. On the other hand, undergraduates showed more bias towards

using one of the homophones in a pair or a group. For instance, a pair of homonyms like for (838) and four (1), as well as a group of three homonyms such as buy (20), by (320), and bye (1) showed a discrepancy in their frequency of occurrence. Likewise, there are 21 pairs, and 2 groups of homophones demonstrated this result, while the rest showed an equal or slight difference in their frequency.

#### 4.1.2 – Homographs

Homographs are words that have the same or similar spelling but are pronounced differently. This type of homonym often has more than one meaning, and most of the homographs are content words like nouns, verbs, and adjectives. All definitions presented are from an online dictionary, Oxford Languages (n.d.).

<b>Homographs</b>	<b>IPA</b>	<b>Word Class &amp; Definition</b>	<b>Frequency of Occurrence</b>
<b>Abuse</b>	/ə'bjʊ:z/	(v.) insult	1
	/ə'bjʊ:s/	(n.) persecution	1
<b>Conflict</b>	/'kɒnflikt/	(n.) disagreement	19
	/kən'flikt/	(v.) clash	1
<b>Content</b>	/kən'tent/	(v.) satisfy	1
	/'kɒntent/	(n.) online or digital information	6
		(n.) the things that are included in something	3
		(n.) the amount of an ingredient that exists in a substance	3
<b>Contrast</b>	/'kɒntrɑ:st/	(n.) the state of being totally different from something	4
		(n.) differences in colour	2
	/kən'trɑ:st/	(v.) differ strikingly	1
<b>Impact</b>	/'ɪmpakt/	(n.) marked effect or influence	19
	/ɪm'pakt/	(v.) have a strong effect on someone or something	1
<b>Increase</b>	/ɪn'kri:s/	(v.) become greater in size, amount, or degree	35
	/'ɪŋkri:s/	(n.) a rise in the size, amount, or degree of something	5
<b>Present</b>	/'prez(ə)nt/	(adj.) existing in a place	1
		(adj.) happening now	49
		(adj.) fully focused in doing one thing	1
		(n.) the period of time that occurs now	17

	/prɪ'zɛnt/	(v.) show something for consideration	2
		(v.) show a particular state to people	3
<b>Produce</b>	/prə'dju:s/	(v.) grow	1
		(v.) create something from a chemical process	2
		(v.) cause a specific result to happen	1
	/'prɒdju:s/	(n.) fruits or vegetables	1
<b>Protest</b>	/'prəʊtɛst/	(n.) demonstration	1
		/prə'tɛst/	(v.) express objection
<b>Use</b>	/ju:z/	(v.) utilise	34
	/ju:s/	(n.) the action of using something for a purpose	31
<b>Total</b>			249

*Table 4.2 – Homographs list*

Based on table 4.2, there are only 10 homographs identified in the generated learner corpus. Among these homographs, the word ‘present’ is recorded with the highest frequency of occurrence (73), and it is used as an adjective more frequently when compared to nouns and verbs. The second highest is the word ‘use’ which showed a slight difference in the frequency of occurrence between verb and noun. It is followed by the word ‘increase’ as the third most frequently used homograph that shows the use of a verb was more than a noun. While the words ‘conflict’ and ‘impact’ have the same frequency of occurrence (20), students prefer using both words as a noun more than a verb. Based on this finding, students showed bias towards using a single meaning of the homographs.

#### *4.1.3 – Lexical Homonyms*

Lexical homonyms are words with multiple lexical meanings that are from the same word class like nouns, verbs, and adjectives. This type of homonym is classified based on senses and context of usage to distinguish it from lexico-grammatical homonyms because the homonym might have multiple meanings from a different word class in the dictionary. However, the word usage could differ from the dictionary. Hence, for lexical homonyms, only words that have all meanings in the same word class fall under this category. All definitions presented are from an online dictionary, Oxford Languages (n.d.).

Lexical Homonyms	Word Class	Definition	Frequency of Occurrence
<b>Born</b>	Adjective	a result of birth	2
		having a natural ability to perform work	1
		an idea that brought into existence	1
<b>Case</b>	Noun	an instance of a disease	1
		an instance of a situation	2
		a container	1
		the situation that affects or relates to a person or thing	10
<b>Cover</b>	Verb	a sum of money that is sufficient to pay for expenses	1
		put something on top of another to conceal it	2
		reporting of an event	1
<b>Get</b>	Verb	receive	38
		have the opportunity to do something	28
		succeed in attaining, achieving, or experiencing	26
		reach a particular state or condition	14
		making someone come, go, or move somewhere	2
<b>Light</b>	Adjective	a small quantity of food for easy digestion	3
		low in amount or intensity	1
<b>Live</b>	Verb	reside	19
		a person spends his or her life in a particular way	53
		remain alive for a specified time	4
<b>Low</b>	Adjective	below average in amount or intensity	14
		food that contains lower quantity than the usual ingredients	2
<b>Picture</b>	Noun	idea	2
		photograph	1
<b>Race</b>	Noun	competition	1
		ethnic group	33
<b>Save</b>	Verb	avoid overspending	1
		maintain healthy	1
<b>Sense</b>	Noun	aware of the presence or importance of something	14
		meaning	1
		have a feeling	3
<b>Take</b>	Verb	use as an example	1
		accept something	16
		act on the chances given	5
		react to an event	1
		achieve a result	1
		require certain amount of time	4
		consume food	6
perform an action or task	41		
<b>Total</b>			<b>358</b>

Table 4.3 – Lexical homonyms list

Based on table 4.3, there are in total 12 lexical homonyms identified in the generated learner corpus. The word ‘get’ is recorded with the highest frequency of occurrence (108) in six different meanings. Although ‘take’ is the second most frequently used lexical homonym, students knew more senses when compared to the word ‘get’, especially using it to indicate the meaning of performing an action or task. Unlike 'get' and 'take' are verbs, ‘race’ recorded as the third most frequently used lexical homonym is a noun which refers to the ethnic group rather than competition.

#### 4.1.4 - Lexico-Grammatical Homonyms

Lexico-grammatical homonyms are words that have multiple meanings and are from different word classes. As this type of homonym is similar to lexical homonyms, the grammatical class and precise sense of the homonym are analysed based on the context. Apart from determining the context of usage, observing the word class of the homonyms from the sentence structure is also prominent in counting the frequency of occurrence for each word class of the homonyms.

<i>Lexico-Grammatical Homonyms</i>	<b>Word Class</b>	<b>Frequency of Occurrence</b>	<b>Word Class</b>	<b>Frequency of Occurrence</b>	<b>Word Class</b>	<b>Frequency of Occurrence</b>
<i>Act</i>	Noun	3	Verb	7		
<i>Address</i>	Verb	4	Noun	2		
<i>Balance</i>	Verb	9	Noun	13		
<i>Bite</i>	Noun	12	Verb	4		
<i>Break</i>	Verb	6	Noun	2		
<i>Care</i>	Verb	2	Noun	7		
<i>Date</i>	Noun	3	Verb	1		
<i>Deal</i>	Verb	4	Noun	1		
<i>Drink</i>	Verb	5	Noun	1		
<i>Extra</i>	Adj.	22	Adv.	1	Noun	1
<i>Face</i>	Verb	23	Noun	1		
<i>Fast</i>	Adv.	1	Adj.	21	Noun	1
<i>Fine</i>	Adj.	3	Verb	1	Adv.	2
<i>Firm</i>	Adj.	4	Noun	1		
<i>Fit</i>	Adj.	1	Verb	3		

<i>Folk</i>	Adj.	1	Noun	1		
<i>Form</i>	Noun	4	Verb	3		
<i>Forward</i>	Adv.	5	Verb	2		
<i>Free</i>	Adj.	10	Adv.	1	Verb	2
<i>Great</i>	Adj.	68	Adv.	3		
<i>Lead</i>	Verb	49	Noun	1		
<i>Like</i>	Verb	61	Prep.	106	Conj.	11
<i>Line</i>	Noun	6	Verb	1		
<i>Match</i>	Noun	6	Verb	8		
<i>Mind</i>	Noun	20	Verb	1		
<i>Need</i>	Verb	82	Noun	7		
<i>Note</i>	Noun	1	Verb	2		
<i>Place</i>	Verb	5	Noun	11		
<i>Release</i>	Verb	5	Noun	2		
<i>Report</i>	Verb	2	Noun	1		
<i>Set</i>	Verb	7	Noun	4		
<i>Show</i>	Verb	12	Noun	1		
<i>Step</i>	Noun	16	Verb	5		
<i>Supply</i>	Verb	1	Noun	1		
<i>Through</i>	Prep.	35	Adv.	3		
<i>Visit</i>	Verb	5	Noun	1		
<i>Whole</i>	Noun	5	Adj.	17		
<b>Subtotal</b>		<b>508</b>	<b>Subtotal</b>	<b>248</b>	<b>Subtotal</b>	<b>17</b>
				<b>Total</b>		<b>773</b>

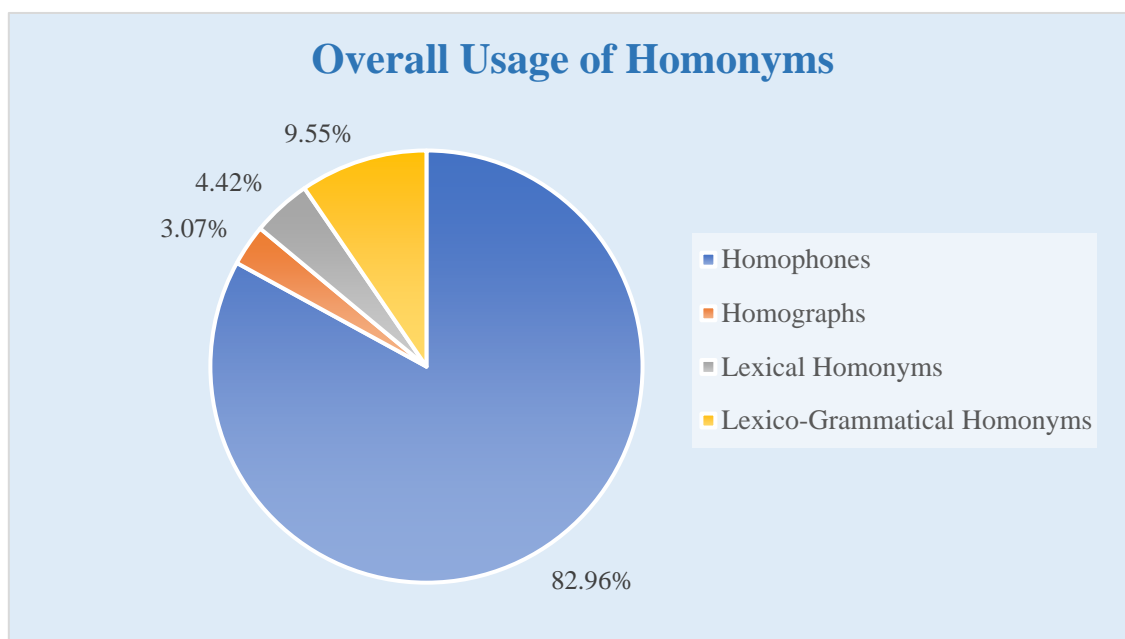
*Table 4.4 – Lexico-grammatical homonyms list*

Based on table 4.4, there are 37 lexico-grammatical homonyms found in the generated learner corpus. The most frequently used homonym is ‘like’ with the highest record of 178, and this word is used as a preposition more than as a verb and conjunction. Next, the word ‘need’ has the second highest record (89), showing an inclination result of using the word as a verb rather than a noun. It is followed by the word ‘great’ recorded as the third highest frequently used homonym (71). Similar to the previous word, ‘great’ is used as an adjective but less likely used as an adverb. The same situation of bias towards using a word class also applied to the word ‘lead’ and ‘through’. Both words are used more frequently as a verb and a preposition, respectively.



## 4.2 – Usage Pattern of Homonyms

The present study will look into the usage pattern of homonyms by analysing and interpreting the common usage, collocational pattern, and misuse pattern of homonyms found in the generated learner corpus. Besides, the current study also presented the overall usage of the homonyms as followed:



*Figure 4.1 – Chart of the overall usage of all types of homonyms*

Overall, there are in total 8098 homonyms found in the corpus. Among all four types of homonyms, homophones are the most frequently used homonym that comes in a pair or in a group of three, then it is followed by lexico-grammatical homonyms in the second place. While lexical homonym comes in the third place that has around 1.3% more than homograph which is in the last place.

### 4.2.1 – Common Usage of Homonyms

Despite the number of homonyms, this section only focused on reporting a few of these homonyms that showed significant results. Students are more familiar with using function words instead of content words in the learner corpus. Although some of the homonyms included are closed class words, the current study will only focus on reporting common usage

of homonyms that are content words like adjectives, nouns, and verbs. All definitions presented are from an online dictionary, Oxford Languages (n.d.).

### Common Usage of Homophones

Word Class & Homophone	Meaning	Key Word in Context (KWIC)
(v.) <b>Affect</b>	Influence someone or something	<p>Many people have voiced their opinion that pineapple topping on pizza is seriously <b>affect</b> the overall taste of pizza, ...</p> <p>When there are too many fakes news on the social medias, it will create uncertainties for people which will <b>affect</b> their decision making...</p> <p>In this essay, we will be discussing how multiculturalism will <b>affect</b> the education, tourism industry and the lifestyle of Malaysia.</p>
(n.) <b>Effect</b>	Causing a change resulted from the consequence of an action	<p>It can be stated that the positive <b>effect</b> of multicultural society outweighs the negative <b>effect</b> of accepting a new environment for minority groups and offence occurs...</p> <p>The first <b>effect</b> is Malaysian is high-ambiguity tolerant compare with the citizen from others country such as...</p> <p>First and foremost, one significant <b>effect</b> of fake news and false information is it will damage the credibility and image of the media...</p>

*Table 4.5 – Common usage of homophones ‘affect’ and ‘effect’*

‘Affect’ and ‘Effect’ are homophones because they have similar pronunciation and only a slight difference in spelling which has confused the learners in distinguishing them. The grammatical classes of both words are different. Although the word 'affect' can be used as a noun, it is more commonly used as a verb instead of a noun. This common usage is found in this corpus whereby all students used it as a verb because it precedes a predicate that further describes something about the subject. However, students are biased towards only one of the meanings. Based on the table, they mainly used the meaning of influencing someone or something to construct sentences even though it has other senses in a verb like pretend to have or feel something and feeling upset. On the other hand, ‘effect’ also has two different word

classes, verb and noun, but it is identified as a noun rather than a verb in the corpus. Based on the lines, ‘effect’ is a noun that used along with adjectives like positive, negative, and significant, or cardinal number such as first and second. Similarly, even though ‘effect’ also has several meanings, students preferred to use only a single sense which is the meaning of causing a change resulting from the consequence of an action.

<b>Word Class &amp; Homophone</b>	<b>Meaning</b>	<b>Key Word in Context (KWIC)</b>
<b>(v.) Advise</b>	Recommend	But it was <b>advise</b> without substance, <i>advice</i> that, when examined closely, swiftly disintegrates and reveals itself to be all dazzling style...
<b>(n.) Advice</b>	Provide guidance or suggestion to other people	What if, in this case, your friend is having trouble in choosing a course that will send them on the right career path, saying “Just do it” is useless <b>advice</b> .  Just do it is an obvious <b>advice</b> but “doing the things that you’re capable of” would be a better <b>advice</b> .  When people are stuck and need help or <b>advice</b> or maybe just some motivation from anyone, ...

Table 4.6 – Common usage of homophones ‘advise’ and ‘advice’

Similar to the previous homophone pair, ‘Advise’ and ‘Advice’ are other homophone pair that not only has similar pronunciation and spelling but also has alike meaning. The only difference that can be noticed between this pair of homophones is their word classes. ‘Advise’ is used as a verb to indicate recommend, while ‘advice’ is used as a noun which refers to providing guidance or suggestion to other people. Although this pair of homophones has two to three different meanings in a word class, students preferred to use one of them only in each word class. Since this homophone pair has a smaller number of meanings, students are less likely confused with the uses of these words. Based on the examples provided above, students used both words correctly in a different contexts. Nevertheless, as the frequency of occurrence for ‘advise’ is low, it is still possible for some students to get confused with this homophone pair.

## Common Usage of Homographs

Word Class & Homograph	Meaning	Key Word in Context (KWIC)
<b>Present (adj.)</b>	Existing in a place	The amount of sugar in bread is higher than the sugar <b>present</b> in pizza crust, ...
	Happening now	They stated that enjoying the <b>present</b> moment is cowardly to escape from the future.  ...concentrate the <b>present</b> life is the most important things in life.
	Fully focused in doing a thing	If we practise the spirit of the slogan, we can be trained to be <b>present</b> and make complicated thoughts in an easier way.
<b>(v.) Present</b>	Show something for consideration	Pineapple maybe rich in nutrients, maybe appealing to see or <b>present</b> and most importantly...  Although some do <b>present</b> customer with the option to remove or add certain ingredients namely, ...
	Exhibit a particular state or appearance to people	Some may argue that a non-detailed plan of execution will make people easy to fail and it will <b>present</b> a crisis in their future.  ...multicultural education provides Malaysian with educational experiences that helps to maintain their commitment in cultural community that acquire to <b>present</b> a racial and ethnic harmony community.
<b>(n.) Present</b>	The period of time that occurs now	...financial planning is essential rather than living in the <b>present</b> and getting what you want in life right now.  There are individuals who assume living at the moment is equivalent to enjoy the <b>present</b> and need not to worry about the future.

*Table 4.7 – Common usage of homograph ‘present’*

Next, ‘Present’ is identified as a homograph because it has the same spelling but has two different pronunciations for its word classes. As each word class has more than one meaning except for nouns, it is crucial to look at how students used these words with different meanings. Based on the examples given, students used ‘present’ in three different word classes,

adjective, verb, and noun. Despite the ambiguity, students did not encounter any problem in using ‘present’ to construct sentences since they demonstrated a high level of accuracy for all usage of the word in the generated corpus. This result indicates students would not easily get confused with the word used when the homonym is a common term used by students to construct sentences. In other words, as students are exposed more to the word, they could understand the meanings of each word class better and possibly use them correctly even if it has multiple senses.

<b>Word Class &amp; Homograph</b>	<b>Meaning</b>	<b>Key Word in Context (KWIC)</b>
<b>(v.) Impact</b>	Have a strong effect on people or thing	For example, culture can <b><u>impact</u></b> speaking inclinations such as figurative language, speed, and presentation.
<b>(n.) Impact</b>	Marked effect or influence	<p>...multicultural society can leave a positive <b><u>impact</u></b> on a country as well as the citizen of that country.</p> <p>It brings a great idea to decrease the people to spreading the fake news, because the fakes news brings a lot of the negative <b><u>impact</u></b> to Malaysia’s society, ...</p> <p>The <b><u>impact</u></b> of cultural conflict can be serious.</p>

*Table 4.8 – Common usage of homograph ‘impact’*

‘Impact’ is classified under homograph as it also has two different pronunciations for both verbs and nouns. Although it has two different meanings for each word class, students only used one of the meanings for each word class. Based on the table, the meaning of the verb and the noun is almost the same, the only difference here is the word class used. ‘Impact’ is certainly easier to use in building sentences as it is not affected by the similar meanings of verbs and nouns. Students could use the word precisely in different contexts regardless of the word usage in both verb and noun.

### Common Usage of Lexical Homonyms

Word Class & Homograph	Meaning	Key Word in Context (KWIC)
(v.) Live	Reside	<p>We <b>live</b> in a society where people are judged based on their gender, skin colour...</p> <p>...the living cost of Malaysia is very low, which causes many foreigners to want to <b>live</b> in Malaysia.</p> <p>Even if we bring our families to <b>live</b> overseas, we must consider whether or not they will be able to grasp the contemporary atmosphere.</p>
	A person spends his or her life in a particular way	<p>...people from the west <b>live</b> a <i>life</i> full of freedom, while people from the east <b>live</b> a <i>life</i> full of labour and work...</p> <p>People of different cultures can <b>live</b> in harmony and thus become a more knowledgeable and secular society.</p> <p>...they will achieve the sense of joyfulness as they <b>live</b> in the way they desire and the way they think is the best.</p>
	Remain alive for a specified time	<p>The reason given is that parents <b>live</b> longer than young people, thinking that they know what is better for young people...</p> <p>It is true as there are a lot of uncertainty exist in our <i>life</i> and we just only <b>live</b> once.</p>
Live in the past (phrase)	Reminisce about past events.	<p>...when people keep considerate the unforeseen consequences and <b>live in the past</b>, they cannot live with joy.</p> <p>...they cannot prove that neither all people that don't live in the present will fail in their future nor all people that <b>live in the past</b> will succeed in the future.</p> <p>Mature ones should <b>live in the past</b> to allow you to design a better future.</p>
Live with (phrase)	Tolerate or accept something unpleasant	<p>most of them <b>live with</b> regrets in their late lives, regretting for not accomplishing their longing desire...</p>

Table 4.9 – Common usage of lexical homonym 'live'

A lexical homonym refers to a word that has different meanings in the same word class. 'Live' is classified under this category because it has multiple meanings under the same word class. Supposedly, it is categorised under homograph because it has different pronunciations for two word classes. However, in this corpus, 'live' is only used as a verb instead of an adjective. Although students do not use this word as an adjective, its pronunciation is similar to another word, life. In this corpus, students showed confusion in using these two words not because of similar pronunciation but spelling. Hence, it is vital to observe the common usage of 'live' used by students. In this case, 'live' is a lexical homonym that is mainly used as a verb, and only a minority of it is used in phrases. For the verb usage, students used it in three different meanings, while the meanings of phrases are not the same.

Word Class & Lexical Homonym	Meaning	Key Word in Context (KWIC)
(v.) Get	Receive or come to have something	<p>There are also instances where people faked cancer to <b>get</b> a large donation from the public to <b>get</b> treatment from the hospital.</p> <p>Every time I <b>get</b> a salary, my parents do not have to worry about my study fees and living expenses, ...</p> <p>...some bloggers will exaggerate their words or even write fake messages to <b>get</b> more attention and followers.</p>
	Obtain or succeed in achieving something	<p>...people should enjoy the present moment and <b>get</b> what they want in life right now.</p> <p>The students need to retake again in order to pass the exam and <b>get</b> the SPM certificate.</p> <p>... he could face the interview directly by not thinking whether he can <b>get</b> the job or not but knowing...</p>
	Reach a particular state or condition	<p>When people know the real facts and find that they have been deceived, they will <b>get</b> angry and...</p> <p>As people <b>get</b> older, people will get countless things to worry.</p>

		Many pizza shop did promote healthy pizzas like low fat pizza which may decrease the percentage for people to <b>get</b> obesity.
	Have the opportunity to do something	<p>Pizza has always been popular because customers <b>get</b> a chance to customize toppings on a pizza.</p> <p>Customers <b>get</b> to choose any topping that suits their preferences, which makes it unique for each customer.</p> <p>Malaysians <b>get</b> to learn different culture easily as they have close contacts with other ethnics in daily life.</p>
<b>Get along (phrase)</b>	Maintain a harmonious relationship	<p>... racial and ethnic harmony can be achieved, meaning that individuals from diverse backgrounds <b>get along</b> well.</p> <p>Every Malaysian need to learn how to accept and <b>get along</b> with others race while living in a multicultural society.</p>
<b>Get rid (phrase)</b>	Taking action to remove unwanted things	<p>The children might try to eat fruit because the attractiveness of the pizza which can <b>get rid</b> of the kids' bad habits of picky eating.</p> <p>This advertisement told you that you need to <b>get rid</b> of your laziness and follow your heart decision.</p>
<b>Get into (phrase)</b>	Feelings that influence or affect a person	...everyone will <b>get into</b> deeper relationship with gain more trust and confidence with one another and...
<b>Get over (phrase)</b>	Overcome a difficulty	<p>As long as you keep trying, you will <b>get over</b> every obstacle and eventually make it.</p> <p>...it is most applicable as a motivation booster, and to help people <b>get over</b> obstacles in life, ...</p>
<b>Get up (phrase)</b>	Rise after awaking from bed	The slogan makes me <b>get up</b> early in the morning to finish my daily tasks such as housework and workouts, ...
<b>Get out (phrase)</b>	Leave	The greatest way to avoid feeling hopeless is to <b>get out</b> of bed and do something.

*Table 4.10 – Common usage of lexical homonym 'get'*

'Get' has only a one-word class with many different meanings, it is categorised under lexical homonym. Based on the lines given above, 'get' is mainly used as a verb with several meanings, as well as used in multiple different phrases. The word is commonly used as a verb,



students knew most of the word senses, and they managed to use it precisely in different contexts and phrases. Students showed high accuracy in using this word because it is a general term that is frequently used to construct sentences, and more exposure to the word helps to strengthen their memory towards the usage of the word in different contexts.

### Common Usage of Lexico-Grammatical Homonyms

<b>Word Class &amp; Lexical Homonym</b>	<b>Meaning</b>	<b>Key Word in Context (KWIC)</b>
<b>(v.) Lead</b>	Causing an event or consequence	<p>...borrowing money or owning debts for the use of present may <b>lead</b> to the problem of bankruptcy in future.</p> <p>...fighting corruption among the Malaysians, the socio-economic may be affected and this could <b>lead</b> to a problem...</p> <p>... processed meat contains fat that might <b>lead</b> people to obesity if people consume it for a long period.</p> <p>no matter how determined or robust we believe we are, our brain will <b>lead</b> us to give up more often than it will encourage us to succeed.</p>
<b>(n.) Lead</b>	A competitor is ahead of the others with a certain amount	For instance, the analysis discovered that the fictional Chinese candidates had a massive <b>lead</b> in callbacks.

*Table 4.11 – Common usage of lexico-grammatical homonym ‘lead’*

‘Lead’ is also supposedly classified under homographs as it has two different pronunciations /li:d/ is used as a verb, while /lɛd/ is a noun. Nevertheless, in this corpus, no student used the word meaning of the second pronunciation, which refers to the metallic element. Most of the students are more preferred to use the word senses of the first pronunciation. As the first pronunciation has two different word classes, the lexical and grammatical meanings also differ. Therefore, ‘lead’ is categorised under the lexico-grammatical homonym instead of homograph. Based on the examples provided, most students used the word as a verb in a different context, while only a student used it as a noun. It showed

that the students preferred to use the word as a verb rather than a noun. It indicates that students did not have the same exposure to meanings for both verbs and nouns.

Word Class & Lexical Homonym	Meaning	Key Word in Context (KWIC)
(v.) Need	Require something which is very important or essential	...we <b>need</b> to be highly motivated and work hard toward our target, so we can get what we want in life right now.  ...the founder of Nike think that this brand will <b>need</b> a tagline or slogan that give unity to everyone, ...  Sometimes I believe what we truly <b>need</b> is a drill sergeant in our heads continually pushing us to buckle down and...
(n.) Need	Require or want a thing	Consumption of this cheese helps to satisfy the immediate nutritional <b>need</b> .  ...the pizza will always be there for those people who are in <b>need</b> .
	Circumstances in which something is necessary	...there is a strong <b>need</b> for government or mass media to organize an awareness campaign on the importance of living in a multicultural society.

Table 4.12 – Common usage of lexico-grammatical homonym ‘need’

Supposedly, ‘need’ is categorised under homophones as it has the same pronunciation as another word, ‘knead’. However, students did not use this word, but they used the word ‘need’ as nouns and verbs in the learner corpus. Therefore, when a word like ‘need’ has meanings in different word classes, it is classified as a lexico-grammatical homonym. Unlike ‘lead’, students are not biased in using either verb or noun of ‘need’. As the senses of the word in both verbs and nouns are similar, students used them without any problem or error. It indicates that they are familiar with the lexical item and its meaning.

#### 4.2.2 – Misuse Pattern of Homonyms

As homonyms are ambiguous, students are confused by their similarity in spelling, pronunciation or meaning, which has caused them to misuse homonyms. In this corpus, students misused all types of homonyms except for homographs. Based on the learner corpus, students have committed errors related to semantics, word choice, L1 interference, grammar, and word class. The keyword-in-context (KWIC) tool is utilised to report the misuse pattern of the homonyms in the learner corpus.

#### Misuse of Homophones

20	88.txt	there are a few effects of living in a multicultural society in Malaysia. The first	affect	of living in multicultural society in Malaysia is multicultural society helps to gain more
28	71.txt	might face some problems in the relationship and lifestyle between them. Component that can be	effect	living in a multicultural society in Malaysia by various cultural, type of languages and
29	71.txt	Malaysia by various cultural, type of languages and education. One of the components that can	effect	living in a multicultural society in Malaysia which is different types of cultural. Malaysia
30	71.txt	society has shows the cultural of different in Malaysia. Languages is another components that can	effect	living in a multicultural society in Malaysia. In Malaysia, the most common languages which
31	71.txt	multicultural society in Malaysia. In conclusion, different cultural, types of languages, and higher education can	effect	living in a multicultural society in Malaysia. Malaysia is a special country as it
32	82.txt	of different races, so we need to use different languages to communicate. In conclusion, the	effect	living in a multicultural society in Malaysia is what I mentioned in above that
33	84.txt	well as the national language of Malaysia which is Malay need to learn are the	effect	living in a multicultural society in Malaysia. Those effects need to pay attention as

*Figure 4.2 – Misuse lines of homophones ‘affect’ and ‘effect’ taken from the concordance software (AntConc)*

The first homophone pair that is misused by students is ‘affect’ and ‘effect’. Although ‘effect’ occurred more than ‘affect’ in the corpus, it is more frequently misused by students. Based on the examples given, students show confusion in using both words. For instance, a student misuse the word ‘effect’ to represent the correct verb ‘affect’ from lines 28 to 31. Similarly, another student misuse ‘affect’ as the noun ‘effect’ in line 20. Both words are misused interchangeably by students which have demonstrated that students are unaware of word class could be used to identify

their differences. Apart from that, students also made the grammatical mistake of not adding a preposition ‘of’ to connect both nouns, ‘effect’ and ‘living’ in lines 32 and 33.

1 38.txt	of dough afterwards, creating an uncomfortable stomach, all because of the desire to cleanse their	palette,	which can be achieved by pineapple on pizza. Other than that, pineapple on pizza
----------	--	----------	--

*Figure 4.3 – Misuse line of homophone ‘palette’ taken from the concordance software (AntConc)*

Unlike the first pair of homophones, the incorrect usage only falls on the word ‘palette’. The meaning of the word refers to different colours. However, in this line, it means the upper part of the inner mouth, but instead of using the correct word ‘palate’, the student chose to use its homophone pair, palette. This wrong usage indicates that the student is not fully aware of the multiple meanings of the word ‘palate’ which caused the student to select another homophone ‘palette’ that has a similar sound with its correct form.

### **Misuse of Lexical Homonym**

2 34.txt	also does not take long, which gives it a bonus point. Also, chicken does not	cover	the taste of the sauces or other existing toppings on the same pizza as
3 79.txt	to be more open minded, flexible and tolerant because Malaysia as a country alone, have	cover	a lot of unique and amazing races and people can have a better understanding

*Figure 4.4 – Misuse lines of lexical homonym ‘cover’ taken from the concordance software (AntConc)*

‘Cover’ is usually used to indicate multiple meanings such as putting something on top of another to hide it, reporting an event, and having enough money to pay for expenses. However, based on the lines given, the meaning of the word ‘cover’ is misinterpreted by students who thought that the sense of the English word ‘cover’ has the same meaning as the word in their mother tongue. Both misuse lines here are affected by the Mandarin language. In line 2, the word ‘cover’ has its Mandarin equivalent word, 盖 /gài/. This word was nativized and experienced a semantic

shift. It means another substitute has replaced the food's original flavour, and such sense is localised and usually appears in conversation. While in line 3, 'cover' is believed to be directly translated from another Mandarin word, 涵盖 /hángài/ with a different meaning which refers to comprise. This result proved that although the English word has its Mandarin equivalent, the sense of the Mandarin word is different from the English word. These two misuse examples showed that the students are influenced by their first language while using English words to build sentences. In other words, students experienced L1 interference while using homonyms to write essays.

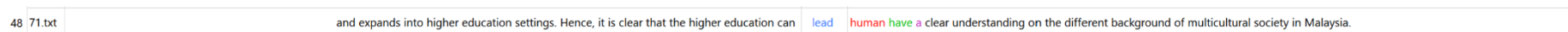
1	126.txt	to do. "He said, "One day, you'll leave this world behind, so live a	live	you will remember". Taken from the lyrics of the popular song "The Nights", written
22	95.txt	languages in Malaysia, for instance, Malay, English, and Chinese, other than that, the people who	live	in Chinese culture can learn Hokkien and Mandarin. In addition, the English language is
64	106.txt	think about the consequence when they do something. The bad consequence might affect people's	live	and have a negative impact on themselves. However, when people think too much on
66	29.txt	sword because consuming too much of pizza with unhealthy toppings can be detrimental to humans	live	but at the same time pizza can be a healthy food with healthy toppings.
67	145.txt	experiences or different value that preferring on alternative . Since, People have different experience in their	live	but should enjoy the present moment in life right now. All of these selected
69	135.txt	slogan of a prestigious sportswear company – Nike. A common perspective of people nowadays is that	live	is too short, so we should live every day to the fullest and leave
74	72.txt	a multicultural country. However, some people claim that multicultural society brings some effect to their	live.	Most of Malaysian are not aware of the positive effects of multicultural society in
95	55.txt	fusion of Malay, Chinese, Indian, and European elements. In Malaysia, different race has their own	live	style, we keep touch with each other although the culture is different. However, having
96	55.txt	conditioning. Overall, multicultural society have a positive effect to Malaysia. It can affect a person	live	style. It encourage people to more open to different cultures. As such this paper
97	128.txt	also despise crisis. First of all, "just do it" can motivate teenagers to achieve the	live	that they wanted. Therefore, they will put more effort on the target process. For

*Figure 4.5 – Misuse lines of lexical homonym 'live' taken from the concordance software (AntConc)*

The word 'live' has a pronunciation /lɪv/ which is similar to the word 'life' and its plural form 'lives', some of the students do not know how to distinguish their differences. 'Live' can be used as a verb and an adjective, but in the corpus, no student used 'live' as an adjective. Most of the students used the word as a verb which has a different pronunciation from the noun 'life' and 'lives'. In the corpus, students misuse 'live' as the singular form of the noun, 'life' in lines 1, 69, 95, and 96. On the other hand, 'lives' is the plural form of 'life' which is also misrepresented by 'live' in lines 64, 66, 67, 74, and 97. Based on this result, it pointed out that students are not aware of the word class for both 'live' and 'life'

which caused them to misuse the word 'live'. Besides, a student has committed a word choice error whereby the selected word 'live' does not fit into the context in line 22. In this line, the word 'live' cannot be used in this context as the intended meaning expressed by the writer was incorrect, and the correct word should be 'practice'. This misuse line proved that the student did not know the meaning used was not in the dictionary entry.

### **Misuse of Lexico-Grammatical Homonym – Lead**



48 71.txt and expands into higher education settings. Hence, it is clear that the higher education can lead human have a clear understanding on the different background of multicultural society in Malaysia.

*Figure 4.6 – Misuse line of lexico-grammatical homonym 'lead' taken from the concordance software (AntConc)*

The word 'lead' is misused when the student would like to express the meaning of commanding or being the one in charge in this context. However, the student misunderstands the meaning of the word whereby 'lead' refers to a person or people who are the ones commanding or the person in charge, but not referring to a subject who will give the command. It clearly indicates that the student does not fully comprehend the meaning of 'lead' and does not know how to use this word in a sentence. The incorrect usage of the word could mislead the readers when the intended message is not conveyed accurately through the sense.

### 4.2.3 – Collocational Pattern of Homonyms

The ambiguous nature of homonyms makes the process of acquiring English homonyms more challenging, and learners might not be able to memorise all the meanings and usage of the homonyms. Hence, it is significant to study the collocational pattern of homonyms to know what words come beside or along with the homonyms that could facilitate the learning process. This section will focus on finding collocates that co-occur with different types of homonyms by utilising the collocate tool in the corpus. The collocations presented are explained based on grammatical and lexical collocation.

#### **Collocates of Homophone**

##### **Effect**

Collocate	Rank	FreqLR	FreqL	FreqR
multicultural	1	30	9	21
living	2	19	1	18
society	3	20	7	13
negative	4	9	9	0
malaysian	5	12	1	11
positive	6	6	6	0
brings	7	6	5	1
outweighs	8	2	1	1
second	9	5	5	0
habit	10	2	1	1
another	11	4	3	1

*Figure 4.7 – Collocate list of homophone ‘effect’ in the generated learner corpus*

Based on the table, words that existed at the left or before ‘effect’ are negative, positive, and second. Positive and negative are adjectives that usually modify the noun, as ‘effect’ is a noun, so these two words precede it. While second is an ordinal number which can be an adjective or verb. The word is used as an adjective instead of a verb which appears before the word ‘effect’ in the learner corpus. Moreover, other words that appeared on both sides or before

and after ‘effect’ are bring and another. Both words are commonly used at the left or before ‘effect’ and seldomly come after. Hence, the only way to recognise ‘effect’ as a noun is an adjective that comes before the word because another homophone pair, ‘affect’ is a verb that is usually not used together with an adjective. Other collocates like multicultural, living, society, Malaysian, outweighs, and habit appeared on both sides of the homophone. These words are often found within a sentence that rarely existed besides the homophone. The uses of these words also signified various contexts of usage found in the learner corpus.

### **Collocates of Homograph**

#### **Conflict**

Collocate	Rank	FreqLR	FreqL	FreqR
ethnic	1	5	4	1
tension	2	2	2	0
between	3	3	0	3
cause	4	3	3	0

*Figure 4.8 – Collocate list of homograph ‘conflict’ in the generated learner corpus*

In the generated corpus, ‘conflict’ is usually surrounded by words like ethnic, tension, between, and cause. Among these collocates, only the preposition ‘between’ appeared after the word ‘conflict’. This word usually describes two parties that are having conflicts. On the other hand, ‘cause’ is a verb that usually appears before a noun, and it is attached to the noun ‘conflict’ that comes as a pair. Besides, other words that correlate to ‘conflict’ are ‘ethnic’ and ‘tension’. Both lexical items are from different word classes, but when these words co-exist together in a text, the students can predict the homonym used since they are related to each other.



## Collocates of Lexical Homonym

### Live

Collocate	Rank	FreqLR	FreqL	FreqR
together	1	10	0	10
people	2	33	26	7
we	3	18	10	8
in	4	45	8	37
harmoniously	5	3	0	3
pizza	6	2	0	2
moment	7	8	1	7
partner	8	2	1	1
it	9	6	2	4

*Figure 4.9 – Collocate list of lexical homonym ‘live’ in the generated learner corpus*

Based on the collocate list, words that usually appeared at the right of the word ‘live’ are together and harmoniously are used to modify the verb. As these two words are adverbs, they usually come after a verb. While other words that commonly exist on both sides or before and after the word ‘live’ are people, we, in, and it. These words are nouns, pronouns and prepositions that could appear in any part of a sentence, but ‘live in’ is commonly used by learners which means stay in a place. Other collocates such as pizza, moment, and partner are words that do not usually co-occur with ‘live’, but provide a clue on the context. Since learners are often confused between ‘live’ and ‘life’, the only method used to differentiate them is using an adverb that comes after a verb.

## Collocates of Lexico-Grammatical Homonym

### Lead

Collocate	Rank	FreqLR	FreqL	FreqR
to	1	52	2	50
will	2	15	15	0
obesity	3	3	0	3
may	4	8	8	0
accusations	5	2	0	2

*Figure 4.10 – Collocate list of lexico-grammatical homonym ‘lead’ in the generated learner corpus*

The words that appeared at the right or after the word ‘lead’ are will and may. These two words are modal verbs that are next to verbs most of the time. When modal verbs come after the word ‘lead’, it is used as a verb instead of a noun. On the other hand, the preposition to existed before and after the word ‘lead’ but appeared more after the word. For example, ‘lead to’ is commonly used by students to denote the meaning of causing a consequence or an event. As other collocates like obesity and accusations mean the result or outcome of actions, they usually come after the verb form of ‘lead’, which can further clarify its grammatical class and meaning since it has multiple senses in two different word classes.

## CHAPTER 5 – DISCUSSION

### 5.1 – Research Objectives

The current study highlights two different research objectives. The first research objective is to identify the type of homonyms used in the writing of Malaysian ESL learners. While the second research objective is to report the usage pattern of homonyms in the writing of Malaysian ESL learners. The researcher will discuss data and findings pertaining to the research objectives and point out several aspects that contribute to the problems or issues faced by Malaysian ESL learners while using homonyms in their writings.

#### *5.1.1 – Research Objective 1*

As the first research objective is to identify the types of homonyms used in the writing of Malaysian ESL learners, the researcher has identified 4 types of homonyms – homophones, homographs, lexical homonyms, and lexico-grammatical homonyms in the generated learner corpus. Each category of the homonym has its own traits or properties, only homonyms that have the same or similar characteristics are classified under the same category. According to the Usage-Based Theory, the categorisation process aims to help learners to connect varied inputs into equivalent classes when they exhibit the same or similar properties (Bruner et al., 1956, as cited in Bybee & Beckner, 2015). For instance, words like ‘present’, and ‘impact’ are homographs that have the same spelling but are pronounced differently for a different meaning. During the categorisation process, the researcher needs to analyse different meanings of homonyms. This process involved the use of dictionary and projection rules as mentioned in Katz's Semantic Theory to disambiguate meanings while observing the collocates used in the sentence and examining whether the sense of the homonym used fits into the context (Falkum, 2011).

Apart from applying Katz's Semantic Theory while classifying homonyms, the context of usage also affects the interpretation of the word senses as the contextual information would help to disambiguate different meanings of homonyms. The presence of context helped activate relevant word senses in the mental lexicon that caused a constraint on semantic selection (Tabossi et al., 1987). Although homonyms have more than one sense in the dictionary, learners would not select most or all meanings but are inclined to choose a sense that matches the context to form sentences. The current findings suggested that most students access the senses of homonyms under the influence of dominance and context effects. As such, context deals with the relevancy of the dominant meaning based on the semantic information provided in the text, which is crucial in examining whether the sense of the homonym used is appropriate in a particular context (Tabossi et al., 1987; Simpson, 1984). Thus, the context helped to determine the exact sense of the homonym, while the dictionary was only used as a reference to clarify the meaning used in a sentence (Nouraldeen, 2015). Without the context, learners cannot identify the differences in the senses of the homonyms used in the text (Ushiro et al., 2010). Therefore, sense and context played a significant role in the categorization process because they are part of the components used to classify homonyms into different categories.

Similarly, Simpson (1984) stated that not only context and meaning are crucial in semantic activation, but the meaning frequency is also equally important to the access of multiple senses of homonyms. Each homonym has included the frequency of occurrence for meanings and words. Based on the frequency of homonyms, the researcher decided to focus on these two aspects, function words and content words. Although a minority of the homophones and lexico-grammatical homonyms are closed-class words, they showed a higher frequency of occurrence when compared to the content words. In other words, students are more familiar with using the function word homonyms as there are disparities in the frequency of occurrence between content words and function words for homophones and lexico-grammatical homonyms in the

generated corpus. As function words are commonly used in the English language; therefore, students recognise their usage in the text (Pennebaker et al., 2014). In the current study, the function word homonyms demonstrated a high frequency of occurrence in their respective categories. This finding justified that learners tend to retain memories of high-frequency items when these words are presented in a category as stated in the Usage-Based Theory (Nosofsky, 1988, as cited in Bybee & Beckner, 2015). Similarly, another study asserted that high-frequency words would leave a positive effect in facilitating vocabulary learning when these lexical items are presented repeatedly in front of the learners (Peters, 2020).

On the other hand, while noticing the frequency of the content word homonyms, the researcher found that more than half of the content words for homographs and lexical homonyms showed a higher frequency of occurrence in one meaning when compared to the others. This result demonstrated that students are biased towards a sense when using homonyms (Ushiro et al., 2013). In other words, learners are usually dominant in using a meaning, and the bias level toward one sense affects the word selection process. This assumption is supported by scholars whereby learners are more likely to choose a subordinate meaning of the homonym when the contextual information has a stronger biased toward this meaning otherwise dominant sense will be selected (Simpson & Burgess, 1985; Duffy et al., 2001). Besides, the evidence of bias can be seen in the frequency of occurrence for the word class of lexico-grammatical homonyms. While for homophones, students preferred to use one of the homophone pairs. All these biases indicated that students are more favourable to learning words in the form of one-to-one mapping between the lexical and semantic structures even though homonyms are processed as many-to-one mappings from meanings to signals (Smith, 2004). This result showed that students are still under exposure to the subordinate meanings of homonyms and if students learn more about different senses of homonyms through their usage patterns in multiple contexts.

### *5.1.2 – Research Objective 2*

Next, the second research objective in the current study is to report the usage pattern of homonyms in the writings of Malaysian ESL learners. The researcher discovered the usage pattern of homonyms through the overall usage, common usage, misuse pattern, and collocational pattern. All these aspects would enable learners to recognize the differences in usage for different types of homonyms.

In the generated learner corpus, the overall usage of homonyms is relatively low as it constituted only about 9.5% out of the total of 84940 words. This finding contradicts the past study, Rodd et al., 2002, (as cited in Demir, 2020) mentioned that over 80% of the common English words are homonyms. The corpus size and the meaning used by students are the factors that caused a low result in the finding. Firstly, the corpus size is relatively small, which affects the number of homonyms found in the corpus. Secondly, students only used a single meaning of the homonyms to construct sentences. These homonyms are excluded from the findings as it does not reflect the ambiguous nature of the homonyms. Regarding the overall usage of homonyms, homophones are the most frequently used homonym partly because they are in a pair or a group of three and comprise function words and content words that are absent in other homonym categories except the lexico-grammatical homonyms.

Next, the common usage of the homonyms reported in the present study seeks to analyse the meanings of homonyms used by students to form sentences. It often involves how students use multiple senses of homonyms in different contexts regardless of the frequency. In the learner corpus, a few of the homonyms are general terms that learners are more likely to use, and the meanings of these terms are simple to comprehend even though they are ambiguous. When a specific context is available, learners can identify the exact sense of the homonyms effortlessly. However, sometimes this could pose difficulties for learners to use homonyms accurately as they might be confused with the similar traits of different types of

homonyms. Therefore, learners tend to use generalised words in different contexts since the meaning of the words is not restricted (Blum & Levenston, 1978). As these homonyms are commonly used by learners, with frequent exposure to the usage, they seem to be more familiar with the homonyms. Eventually, they are less likely to commit errors when using homonyms since these repeated lexical items would retain in their memories (Laufer, 1990).

Moreover, while analysing and interpreting the senses of homonyms used within the sentence, Katz's Semantic Theory played a significant role in disambiguating the meanings of the homonyms to ensure they are used accurately in different contexts (Falkum, 2011). Apart from determining the sense of the homonym through context, the grammatical classes of the word assist learners in distinguishing different homonym categories except for lexical homonyms that have several meanings in the same word class. When a homonym uses along with other words, it usually forms a subject-verb-object sentence signalling the word classes of the homonym and co-occurrence words. These co-occurrence patterns of the grammatical classes have facilitated the syntactic processing of the texts, especially for sentences that contain ambiguous words (Hirschman et al., 1975). Besides, it also suggested that word classes are the features that co-occurred in most of the homonym categories except for lexical homonyms that could help in maintaining knowledge about the meanings of each homonym (Malt & Smith, 1984, as cited in Bybee & Beckner, 2015). It has highlighted the importance of learning word class and word meaning together while acquiring homonyms because without word class knowledge, learners might misuse homonyms to build sentences.

Furthermore, the common usage allowed the researcher to interpret multiple senses of the homonym used by learners in different contexts through the corpus. The concordance lines presented in the finding show how different meanings of homonyms are used to construct sentences as contextual information is available to disambiguate meanings (Al-Harbi et al., 2011). Based on the findings, a few homonyms showed semantic similarity even though they

are from different word classes. Since the meanings of the homonym are similar, it does not pose difficulties for students to use this kind of homonym accurately in multiple contexts. This result has demonstrated that the semantic similarity between senses of different word classes of some homonyms would facilitate the process of acquiring and retrieving new meanings as they share the same feature that could help to integrate the new senses into the semantic network easily (Eddington & Tokowicz, 2015). Nevertheless, the present study shows contrast with Pustejovsky, 1995, (as cited in Yurchenko et al., 2020) stated that homonyms have unrelated senses that would delay the word recognition process. In fact, only some of the homonyms in the current study showed semantic similarity, and these similar senses have two meanings from each different word class. Hence, students could easily recognise the homonyms without paying attention to the word classes.

On the other hand, the ambiguous nature and similarity of the lexical and phonetic structures of homonyms often cause confusion among learners. Most of the time, scholars believed that learners would certainly confuse different types of homonyms, especially homophones and homographs. Vethamaiccam & Ganapathy (2017) stated that students are confused about the homophone pairs due to the similarities in pronunciation. In another past study, Demir (2020) claimed that the numerous senses of the homographs hindered students from learning them. However, the present study found that confusion not only applied to the homonym categories but also extended to non-homonymous words that have identical spelling to the homonyms. For instance, 'live' is a lexical homonym, while 'life' is not a homonym in the generated learner corpus, but both words are misused interchangeably by students as they are confused about the word classes of these two words. As students were not fully aware of the grammatical classes, they did not realise the differences in the meanings of both words that caused them to misuse the homonym. This result is similar to Alghamdi (2021) whereby learners cannot identify the word classes of homographs that led to the misuse of the



subordinate meaning. Thus, it inferred that learners need to expose more to the word classes and usage of the homonyms in different contexts to reduce the probability of misuse.

Additionally, students also committed word choice errors of choosing inappropriate words to use in a particular context, for instance, words like 'live' and 'life'. This problem could have resulted from the learners' confusion in understanding multiple senses of the homonyms as mentioned in the previous studies (Jacobson et al., 2007; Saleh, 2017). Nevertheless, another possible cause of misuse is the intended meaning expressed by the learner does not tally with the definition in the dictionary. This finding implied that learners are unaware of all the word meanings in the dictionary and do not fully comprehend the meanings when there is minimum contextual information provided in the dictionary. Rodd (2017) asserted that a dictionary only included a small subset of the word usage used in natural conversation. As a result, learners could not gain sufficient information on how the word meanings are used in different contexts. In the current study, students also showed confusion in using homophone pairs. For instance, 'palate' is misrepresented by its homophone pair 'palette'. This result indicated that learners are unaware of the meanings of the homophone pairs, and they are confused about the similarity in pronunciation. Yet, Starr and Fleming (2001) believed that the word knowledge of the learners also affects their decision in selecting the right homophone pair that fits the context of usage.

Besides, the word acquiring process became more challenging for students as homonyms usually have a few unrelated senses. As homonyms are ambiguous, learners hardly remember all the senses of the homonyms. Even if they manage to memorise the meanings, it is still possible for them to misunderstand the word's sense when the homonym is in isolation without any context. This issue of misunderstanding word senses happened in the previous study. Ibrahim (2018) stated that undergraduates did not understand the meanings of homophones and homographs which has affected how they use them in their writing. Similarly,

the present study has also found that learners misused lexico-grammatical homonyms in terms of not fully comprehending the word's meaning and not knowing how to use it to build sentences. When this misunderstanding of sense occurs, the right message is not conveyed accurately to the readers because writers produce misleading sentences. Therefore, scholars suggested that if background information is present, it might help learners to avoid the issue of misunderstanding meanings when they encounter ambiguous words (Liu, 2012).

Apart from that, most Malaysian students learned English as their second language, and the differences in the language system between the English language and their mother tongue have directly affected how learners process and use the language. As most of the UTAR undergraduates are Chinese students, the essays included in the generated corpus for the present study are mostly students from this ethnicity. Based on the result, students experienced L1 interference as learners directly translated words from the first language to the English language. This problem occurred when learners thought that the translated words fit into the context of usage (Vethamaiccam & Ganapathy, 2017). Yet, they are unaware that English and Mandarin words might not share the same meaning (Dipolog-Ubanan, 2016). Both studies have supported the current finding whereby students thought the meaning of the English homonym is the same as the Mandarin equivalent and matched with the context. In fact, they experienced L1 interference that caused them to misuse the word. Besides experiencing L1 interference, learners also tend to nativize their first language as the original meaning of the Mandarin word change to a new localised meaning (Mohd Nasir, 2021). This semantic shift resulting from their mother tongue is transferred directly to the English language, and it eventually causes the misuse of the homonym.

Since the learners experienced difficulties in acquiring homonyms, they must study the words that co-occur with the homonyms. These co-occurrence words help learners to make assumptions about the meaning of the homonym used in different contexts. Scholars used this

collocation benefit in their study to examine the relationship between prepositions and co-occurrence words that help to disambiguate sentence meaning (Yunus & Awab, 2012). Unlike the previous study, the current research found that most of the collocates are commonly used by learners in different contexts. Although the frequency of usage for some collocates is not high due to the small corpus size, the collocation found in this learner corpus showed authentic language use. It highlighted the collocational pattern used in the local context. Daskalovska (2015) stated that authentic text generated by local learners would facilitate collocation learning when they utilised corpus as a tool to acquire homonyms. The collocate tool provided a list of co-occurrence words which allowed learners to observe and identify the differences between homonym categories. Apart from that, each of the homonyms demonstrated in the present study has a different set of collocations that rarely overlap with each other. It could be a practical tool in identifying the word classes and disambiguating the senses of the homonyms (Bolshakov & Gelbukh, 2001). This assumption is supported by Yarowsky (1993) claimed that a collocation has a higher chance of identifying only a single meaning for an ambiguous word.

Furthermore, the current study highlighted grammatical and lexical collocations. Grammatical collocation refers to the fixed rules of word classes of a set of collocations that have grammatical functions (Aarts et al., 2014). For example, a noun, verb, or adjective is followed by a preposition, adverb, or infinitive in a phrase (Diah Moehkardi, 2012). In the current study, phrases like 'live in' and 'lead to' are examples of verb patterns of grammatical collocation found in the generate learner corpus. While lexical collocation usually showed a predictable connection between a word class and a lexical item that is a combined form of verb and noun, adjective and noun, verb and adverb, adverb and adjective (Aarts et al., 2014; Diah Moehkardi, 2012). Lexical collocation is a combination of words that have different word classes and does not contain any grammatical element. Based on the finding, most collocates identified in the learner corpus is lexical collocation. For instance, phrases like 'bring effect',

‘cause conflict’, ‘live harmoniously’, ‘positive effect’, and many others are examples of lexical collocation reported in the present research. As both types of collocations emphasise the word classes used, it is vital to observe the words that co-occur with the homonyms since the accompanied words provided clues to the learners on the grammatical class of the homonym. Once the word class is identified, it is easier for learners to look for semantic restrictions that homonyms impose on different contexts as this restriction could decide the words used together with homonyms based on their senses (Aarts et al., 2014). Thus, this step is significant for learners to ensure that the meaning of the homonym used is appropriate in different contexts.

On the other hand, homonyms are disambiguated based on their collocation register as a collocation not only predicts the structure, word classes, semantics, and the position of a word in a language system but also shows the restriction of compatibility and declination on the words used in a context (Perebiynis & Bobkova, 2008, as cited in Bobkova, 2016). By looking at these linguistic features, learners can foresee the compatibility of the homonyms with other words in different contexts through collocations. Besides, the function of the lexical collocation allowed the researcher to observe any other words that are related to the word class of the homonyms used which could form a predictable connection between the homonyms and co-occurrence words (Aarts et al., 2014). Hence, collocations played a significant role in disambiguating multiple senses of the homonyms as it allows disambiguation of a minimum of 70% of words which could be an effective tool to use when learning homonyms (Bobkova, 2016).

## 5.2 – Recommendation for Future Research

The essays selected for the present study have created a learner corpus that was relatively small because it only involved undergraduates from two different faculties instead of all faculties in the university. In other words, the finding presented in this research cannot

represent how most ESL learners used homonyms since the number of homonyms found is affected by the size of the learner corpus. Therefore, future research should consider the effect of corpus size as it might directly influence the finding. In the previous study, Mizumoto et al. (2012) proved that a large corpus promotes the accuracy and recall of the learners and enhances their performance in making errors, especially for word choice errors that require more lexical knowledge. Since learners are often confused about homonyms, a larger corpus size could be an effective tool to improve the word choice errors related to homonyms.

Although semantic lexical ambiguity is sub-divided into homonym and polysemy, the focus of the current study only falls on the homonym, polysemy is not covered in this research. However, knowing the concept of semantic lexical ambiguity through homonym and polysemy is crucial for learners to distinguish their differences as these two sub-classes sometimes overlap due to their similar characteristics. Tarp (2009) pointed out that even lexicographers are confused with homonyms and polysemy in practical lexicography. This problem or issue would extend to ESL learners since they are in the mid of learning the concept. Consequently, learners might not fully comprehend the semantic lexical ambiguity concept if one of them, homonym or polysemy is not incorporated in the study. Both homonyms and polysemy can be further classified into different categories. If one of them is excluded from the research, learners might not fully aware of the types of homonyms or polysemy. Hence, both sub-categories of semantic lexical ambiguity, homonym and polysemy should be the focus of future research.

Moreover, the current study only emphasised the written context, specifically in analysing the usage pattern of homonyms in the essay writing of undergraduates. As homonyms could appear in both written and spoken contexts, it is also equally important to look at the usage of homonyms in the spoken context. Berdiyeva (2022) believed that the disambiguation of homonyms needs to be carried out in the communication process to determine the exact sense of the homonym that matches the context of speech. Meanwhile, the

researcher of the same study also asserted that the speaker faced some problems in using homonyms in their utterances whereby they randomly chose other unambiguous words to make their statement more comprehensible to people (Berdiyeva, 2022). However, homonyms are unavoidable in the English language as most of them are common English words used by learners in both written and spoken contexts. People are often confused with homonyms due to their ambiguous nature. This phenomenon applies to native speakers and language learners (Readence et al., 1986, as cited in Jacobson et al., 2007). These previous studies have highlighted the reasons for learning homonyms in the spoken context since it would directly affect learners' speaking and reading comprehension skills. While the present study did not highlight the usage pattern of homonyms in the spoken context, it should be conducted for future research.

Besides, undergraduates are the focus group in the current study, the findings cannot represent all Malaysian ESL learners in general since it was not a combined data of the usage pattern of homonyms used by students from primary to tertiary education levels. As the ambiguous nature of the homonyms affects learners of all ages, it acknowledges the importance of learning homonyms in helping learners to distinguish their differences in the lexical, phonetic, and semantic structures. A past study pointed out that primary school students in Malaysia faced problems while recognising the spelling of homophones due to their similarity in pronunciation (Bakar et al., 2018, as cited in Lau & Mohamad, 2020). Similarly, another study mentioned that Malaysian students have difficulties in recognising homophones which has affected their listening skills performance (Singaravelu & Paramasivam, 2016). As the previous studies only highlighted homophones, future research should look at all types of homonyms and their usage patterns used by primary and secondary school students and undergraduates.

### 5.3 – Conclusion

Overall, the present study concluded that there are several significant results which could assist learners to distinguish multiple types of homonyms. There are several ways to resolve the problems encountered by the learners in acquiring homonyms. Based on the homonym categories, learners showed several forms of biases through meaning, word class, and homophone pair (Ushiro et al., 2013). These bias patterns proved that students overlooked the subordinate meanings of the homonyms and other homonyms with a lower frequency of occurrence. This problem should be addressed as learners need more exposure to different senses and the usage of multiple types of homonyms to avoid the issue of misinterpreting homonyms. While looking at the frequency of occurrence, learners can know exactly which homonyms they need to pay extra attention to their meanings and usage which could help them to learn and distinguish different categories of homonyms.

Knowing the common usage of homonyms through their word classes could help in addressing learners' confusion on different types of homonyms. The learners are not only confused about the similar characteristics of the homonyms but also unaware of the meanings of the word classes. However, when the homonyms are presented in the form of a sentence, not in isolation, learners can identify the word class of the homonym through its co-occurrence words. This way not only facilitates the syntactic processing of the texts but also minimizes their confusion in determining the exact sense of the homonym used when they know about the word class since the meanings of each word class vary (Hirschman et al., 1975). Nevertheless, learners should also be aware that identifying the sense of the homonyms through word class can be applied to most homonym categories, except for lexical homonyms.

Some of the Malaysian ESL learners in the current study were confronted with the issue of L1 interference when using homonyms to construct sentences. In this case, learners often directly translate the English homonym from the same word in their mother tongue, and they

assumed that the word sense is the same and fits into the context of usage, but they have misused the homonym with an incorrect meaning (Vethamaiccam & Ganapathy, 2017). Most of the time, students are unaware of the misuse issue and the correct usage of the homonyms. By using corpus, learners can observe different misuse patterns of the homonyms that can help them to rectify the homonym mistakes. Hence, the researcher would like to suggest educators utilise learner corpus as a pedagogical tool in the classroom since scholars have adopted corpus in their studies and highlighted the benefits of using a corpus to teach students about the meaning and usage of words (Levy & Bullinaria, 2001; Poole, 2016). Teachers can introduce the use of the learner corpus to their students by highlighting the misuse pattern of homonyms and their correct usage. After frequent exposure to the misuse pattern, it would eventually increase their awareness of not committing the same mistakes again while using English homonyms.

A corpus not only helps learners to learn misuse patterns of homonyms but also allows learners to predict the word class and meaning of the homonyms used in a context through the collocate tool. The collocations found in the current study are not limited to words that appeared next to the homonyms, other words related to the homonyms are included as part of the findings. This result justified that collocates disambiguate multiple senses of homonyms by forming a predictable connection between different linguistic features shown in the co-occurrence words (Perebiynis & Bobkova, 2008, as cited in Bobkova, 2016). Since a corpus supports the acquisition of homonyms from semantic, grammatical class, and context and even addresses several issues learners would encounter during the learning process, the researcher highly recommends learners and educators use the corpus to learn homonyms. Although homonyms are more challenging to acquire, it should not be a drawback for learners to learn different homonym categories as there is always a solution to tackle the problems encountered during the learning process.



## REFERENCES

- Aarts, B., Chalker, S., & Weiner, E. (2014). *The Oxford dictionary of English grammar* (2nd ed.). Oxford University Press.
- Abuse. (n.d.). *Oxford Languages dictionary*. Retrieved 24 July, 2022, from <https://languages.oup.com/google-dictionary-en/>
- Adha, A. Y., & Widyaningtyas, A. (2017). Students' error in lexical meaning in communication through conversation in the classroom. *Advances in Social Science, Education and Humanities Research*, 104, 442-445. <https://doi.org/10.2991/aisteel-17.2017.95>
- Advice. (n.d.). *Oxford Languages dictionary*. Retrieved 31 July, 2022, from <https://languages.oup.com/google-dictionary-en/>
- Advise. (n.d.). *Oxford Languages dictionary*. Retrieved 31 July, 2022, from <https://languages.oup.com/google-dictionary-en/>
- Affect. (n.d.). *Oxford Languages dictionary*. Retrieved 31 July, 2022, from <https://languages.oup.com/google-dictionary-en/>
- Afzal, N. (2019). A study on vocabulary-learning problems encountered by BA English majors at the university level of education. *Arab World English Journal*, 10(3), 81-98. <http://dx.doi.org/10.2139/ssrn.3465990>
- Al-Harbi, O., Jusoh, S., & Norwawi, N. M. (2011). Lexical disambiguation in natural language questions (NLQs). *IJCSI International Journal of Computer Science*, 8(4), 143-150. [https://www.researchgate.net/publication/221936502\\_Lexical\\_Disambiguation\\_in\\_Natural\\_Language\\_Questions\\_NLQs](https://www.researchgate.net/publication/221936502_Lexical_Disambiguation_in_Natural_Language_Questions_NLQs)
- Alekse, L. (2017). *Distinguishing between polysemy and homonymy: A critique of a common dictionary approach* [Bachelor's thesis, University of Iceland]. Skemman. <http://hdl.handle.net/1946/26712>

- Alghamdi, L. H. M. (2021). The challenges of homographs among EFL Learners of the college of foreign languages. *Journal of English Language Teaching and Applied Linguistics*, 3(12), 1-8. <https://doi.org/10.32996/jeltal.2021.3.12.1>
- Anthony, L. (2020). *AntConc (Windows, Macintosh OS X, and Linux)*. Laurence Anthony's Website. <https://www.laurenceanthony.net/software/antconc/>
- Azmi, N. A., Tan, K. H., Abdullah, I. H., & Azman, H. (2021). The efficacy of Kibbitzer in solving specific language problems among ESL undergraduates. *International Journal of English Language and Literature Studies*, 10(2), 83-93. [https://www.researchgate.net/publication/349285817\\_The\\_Efficacy\\_of\\_Kibbitzer\\_in\\_Solving\\_Specific\\_Language\\_Problems\\_among\\_ESL\\_Undergraduates](https://www.researchgate.net/publication/349285817_The_Efficacy_of_Kibbitzer_in_Solving_Specific_Language_Problems_among_ESL_Undergraduates)
- Balloqui, M. R. (2017). *Polysemy*. Digital Archive Learning Researching. <http://hdl.handle.net/10810/21400>
- Berdiyeva, M. (2022). The usage of homonyms in English and Uzbek languages. *Science and Education*, 3(7), 461-464. <https://openscience.uz/index.php/sciedu/article/view/4105>
- Biber, D., Conrad, S., & Reppen, R. (1994). Corpus-based approaches to issues in applied linguistics. *Applied Linguistics*, 15(2), 169–189. <https://doi.org/10.1093/applin/15.2.169>
- Blum, S., & Levenston, E. A. (1978). Universals of lexical simplification. *Language Learning*, 28, 399-415. <https://doi.org/10.1111/j.1467-1770.1978.tb00143.x>
- Bobkova, A. (2016, April 26-27). *The use of collocations for English nouns disambiguation* [conference session]. 4<sup>th</sup> scientific conference “Thought Elaboration: Linguistics, Literature, Media Expression: Tell Me’ 16, Kaunas, Lithuania. [https://www.knf.vu.lt/dokumentai/failai/katedru/germanu/TELL\\_ME\\_2017.pdf#page=64](https://www.knf.vu.lt/dokumentai/failai/katedru/germanu/TELL_ME_2017.pdf#page=64)
- Bolshakov, I., & Gelbukh, A. (2001). A very large database of collocations and semantic links. In M. Bouzeghoub, Z. Kedad, E. Métais (Eds.), *Natural language processing and information systems: Lecture notes in computer Science 1959* (pp. 103–114), Springer. [https://doi.org/10.1007/3-540-45399-7\\_9](https://doi.org/10.1007/3-540-45399-7_9)

- Born. (n.d.). Oxford Languages dictionary. Retrieved 24 July, 2022, from <https://languages.oup.com/google-dictionary-en/>
- Budiwiyanto, A. & Suhardijanto, T. (2019). Frequency and structure of Indonesian lexical bundles on academic prose in legal studies: A driven-corpus approach. In P. A. W. Wibowo, S. Endraswara, T. Supriyanto, K. Saddhono, & Pardi (Eds.), *BASA 2019: Proceedings of the third international seminar on recent language* (pp. 101-107). European Alliance for Innovation. [https://books.google.com.my/books/about/BASA\\_2019.html?id=bMv2DwAAQBAJ&redir\\_esc=y](https://books.google.com.my/books/about/BASA_2019.html?id=bMv2DwAAQBAJ&redir_esc=y)
- Bybee, J. L. & Beckner, C. (2015). Usage-based theory. In B. Heine & H. Narrog (Eds.), *The Oxford handbook of linguistic analysis* (2nd ed., pp. 953-979). Oxford University Press. <https://doi.org/10.1093/oxfordhb/9780199677078.013.0032>
- Case. (n.d.). Oxford Languages dictionary. Retrieved 24 July, 2022, from <https://languages.oup.com/google-dictionary-en/>
- Conflict. (n.d.). *Oxford Languages dictionary*. Retrieved 24 July, 2022, from <https://languages.oup.com/google-dictionary-en/>
- Content. (n.d.). Oxford Languages dictionary. Retrieved 24 July, 2022, from <https://languages.oup.com/google-dictionary-en/>
- Contrast. (n.d.). Oxford Languages dictionary. Retrieved 24 July, 2022, from <https://languages.oup.com/google-dictionary-en/>
- Cover. (n.d.). Oxford Languages dictionary. Retrieved 24 July, 2022, from <https://languages.oup.com/google-dictionary-en/>
- Crosthwaite, P., Wong, L., & Cheung, J. (2019). Characterising postgraduate students' corpus query and usage patterns for disciplinary data-driven learning. *ReCALL*, 31(3), 255-275. <https://doi.org/10.1017/S0958344019000077>
- Cruse, A. (2010). *Notes on meaning in language*. PhilArchive. <https://philarchive.org/archive/CRUMIL-2>

- Daniel, J. (2012). *Sampling essentials: Practical guidelines for making sampling choices*. SAGE Publications. <https://www.doi.org/10.4135/9781452272047>
- Daskalovska, N. (2015). Corpus-based versus traditional learning of collocations. *Computer Assisted Language Learning*, 28(2), 130-144. <https://doi.org/10.1080/09588221.2013.803982>
- Demir, C. (2020). Lexical and structural ambiguities in student writing: An assessment and evaluation of results. *African Educational Research Journal*, 8(3), 100-108. [https://www.researchgate.net/publication/345940501\\_Lexical\\_and\\_structural\\_ambiguities\\_in\\_student\\_writing\\_An\\_assessment\\_and\\_evaluation\\_of\\_results](https://www.researchgate.net/publication/345940501_Lexical_and_structural_ambiguities_in_student_writing_An_assessment_and_evaluation_of_results)
- Deng, K. (2011). Technical university students' difficulties in understanding polysemes and homonyms. *Chaoyang Journal of Humanities and Social Sciences*, 9(1), 301-331. <http://ir.lib.cyut.edu.tw:8080/handle/310901800/12208>
- Diah Moehkardi, R. (2012). Grammatical and lexical English collocations: Some possible problems to Indonesian learners Of English. *Humaniora*, 14(1), 53-62. <https://doi.org/10.22146/jh.745>
- Dipolog-Ubanan, G. F. (2016). L1 Influence on writing in L2 among UCSI Chinese students: A case study. *Pertanika Journal of Social Science & Humanities*, 24(4), 1835 – 1847. [https://www.researchgate.net/publication/313314031\\_L1\\_influence\\_on\\_writing\\_in\\_L2\\_among\\_UCSI\\_Chinese\\_students\\_A\\_case\\_study](https://www.researchgate.net/publication/313314031_L1_influence_on_writing_in_L2_among_UCSI_Chinese_students_A_case_study)
- Duffy, S. A., Kambe, G., & Rayner, K. (2001). The effect of prior disambiguating context on the comprehension of ambiguous words: Evidence from eye movements. In D. S. Gorfein (Ed.), *On the consequences of meaning selection: Perspectives on resolving lexical ambiguity* (1st ed., pp. 27-43). American Psychological Association.
- Eddington, C. M., & Tokowicz, N. (2015). How meaning similarity influences ambiguous word processing: The current state of the literature. *Psychonomic Bulletin & Review*, 22(1), 13–37. <https://doi.org/10.3758/s13423-014-0665-7>
- Effect. (n.d.). Oxford Languages dictionary. Retrieved 31 July, 2022, from <https://languages.oup.com/google-dictionary-en/>

- Falkum, I. L. (2011). *The semantics and pragmatics of polysemy: A relevance-theoretic account* [Doctoral dissertation, University College London]. UCL Discovery. <https://discovery.ucl.ac.uk/id/eprint/1139079>
- Gablasova, D., Brezina, V., & McEnery, T. (2017). Collocations in corpus-based language learning research: Identifying, comparing, and interpreting the evidence. *Language Learning*, 67(S1), 155-179. <https://doi.org/10.1111/lang.12225>
- Gautam, P. (2019). Writing skill: An instructional overview. *Journal of NELTA Gandaki*, 2, 74-90. <https://doi.org/10.3126/jong.v2i0.26605>
- Get. (n.d.). Oxford Languages dictionary. Retrieved 24 July, 2022, from <https://languages.oup.com/google-dictionary-en/>
- Hirschman, L., Grishman, R., & Sager, N. (1975). Grammatically-based automatic word class formation. *Information Processing and Management*, 11(1-2), 39-57. [https://doi.org/10.1016/0306-4573\(75\)90033-3](https://doi.org/10.1016/0306-4573(75)90033-3)
- Hussein, K. S. (2015). The potentialities of corpus-based techniques for analyzing literature. *International Journal of English Language Linguistics Research*, 2(2), 74-86. [https://www.researchgate.net/publication/315800356\\_The\\_potentialities\\_of\\_corpus-based\\_techniques\\_for\\_analyzing\\_literature](https://www.researchgate.net/publication/315800356_The_potentialities_of_corpus-based_techniques_for_analyzing_literature)
- Ibrahim, A. M. A. (2018). *Investigating the problems that result of using homophones and homographs among students of the college of languages* [Master's thesis, Sudan University of Science and Technology]. SUST Repository. <http://repository.sustech.edu/handle/123456789/21143>
- Impact. (n.d.). Oxford Languages dictionary. Retrieved 24 July, 2022, from <https://languages.oup.com/google-dictionary-en/>
- Increase. (n.d.). Oxford Languages dictionary. Retrieved 24 July, 2022, from <https://languages.oup.com/google-dictionary-en/>
- Jackson, P. (2020). Understanding understanding and ambiguity in natural language. *Procedia Computer Science*, 169, 209-225. <https://doi.org/10.1016/j.procs.2020.02.138>

- Jacobson, J., Lapp, D., & Flood, J. (2007). A seven-step instructional plan for teaching English-language learners to comprehend and use homonyms, homophones, and homographs. *Journal of Adolescent & Adult Literacy*, 51(2), 98-111.  
<https://doi.org/10.1598/JAAL.51.2.2>
- Katz, J. J., & Fodor, J. A. (1963). The structure of a semantic theory. *Language*, 39(2), 170–210. <https://doi.org/10.2307/411200>
- Khisamova, V., Abdullina, L., Nurgalieva, L., & Khabibullina, E. (2020). Classification of homonymic terms in medical terminology of English, Russian and Tatar languages. *Journal of Educational and Social Research*, 10(6), 18-34.  
[https://pdfs.semanticscholar.org/1b61/e77dad2954777664ce3f40952c049d771741.pdf?\\_ga=2.69199058.258641491.1629711460-401166654.1613185095](https://pdfs.semanticscholar.org/1b61/e77dad2954777664ce3f40952c049d771741.pdf?_ga=2.69199058.258641491.1629711460-401166654.1613185095)
- Klepousniotou, E. (2002). The processing of lexical ambiguity: Homonymy and polysemy in the mental lexicon. *Brain and language*, 81(1-3), 205–223.  
<https://doi.org/10.1006/brln.2001.2518>
- Kreidler, C. W. (1998). *Introducing English semantics*. Routledge.
- Kristian, Y. A. K. (2018). *Lexical and structural ambiguity found in The Jakarta Post articles on the E-Id graft case* [Bachelor's thesis, Sanata Dharma University]. Repository Universitas Sanata Dharma. [https://repository.usd.ac.id/31560/2/144214103\\_full.pdf](https://repository.usd.ac.id/31560/2/144214103_full.pdf)
- Kulkarni, A., Heilman, M., Eskenazi, M., & Callan, J. (2008). Word sense disambiguation for vocabulary learning. In B. P. Woolf, E. Aïmeur, R. Nkambou & S. Lajoie (Eds.), *ITS 2008: Intelligent tutoring systems* (pp. 500-509). Springer.  
[https://doi.org/10.1007/978-3-540-69132-7\\_53](https://doi.org/10.1007/978-3-540-69132-7_53)
- Lau, Y. Y., & Mohamad, M. (2020). Utilising e-learning to assist primary school ESL pupils in learning to spell during COVID-19 pandemic: A literature review. *Creative Education*, 11(8), <https://doi.org/10.4236/ce.2020.118091>
- Laufer, B. (1990). Why are some words more difficult than others? – Some intralexical factors that affect the learning of words. *IRAL – International Review of Applied Linguistics in Language Teaching*, 28(4), 293-308.  
<https://doi.org/10.1515/iral.1990.28.4.293>

- Lead. (n.d.). Oxford Languages dictionary. Retrieved 31 July, 2022, from <https://languages.oup.com/google-dictionary-en/>
- Levy, J. P., & Bullinaria, J.A. (2001). Learning lexical properties from word usage patterns: Which context words should be used?. In R. M. French & J. P. Sougné (Eds.), *Connectionist models of learning, development and evolution: Perspectives in neural computing* (pp.273-282). Springer. [https://doi.org/10.1007/978-1-4471-0281-6\\_27](https://doi.org/10.1007/978-1-4471-0281-6_27)
- Light. (n.d.). Oxford Languages dictionary. Retrieved 24 July, 2022, from <https://languages.oup.com/google-dictionary-en/>
- Liu. C. (2012). The linguistic analysis of sound ambiguity and meaning ambiguity. *International Journal of Education and Management Engineering*, 2(10), 39-44. <https://doi.org/10.5815/ijeme.2012.10.07>
- Live. (n.d.). Oxford Languages dictionary. Retrieved 24 July, 2022, from <https://languages.oup.com/google-dictionary-en/>
- Loebner, S., & Hooimeijer, F. (2013). *Understanding semantics* (2nd ed.). Taylor & Francis Group.
- Low. (n.d.). Oxford Languages dictionary. Retrieved 24 July, 2022, from <https://languages.oup.com/google-dictionary-en/>
- Mahadzir, N. H., Omar, M. F., & Mohd Nawi, M. N. (2018). Semantic similarity measures for Malay-English ambiguous words. *Journal of Telecommunication, Electronic and Computer Engineering*, 10(1-11), 109–112. <https://jtec.utem.edu.my/jtec/article/view/3860>
- Mamedova, M. (2019). Classification of homonyms of the English language. *European Journal of Research and Reflection in Educational Series*, 7(12), 1-5. <http://www.idpublications.org/wp-content/uploads/2020/01/Full-Paper-CLASSIFICATION-OF-HOMONYMS-OF-THE-ENGLISH-LANGUAGE.pdf>
- McEnery, T., Xiao, R., & Tono, Y. (2006). *Corpus-based language studies: An advanced resource book*. Routledge.

- Mizumoto, T., Hayashibe, Y., Komachi, M., Nagata, M., & Matsumoto, Y. (2012, December 8-15). *The effect of learner corpus size in grammatical error correction of ESL writings* [conference session]. COLING 2012, 24<sup>th</sup> International Conference on Computational Linguistics, Mumbai, India. <https://aclanthology.org/C12-2084>
- Mohd Nasir, N. F. W. (2021). Nativised English lexemes and semantic shift in Malaysian English. *International Journal of Modern Languages and Applied Linguistics (IJMAL)*, 5(3), 77-94. <https://ir.uitm.edu.my/id/eprint/6110/1/6110.pdf>
- Mukoroli, J. (2011). *Effective vocabulary teaching strategies for the English for academic purposes ESL classroom* [Master's thesis, SIT Graduate Institute]. SIT Digital Collections. [https://digitalcollections.sit.edu/ipp\\_collection/501](https://digitalcollections.sit.edu/ipp_collection/501)
- Nagy, W. E. (1995). *On the role of context in first-and second-language vocabulary learning* (Technical Report No. 627). College of Education, University of Illinois at Urbana-Campaign. <https://eric.ed.gov/?id=ED391152>
- Need. (n.d.). Oxford Languages dictionary. Retrieved 31 July, 2022, from <https://languages.oup.com/google-dictionary-en/>
- Ngula, R. S. (2018). Corpus methods in language studies. In D. D. Kuupole (Ed.), *Perspectives on conducting and reporting research in the humanities* (pp. 205-223), University of Cape Coast Press. [https://www.researchgate.net/publication/327234896\\_CORPUS\\_METHODS\\_IN\\_LANGUAGE\\_STUDIES](https://www.researchgate.net/publication/327234896_CORPUS_METHODS_IN_LANGUAGE_STUDIES)
- Noack, R., & Gamio, L. (2015, April 23). *The world's languages, in 7 maps and charts*. The Washington Post. <https://www.washingtonpost.com/news/worldviews/wp/2015/04/23/the-worlds-languages-in-7-maps-and-charts/>
- Nouraldeen, A. S. (2015). Meaning and context-three different perspectives. *British Journal of English Linguistics*, 3(2), 13-17. <https://www.eajournals.org/wp-content/uploads/Meaning-and-Context-Three-Different-Perspectives..pdf>
- O'Keeffe, A., McCarthy, M., & Carter, R. (2007). *From corpus to classroom: Language use and language teaching*. Cambridge University Press. <https://doi.org/10.1017/CBO9780511497650>



- Ovu, B. F. (2011). Lexical sources of ambiguity in English and daily communication. *Oluoha Journal of Languages*, 1(1), 15-26.  
[https://www.researchgate.net/publication/309033841\\_Lexical\\_Sources\\_of\\_Ambiguity\\_in\\_English\\_and\\_Daily\\_Communication](https://www.researchgate.net/publication/309033841_Lexical_Sources_of_Ambiguity_in_English_and_Daily_Communication)
- Owu-Ewie, C., & Williams, R. (2017). Grammatical and lexical errors in students' English composition writing: The case of three senior high schools (SHS) in the central region of Ghana. *Sino-US English Teaching*, 14(8), 463-482. <https://doi.org/10.17265/1539-8072/2017.08.001>
- Parent, K. (2012). The most frequent English homonyms. *RELC Journal*, 43(1), 69–81.  
<https://doi.org/10.1177/0033688212439356>
- Pennebaker, J. W., Chung, C. K., Frazee, J., Lavergne, G. M., & Beaver, D. I. (2014). When small words foretell academic success: The case of college admissions essays. *PloS one*, 9(12), 1-10. <https://doi.org/10.1371/journal.pone.0115844>
- Peters, E. (2020). Factors affecting the learning of single-word items 1. In S. Webb (Ed.), *The Routledge handbook of vocabulary studies* (1st ed., pp. 125-142). Routledge.  
<https://doi.org/10.4324/9780429291586>
- Picture. (n.d.). Oxford Languages dictionary. Retrieved 24 July, 2022, from <https://languages.oup.com/google-dictionary-en/>
- Poole, R. (2016). A corpus-aided approach for the teaching and learning of rhetoric in an undergraduate composition course for L2 writers. *Journal of English for Academic Purposes*, 21, 99-109. <https://doi.org/10.1016/j.jeap.2015.12.003>
- Present. (n.d.). Oxford Languages dictionary. Retrieved 24 July, 2022, from <https://languages.oup.com/google-dictionary-en/>
- Produce. (n.d.). Oxford Languages dictionary. Retrieved 24 July, 2022, from <https://languages.oup.com/google-dictionary-en/>
- Protest. (n.d.). Oxford Languages dictionary. Retrieved 24 July, 2022, from <https://languages.oup.com/google-dictionary-en/>

- Race. (n.d.). Oxford Languages dictionary. Retrieved 24 July, 2022, from <https://languages.oup.com/google-dictionary-en/>
- Rodd, J. (2017). *Lexical ambiguity*. PsyArXiv Preprints. <https://doi.org/10.31234/osf.io/yez6>
- Rohmatillah, R. (2014). A study on students' difficulties in learning vocabulary. *Jurnal Tadris Bahasa Inggris*, 6(1), 69-86. <http://ejournal.radenintan.ac.id/index.php/ENGEDU/article/view/520>
- Safataj, M., & Amiryousefi, M. (2016). Effect of homonymous set of words instruction on vocabulary development and retention of young female elementary learners in Iranian EFL context through metalinguistic awareness. *Theory and Practice in Language Studies*, 6(11), 2092-2101. <http://dx.doi.org/10.17507/tpls.0611.05>
- Saleh, Y. M. (2017). Semantic ambiguity in English language. *Journal of Surra Man Raa*, 13(48), 607-618. <https://www.iasj.net/iasj/article/170488>
- Save. (n.d.). Oxford Languages dictionary. Retrieved 24 July, 2022, from <https://languages.oup.com/google-dictionary-en/>
- Sense. (n.d.). Oxford Languages dictionary. Retrieved 24 July, 2022, from <https://languages.oup.com/google-dictionary-en/>
- Setiani, R. (2019). Students' ambiguity in writing an essay. *Edukasi Lingua Sastra*, 17(1), 29-39. <https://doi.org/10.47637/elsa.v17i1.104>
- Shamsudin, S., Husin, N., & Manan, A. A. (2013). Exploring fundamental engineering word list for engineering students: A literature review. *Procedia - Social and Behavioral Sciences*, 70, 1275-1281. <https://doi.org/10.1016/j.sbspro.2013.01.187>
- Simpson, G. B. (1984). Lexical ambiguity and its role in models of word recognition. *Psychological bulletin*, 96(2), 316-340. <https://doi.org/10.1037/0033-2909.96.2.316>

- Simpson, G. B., & Burgess, C. (1985). Activation and selection processes in the recognition of ambiguous words. *Journal of Experimental Psychology: Human Perception and Performance*, 11(1), 28-39. <https://doi.org/10.1037/0096-1523.11.1.28>
- Singaravelu, G., & Paramasivam, M. (2016). Insight of teachers on problems of students in acquiring communicative skill in English. *Journal on English Language Teaching*, 6(2), 9-17. <https://files.eric.ed.gov/fulltext/EJ1131406.pdf>
- Small, S. L., Cottrell, G. W., Tanenhaus, M. K. (1988). Preface. In S. L. Small, G. W. Cottrell & M. K. Tanenhaus (Eds.), *Lexical ambiguity resolution: Perspective from psycholinguistics, neuropsychology and artificial intelligence* (pp. 3-9). Morgan Kaufmann. <https://doi.org/10.1016/C2009-0-27555-6>
- Smith, K. (2004). The evolution of vocabulary. *Journal of Theoretical Biology*, 228(1), 127–142. <https://doi.org/10.1016/j.jtbi.2003.12.016>
- Starr, M. S., & Fleming, K. K. (2001). A rose by any other name is not the same: The role of orthographic knowledge in homophone confusion errors. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 27(3), 744–760. <https://doi.org/10.1037//0278-7393.27.3.744>
- Storjohann, P. (2005, July 14-17). *Corpus-driven vs. corpus-based approach to the study of relational patterns* [Conference session]. Corpus Linguistics 2005 conference, Birmingham, West Midlands, United Kingdom. <https://www.birmingham.ac.uk/research/activity/corpus/publications/conference-archives/2005-conf-e-journal.aspx>
- Tabossi, P., Colombo, L. & Job, R. (1987). Accessing lexical ambiguity: Effects of context and dominance. *Psychological Research*, 49, 161–167. <https://doi.org/10.1007/BF00308682>
- Take. (n.d.). Oxford Languages dictionary. Retrieved 24 July, 2022, from <https://languages.oup.com/google-dictionary-en/>
- Tan, W. L. (2018). *Improving the academic vocabulary of tertiary learners through direct vocabulary instruction with gamification* [Doctoral dissertation, Wawasan Open University]. Wawasan Open University Library. [https://woulibrary.wou.edu.my/theses-project/PhD2018\\_WLTAN.pdf](https://woulibrary.wou.edu.my/theses-project/PhD2018_WLTAN.pdf)

- Tarp, S. (2009). Homonymy and polysemy in a lexicographical perspective. *Zeitschrift für Anglistik und Amerikanistik*, 57(3), 289-305.  
<https://doi.org/10.1515/zaa.2009.57.3.289>
- Use. (n.d.). Oxford Languages dictionary. Retrieved 24 July, 2022, from  
<https://languages.oup.com/google-dictionary-en/>
- Ushiro, Y., Hasegawa, Y., Nahatame, S., Hamada, A., Kimura, Y., Mori, Y., & Kato, D. (2013). Incremental learning of homonyms in multiple contexts among Japanese EFL readers. *Jacet Journal*, (57), 1-19. <https://ci.nii.ac.jp/naid/110009684903>
- Ushiro, Y., Hoshino, Y., Shimizu, H., Kai, A., Nakagawa, C., Watanabe, F., & Takaki, S. (2010). Disambiguation of homonyms by EFL readers: The effects of primary meaning and context interpretation. *Annual Review of English Language Education in Japan*, 21, 161-170. [https://doi.org/10.20581/arele.21.0\\_161](https://doi.org/10.20581/arele.21.0_161)
- Vethamaiccam, M., & Ganapathy, M. (2017). Analysing errors among form one students' written English composition in a private Chinese school. *Asian Journal of Education and e-Learning*, 5(1), 16-26.  
[https://www.researchgate.net/publication/313763365\\_Analysing\\_Errors\\_among\\_Form\\_One\\_Students'\\_Written\\_English\\_Composition\\_in\\_a\\_Private\\_Chinese\\_School](https://www.researchgate.net/publication/313763365_Analysing_Errors_among_Form_One_Students'_Written_English_Composition_in_a_Private_Chinese_School)
- Yarowsky, D. (1993, March 21-24). *One sense per collocation* [Conference session]. HLT '93: Proceedings of the workshop on Human Language Technology, Princeton, New Jersey, USA. <https://doi.org/10.3115/1075671.1075731>
- Yu, T. (2013). *College freshmen's strategies to solve ambiguities in process analysis writing* [The dialogue between English culture, teaching and learning: Selected papers of the 2013 academic seminar of the department of English]. National Pingtung University.  
<http://140.127.82.166/handle/987654321/15299>
- Yunus, K. & Awab, S. (2012). The effects of the use of module-based concordance materials and data-driven learning (DDL) approach in enhancing the knowledge of collocations of prepositions among Malaysian undergraduate Law students. *The International Journal of Learning*, 18(9), 181-198.  
[https://www.researchgate.net/publication/286075606\\_The\\_Effects\\_of\\_the\\_Use\\_of\\_Module-based\\_Concordance\\_Materials\\_and\\_Data-driven\\_Learning\\_DDL\\_Approach\\_in\\_Enhancing\\_the\\_Knowledge\\_of\\_Collocations\\_of\\_Prepositions\\_among\\_Malaysian\\_Undergraduate\\_Law\\_Students](https://www.researchgate.net/publication/286075606_The_Effects_of_the_Use_of_Module-based_Concordance_Materials_and_Data-driven_Learning_DDL_Approach_in_Enhancing_the_Knowledge_of_Collocations_of_Prepositions_among_Malaysian_Undergraduate_Law_Students)

Yurchenko, A., Lopukhina, A. & Dragoy, O. (2020). Metaphor is between metonymy and homonymy: Evidence from event-related potentials. *Frontiers in Psychology, 11*, 1-12. <https://doi.org/10.3389/fpsyg.2020.02113>

## Appendix A

### Consent Form

You are invited to participate in a research study about how ESL learners used homonyms in constructing sentences while writing essays. The study aims to identify the types of homonyms and report the usage pattern of homonyms in the writings of undergraduate students. As I will generate a learner corpus, the study will collect students' mid-term assessment from UALL1083/2023 Academic Writing subject in May 2021 and January 2022 trimesters as the main data for this research. The types of essays will be collected is argumentative essay. All personal information and other details that are related to an individual's assignment will be kept anonymously.

Please understand that your participation is voluntary, and you have the right to withdraw your consent or discontinue your participation at any time without penalty. You have the right not to answer any specific questions. The private information found in the essays collected will not be included as part of the research, even if identifiers are removed, it will not be used or distributed for future research studies. The individual privacy and confidentiality of the information you provide will be maintained and kept in all published and written data resulting from the research.

If you have any questions about this study, or require any information about the research, you can contact me via email, [summer.law@lutar.my](mailto:summer.law@lutar.my) or contact my supervisor, Ms Deepa a/p Visvanathan via email, [deepav@utar.edu.my](mailto:deepav@utar.edu.my).

I would appreciate it if you could sign this form and return it to me via email. Many thanks in advance for your consideration of this research. Please let me know if you require further information.

I understand that my participation in this project will involve in:

- Assisting Ms Law Hui Qin to select students' mid-term assessment under UALL1083/2023 Academic Writing subject from the Faculty of Arts and Social Science and the Faculty of Business and Finance in May 2021 and January 2022 trimesters.

I understand that the identity of all students whose essays have been selected for the study are kept confidential.

I understand that the identity of students will be treated confidentially by Ms Law Hui Qin and that all information will be stored anonymously and securely. All information appearing in the final report will be anonymous.

I also understand that I am free to discuss any questions I might have with Ms Deepa a/p Visvanathan.

I, Mohd Amir Izuddin bin Mohamad Ghazali, consent to Ms Law Hui Qin proceeding with this study under the supervision of Ms Deepa a/p Visvanathan.

Signature of Lecturer: .....



Date: 6.7.2022

## Appendix B

### Consent Form

You are invited to participate in a research study about how ESL learners used homonyms in constructing sentences while writing essays. The study aims to identify the types of homonyms and report the usage pattern of homonyms in the writings of undergraduate students. As I will generate a learner corpus, the study will collect students' mid-term assessment from UALL1083/2023 Academic Writing subject in May 2021 trimester as the main data for this research. The essays collected could be from these genres – causal analysis essay, compare and contrast essay, or argumentative essay. All personal information and other details that are related to an individual's assignment will be kept anonymously.

Please understand that your participation is voluntary, and you have the right to withdraw your consent or discontinue your participation at any time without penalty. You have the right not to answer any specific questions. The private information found in the essays collected will not be included as part of the research, even if identifiers are removed, it will not be used or distributed for future research studies. The individual privacy and confidentiality of the information you provide will be maintained and kept in all published and written data resulting from the research.

If you have any questions about this study, or require any information about the research, you can contact me via email, [summer.law@lutar.my](mailto:summer.law@lutar.my) or contact my supervisor, Ms Deepa a/p Visvanathan via email, [deepav@utar.edu.my](mailto:deepav@utar.edu.my).

I would appreciate it if you could sign this form and return it to me via email. Many thanks in advance for your consideration of this research. Please let me know if you require further information.



I understand that my participation in this project will involve in:

- Assisting Ms Law Hui Qin to select students' mid-term assessment under UALL1083/2023 Academic Writing subject from the Faculty of Business and Finance in May 2021 trimester.

I understand that the identity of all students whose essays have been selected for the study are kept confidential.

I understand that the identity of students will be treated confidentially by Ms Law Hui Qin and that all information will be stored anonymously and securely. All information appearing in the final report will be anonymous.

I also understand that I am free to discuss any questions I might have with Ms Deepa a/p Visvanathan.

I, \_\_\_\_\_TAN SWEE MEE\_\_\_\_\_ (NAME) consent to Ms Law Hui Qin proceeding with this study under the supervision of Ms Deepa a/p Visvanathan.

Signature of Lecturer: .....  .....

Date: ...18-07-2022.....