A STUDY OF LEXICAL RICHNESS IN ESL STUDENTS' WRITING

JOSEPHINE UMA A/P SELVARAJU

FACULTY OF LANGUAGES AND LINGUISTICS
UNIVERSITY OF MALAYA
KUALA LUMPUR

2019

University of Malaya

A STUDY OF LEXICAL RICHNESS IN ESL STUDENTS' WRITING

JOSEPHINE UMA A/P SELVARAJU

DISSERTATION SUBMITTED IN PARTIAL FULFILMENT OF THE REQUIREMENTS FOR THE DEGREE OF MASTER OF ENGLISH AS A SECOND LANGUAGE

FACULTY OF LANGUAGES & LINGUISTICS
UNIVERSITY OF MALAYA
KUALA LUMPUR

2019

University of Malaya

UNIVERSITY OF MALAYA

ORIGINAL LITERARY WORK DECLARATION

Name of Candidate: Josephine uma d/o Selvaraju

Matric No: TGB140016

Name of Degree: Master of English as a Second Language

Title of ~~Project Paper/Research Report~~/Dissertation/~~Thesis~~ ("this Work"):

A Study of Lexical Richness in ESL Students' Writing

Field of Study: Language Acquisition

I do solemnly and sincerely declare that:

(1) I am the sole author/writer of this Work;
(2) This Work is original;
(3) Any use of any work in which copyright exists was done by way of fair dealing and for permitted purposes and any excerpt or extract from, or reference to or reproduction of any copyright work has been disclosed expressly and sufficiently and the title of the Work and its authorship have been acknowledged in this Work;
(4) I do not have any actual knowledge nor do I ought reasonably to know that the making of this work constitutes an infringement of any copyright work;
(5) I hereby assign all and every rights in the copyright to this Work to the University of Malaya ("~~U~~UM"), who henceforth shall be owner of the copyright in this Work and that any reproduction or use in any form or by any means whatsoever is prohibited without the written consent of UM having been first had and obtained;
(6) I am fully aware that if in the course of making this Work I have infringed any copyright whether intentionally or otherwise, I may be subject to legal action or any other action as may be determined by UM.

Candidate's Signature Date:

Subscribed and solemnly declared before,

Witness's Signature Date:

Name:

Designation:

University of Malaya

A STUDY OF LEXICAL RICHNESS IN ESL STUDENTS' WRITING

ABSTRACT

The significance of the English language carries endless practicability which is known as commonly spoken language in major areas in Malaysia. The relationship between English language vocabulary and writing considered as notion in second language learning. This study aims to look at the lexical richness of Pre University students in their written context to ascertain if lexical richness has an impact on the quality of written essays. This study aims to investigate lexical richness produced by Foundation in Science students in two different semesters such as the first semester and final semester. Data gathered from two groups with a total of 30 numbers of participants and measured using Compleat Lexical Tutor through Vocabprofile by (Cobb, 2009). The lexical investigation is looking at 3 lexical richness categories; lexical variation, lexical density, and lexical sophistication. Lexical variation with the highest TTR value showed the ability to write the essay precisely without repetition of words that used more than once. Lexical density, three students have produced more than 50 % which shows the text is moderately high and has a good sense of writing and expressed ideas cohesive in a text. Lexical sophistication indicates that a higher percentage of high-frequency words showed a lower dimension of performance. Though, there was no much critical diverse between lexical variation, lexical density, and lexical sophistication among the first and final semesters.

University of Malaya

KAJIAN KEKAYAAN LEKSIKAL DALAM PENULISAN ESL PELAJAR

ABSTRAK

Kepentingan penggunaan bahasa Inggeris membawa kebolehlaksanaan yang tidak berkesudahan yang dikenali sebagai bahasa yang biasa digunakan di kawasan utama di Malaysia. Hubungan antara perbendaharaan kata dan tulisan bahasa Inggeris yang dikenali sebagai tanggapan yang mantap dalam pembelajaran bahasa kedua. kajian ini bertujuan untuk mengkaji kekayaan leksikal pelajar Pra Universiti dalam konteks bertulis untuk menentukan sama ada kekayaan leksikal mempunyai kesan ke atas kualiti karangan bertulis. Kajian ini bertujuan untuk mengkaji kekayaan leksikal yang dihasilkan oleh pelajar Asasi Sains dalam dua semester berbeza seperti semester pertama dan semester akhir. Data dikumpulkan dari dua kumpulan dengan sejumlah 30 peserta dan data diukur dengan menggunakan *Compleat Lexical Tutor* melalui Vocabprofile penyesuaian (Cobb, 2009). Kajian leksikal melihat 3 kategori kekayaan leksikal; iaitu variasi leksikal, kepadatan leksikal dan kecanggihan leksikal. Variasi leksikal dengan nilai TTR tertinggi, menunjukkan kebolehan menulis esei dengan tepat tanpa pengulangan kata-kata yang digunakan lebih daripada satu kali. Kepadatan Leksikal, tiga pelajar telah menghasilkan lebih daripada 50% yang menunjukkan teks itu agak tinggi dan boleh menulis dengan baik atau menyatakan idea secara kohesif dalam teks. Kecanggihan leksikal menunjukkan bahawa peratusan tinggi perkataan frekuensi tinggi mencerminkan dimensi yang lebih rendah prestasi. Manakala, tidak terdapat banyak perbezaan yang kritikal antara variasi leksikal, ketumpatan leksikal antara semester pertama dan akhir.

University of Malaya

ACKNOWLEDGEMENTS

First and foremost, I would like to express my gratitude to God who gave me wisdom and understanding with knowledge and all kinds of skills to fulfill this Masters' Research. My love and heartfelt thanks to my parents for their support, love, prayers, caring, and sacrifices for educating and preparing me for my future. I express my thanks to my sisters (Mal & Sobby) for their unending motivation, support and valuable prayers. My family is always ready to help me whenever I needed them. I could not have accomplished anything without them. Also, I am thankful for Pastor John Titus and Pastor Simon for their prayers and spiritual support.

I would like to express my gratitude to my beloved supervisor Dr. Roshidah Hassan for her unending guidance throughout this research. It is with her supervision that this work came into existence. I am also so thankful for my research panels Dr. Ng Lee Kuan and Dr. Chau Meng Huat whose challenges and provide productive critics and suggestions during Research Proposal and Candidature Defence seminars that have provided new ideas to the work.

I‘m very pleased that I finally managed to complete my research during this most difficult time of my life. And in fact, after the busy hours of work, I really enjoyed the time alone working on my Research. I learned a lot and improve my knowledge throughout this Masters education. I am glad that finally, I completed my Dissertation. I dedicate this research to all second language learners.

University of Malaya

TABLE OF CONTENTS

Abstract iv

Abstrak v

Acknowledgements vi

Table of Contents vii

List of Figures x

List of Tables xi

List of Symbols and Abbreviations xii

List of Appendices xiii

CHAPTER 1: INTRODUCTION 1

1.1 Introduction 1

1.2 Education system in Malaysia 1

1.3 Background of the study 3

1.4 Statement of Problem 4

1.5 Research Objective 7

1.6 Research Question 8

1.7 Significance of study 8

1.8 Scope of the study 9

1.9 Limitation of study 10

CHAPTER 2:LITERATURE REVIEW 12

2.1 Introduction 12

2.2 Definition of word 12

2.3 Receptive and Productive vocabulary 14

2.4 Lexical Knowledge 16

University of Malaya

2.5 Importance of learning vocabulary 17
2.6 Dimension of vocabulary 19
2.7 Lexical Breadth 20
2.8 Lexical Depth 22
2.9 Fluency of vocabulary 23
2.10 Vocabulary acquisition 23
2.11 Lexical Richness 25
2.12 Lexical Variation 30
2.13 Lexical Density 33
2.14 Lexical Sophistication 34
2.15 Conclusion 36

CHAPTER 3: METHODOLOGY 37

3.1 Introduction 37
3.2 Research Design 37
3.3 Research Participants 37
3.4 Data Collection Procedure 41
3.5 Instrument 41
3.6 Data Analysis 44
3.7 Ethical Consideration 46
3.8 Conceptual framework 48
3.9 Conclusion 50

CHAPTER 4:ANALYSIS OF DATA AND DISCUSSION OF FINDINGS 51

4.1 Introduction 51
4.2 First semester results of Lexical Variation 53
4.3 Final semester results of Lexical Variation 60

University of Malaya

4.4 First semester results of Lexical Density 67

4.5 Final semester results of Lexical Density 76

4.6 First semester results of Lexical Sophistication 81

4.7 Final semester results of Lexical Sophistication 89

4.8 Conclusion 95

CHAPTER 5: CONCLUSION 96

5.1 Introduction 96

5.2 Summary of findings 96

5.3 Pedagogical implication 97

5.4 Suggestion for future research 98

5.5 Conclusion 98

REFERENCES 100

APPENDIX 114

University of Malaya

LIST OF FIGURES

Figure 3.1: Compleat Lexical Tutor webpage screenshot…… 44

Figure 3.2: VocabProfile webpage screenshot…… 45

Figure 3.3: VocabProfile data analysis……46

University of Malaya

LIST OF TABLES

Table 3.1 Participants details First semester ……. 39

Table 3.2 Participants details Final semester…….40

Table 4.1 Lexical Variation First semester…… ….53

Table 4.2 Lexical Variation Final Semester……60

Table 4.3 Lexical Density First semester……..68

Table 4.4 Lexical Density Final semester……76

Table 4.5 Lexical Sophistication First semester……...83

Table 4.6 Lexical Sophistication Final semester……..89

University of Malaya

LIST OF SYMBOLS AND ABBREVIATIONS

LR	:	Lexical richness
LV	:	Lexical variation
LD	:	Lexical density
LS	:	Lexical sophistication
AWL	:	Academic word list
TTR	:	Type Token Ratio
L2	:	Second language learner

University of Malaya

LIST OF APPENDICES

Appendix A: Research Consent Letter…….114

Appendix B: Research Questionnaire……..115

Appendix C: Samples of data……...116

University of Malaya

University of Malaya

CHAPTER 1: INTRODUCTION

1.1 Introduction

English plays a greater role in Malaysia, especially in academics, among educators, academic scholars, researchers, writers, students, teachers, politicians, and parents. This research aims to examine the English written composition of a second language learner in Malaysia. Specifically, this research intends to study on pre-university students' lexical knowledge. The first chapter gives a brief description of the significance of the English Language and its role in the Malaysia education system. This section offers a review of lexical richness in writing and explains the objectives of the research explicitly.

1.2 Education System in Malaysia

This research focused on second language learners in Malaysia. The English language is considered a significant language that spoken over a wide range and used in the education system in Malaysia. Malaysia has a bilingual education system similar to other Asian countries. Bahasa Malaysia is Malaysia's national language, yet English plays an important role in the education system as a second language. A major change in education made by Tun Dr. Mahathir bin Mohamad on 6th May 2002 stated by (Darus& Kaladevi ,2009).

He had changed Malay medium education into English medium education to emphasize the significance of the English language to the nation. This major shift is due to low English Proficiency among Malaysians who are largely monolingual. Reverently,

University of Malaya

the English language is a compulsory subject in Malaysia at the primary, secondary and tertiary levels of education.

The English language is known as an important language in this country. Although Bahasa Malaysia is well known as Malaysia‘s national language and is commonly used by major Malay ethnicity yet the English language is given the same priority in academics. The English language is indeed relatively common in Malaysia due to British colonization in the past. It is no surprise that the English language has been used in Malaysia for hundreds of years. The use of the English language is a major contribution to Malaysia‘s education, tourism, politics, business, employment, media, law, and many other sectors. The importance of the English language translates into endless practicability in significant areas of Malaysia. Therefore, the importance of the English language cannot be denied.

The English language has been given priority in academics, in the organization of work and our daily lives. The English language considered a second language in Malaysia since the 1940s. The English language has been taught as an official medium in primary and secondary schools since the British colonial government. English language therefore also plays a major role in the academic and working world. Fresh graduates struggle to get good results in English to pursue higher degrees in top universities since the demand for English is high. Graduates eventually focus on getting good grades so that they can compete with others for better opportunities. Not only has that, graduates competed to meet the higher expectations of top companies to have a successful career.

Mokhtar (2017) said that both primary and secondary school practice English as a second language set by the Malaysian government. Therefore, primary education aims to equip students with fundamental skills in the English language such as listening,

University of Malaya

speaking, reading and writing. Additionally, grammar skills are also focused to make them produce quality text for different purposes and situations (Mohd Sofi Ali, 2003).The English language syllabus specifies the learning objectives. Nevertheless, the Malaysian Education System provides 11 years of education for primary and secondary students respectively. A total of 11 years of free education offered to Malaysian students by the Malaysian government. After that, they can apply for further studies such as Form Six, Matriculation, pre-university courses, Diploma and Degree studies to enhance their specialization after they have completed their basic education.

1.3 Background of the study

Past vocabulary studies focused most on vocabulary definition, development of lexical, lexical items, lexical richness and many more. Wilkins (1972) stated that through grammar plenty of information can be transmitted, without vocabulary we do not communicate anything. This statement clarifies the fundamentals of vocabulary in a language. It is a vital medium to express ideas in communication. The communication process cannot be accomplished without words. Students generally master vocabulary in the English language to develop greater fluency and expression. (McNeill, 1994; Lemmouh, 2008) argues the significant of lexical knowledge in determining the level of student skills and producing successful communication is undeniable. Vocabulary is the first focus and a fundamental concept of a language encompasses all the aspects of words. In general, vocabulary was considered the fundamental tool as the expression of thoughts. The connection between English vocabulary and writing, known as established second language learning notion. Vocabulary related research focuses on how students learned vocabulary and what words students should learn, vocabulary learning strategies, vocabulary knowledge, and acquisition process.

University of Malaya

Malaysian learners should have an awareness to master the English language to upgrade their skill level. English language learning is necessary as it demonstrates its omnipresent influence in academic. This study focuses on pre-university students who use English as a second language. Lexical richness used as a term in this research. Generally, lexical richness measures the different aspects of a composition‘s use of vocabulary and its richness. Laufer & Nation (1995) justified lexical richness shows quality of writer‘s vocabulary knowledge produced in written composition. Lexical richness checks the strengths and weaknesses of the learner‘s vocabulary. These measures are important to facilitate the qualification of the desired level of education vocabulary. These measurements determine the quality factors in their writing composition that affect vocabulary knowledge and the quality of the student. Thus, this study aims to examine pre-university student‘s lexical richness in their written context to determine a quality standard of writing through the lexical richness.

1.4 Statement of Problem

One of the main issues measured in the field of ESL is vocabulary. English is the second most important language in Malaysia and widely used in daily communication. The education system also focuses on the importance of this language among students in Malaysia schools and institutions of higher learning. Students learn English for eleven years in primary and secondary school. Nevertheless, Malaysian students have low English Proficiency. This became the main concern among teachers, particularly at the tertiary level due to their poor exam results. The consciousness of mastering English among Malaysians is relatively diminishing.

Students are still unable to have excellent command in the English language. Therefore, it is important to examine the issue or difficulty faced by students in

University of Malaya

academic contexts. It is important to find a solution to improve students‘ academic performance. This issue can now be taken into consideration before students‘ English Language is permanently degraded. Malaysian students' fail to have good English skills is due to focusing too much on grammar and syllabus and not on vocabulary learning skills.

To add on, Nation and Waring (1997) clarify that a second language learner should at least know 3000 words as a beginner to acquire minimum requirement to produce a comprehension text. University students must have 10000 to 11000 words to comprehend university text. As a second language learner, students with limited vocabulary sizes may fail to produce quality text. Eventually, this may impair and affect their academic progress with poor results.

Additionally, it is important to measure the different proficiency levels of students to get a better understanding of lexical knowledge. Beginner levels of students tend to have little ability to complete an essay whereas an advanced level of students able to fulfill the writing requirement. Language teachers can discriminate between groups based on English proficiency levels. This will help them to design appropriate language plans and teaching materials for students instead of using general materials. Language teachers can also stress students with low proficiency and vocabulary. Moreover, different language levels like beginner, intermediate and advanced level of English proficiency needs to be addressed.

From my observation, the students in the Foundation in Science from Windfield International College particularlyhave a low level of English language skills. They have difficulty in producing quality written compositions. To add on, the students have poor writing skills in English such as tests, assessments, assignments, essays and academic writings. This can be seen clearly in their grades in the English language course in the

University of Malaya

first and final semesters. The low English proficiency can affect their academic performance. According to Hamzah & Abdullah (2009) students are poor in language skills because of the poor foundation of the language. Students are not lexically rich and do not have sufficient vocabulary knowledge to accomplish an academic task. Students have less efficiency in English writing skills. Their English skills do not meet the expected level of tertiary level learners. Students graduating from private colleges have poor English Proficiency. The majority of students in Malaysia are struggling to speak the English language (The Sundaily 24 January 2019).

Another problem among these students is the poor breadth of vocabulary knowledge. Their low English skills are due to insufficient English vocabulary knowledge. This was considered as a reason the students are not performing well in their academic writings. Students were not familiar with the vocabulary, which makes them perform badly in the written composition. Lack of familiarization of vocabulary is responsible for their limitations in academic writing and word expression. Justification is given by Beatty's earlier study (2013) poor knowledge of vocabulary makes writing poor. Second language learning success is highly correlated with vocabulary knowledge (Waring & Nation, 2004). Writing an academic text is challenging for the students. Besides, poor command of the English Language results in low achievement in their academic.

This was an obstacle for the students who aspired to study abroad. Students are not familiar with the vocabulary, because they do not have many words expressed in writing, this influences them to limit their writing. A claim by Aliakbari (2002) with poor vocabulary it is impossible for students to write their ideas in a text. These students have a poor breadth of vocabulary knowledge resulting in low grades in their academic courses and final exam. Students lack English language skills and command due to minimal exposure to the language. A poor command of the English language has been a

University of Malaya

stumbling block for students who aspire to study abroad and continue their studies because they are unable to meet the expected requirements.

The main concern of teachers at Windfield International College is the student‘s vocabulary in language learning and writing. Teachers are mainly concerned with the vocabulary of the students because it is interrelated with the skill of the student and their academic performance. Lexis plays various elements in writing, such as producing, expanding, demonstrating ideas precisely into a context. A claim by Douglas (2010) lexical richness has a connection with vocabulary size because students should know how to produce a text with the right vocabulary. Students who write lexically rich text receive a good grade in their academics. Teachers at Windfield International College use various materials and methods to deliver successful and impactful lessons. The teachers use different reading activities, speaking activity, vocabulary assessments and many more exercises in the classroom. However, it is not possible to generalize a successful writing skill. Therefore, investigating the lexical richness and its development throughout this course seemed important for this study. There is less research done investigating the lexical richness of written production in Malaysia.

1.5 Research Objective

The research objective is to find out the lexical richness produced by Foundation in Science students in two different semesters; first semester and final semester. The lexical investigation is looking at 3 lexical richness categories; lexical variation, lexical density, and lexical sophistication. Analyzing lexical richness precisely finds students‘ performance in their academic compositions and to measure how rich the students‘ vocabulary knowledge is and to analyze the level of their English skills. This study will give an insight into the richness of the lexicon in written production produced by students.

University of Malaya

The second objective is to compare the lexical richness between the first semester students and final semester students in lexical variation, lexical density, and lexical sophistication. This is to highlight the differences and similarities in the quality of written work and to identify which group of students have produced lexically rich text. In addition, this is to identify which semester students have performed well in the written composition.

1.6 Research Question

1. What is the lexical richness found in ESL students‘ written production in terms of lexical variation, lexical density, and lexical sophistication?

2. To what extent do the compositions of first semester ESL pre-university students differ from those of final semester ESL pre-university students in terms of lexical variation, lexical density, and lexical sophistication?

1.7 Significance of Study

The significance of this research is to give pre-university course students an awareness of writing importance in second language vocabulary learning. L2 Writing considered a major issue for language researchers. Special focus and instruction will be given to pre-university students. This study provides an insight into the lexical richness of writing among learners and teachers. Teachers can evaluate students and what is lacking to produce a good quality text. Such perception will help teachers adapt their teaching styles according to students‘ common writing errors. Teachers usually focus more on finishing the syllabus rather than on language skills. Teachers can prepare appropriate teaching materials for students that will help them to improve their academic writing to achieve good grades. This study will provide insight for students to

University of Malaya

make additional efforts to learn vocabulary in order to produce good quality text. A claim by Aliakbari (2002) due to a lack of vocabulary students not able to produce a quality written text and nothing can be transmitted without words.

Apart from that, students with writing problems may find it difficult to produce good quality text by using the right words in the right context. Vocabulary knowledge shows the quality of a piece of written text. Some learners do not prefer to learn advanced words once they already acquire basic vocabulary in a language. This study could be an insight to learn more words to produce better quality written text. This will help them to achieve a good grade in their academics and to help students incorporate more lexical text in their writings, which may result in lexically rich text. Moreover, the writing skills and confidence of students may eventually improve grades and text in the English language. This study will enable students to have a good command of vocabulary especially writing.

Milton (2008) stated that analyzing students‘ vocabulary knowledge shows vocabulary development of the learning scale taking place than is possible with other language skills measures. Writing to complete a task is not considered successful but it is more than knowing the quality that they have finished the task. It is difficult to write in English as it is considered as a second language. It‘s about using the right word in the right context to produce good content. In essence, writing has fixed grammatical rules involving all language facts such as grammar, lexical, pragmatic and semantics.

1.8 Scope of the Study

This research investigates the lexical richness of students writing by focusing only on three aspects, lexical variation, lexical density, and lexical sophistication. The scope is

University of Malaya

to analyze students writing in the first and final semesters. This is to identify their vocabulary knowledge in their writing performance. This study mainly focuses on vocabulary and ensures to give an insight into vocabulary understanding of students. Their writing measured using computational methods to identify lexical variation, lexical density, and lexical sophistication. Lexical variation is based on the different types of words used and the lexical density is based on the words used in the text. Furthermore, lexical sophistication stresses the rare words produced in a text to indicate the level of the words. The participants are from first and final semester students, Foundation in Science course. Participants considered second language learners with different levels of English proficiency.

1.9 Limitation of study

This is a small study as this investigates the lexical richness of 30 essays by Pre University students at private college. The total number of participants was 30 with 15 participants per group. So if the number was larger, it would have served a stronger basis for the analysis. Therefore, this outcome cannot be generalized since minimum sampling and participation of respondents was relatively small. Participants from the first semester and final semester are known as second language learners with different proficiency levels like beginner, intermediate and advanced. The comparability of the first semester and final semester group of students consisted of an equal number but different English proficiency levels. Different English proficiency in terms of their SPM results. There was no other language test was conducted to find out students‘ proficiency level. Therefore, this study can only apply to this context and the findings cannot generalized to represent the standard of student or teacher or private universities in Malaysia.Another limitation found in this study was incorrectly spelled

University of Malaya

words and grammar errors in the texts. A principle by Wang (2014) taken into consideration when transcribing the written data. The collected data were transcribed without any grammatical correction, but the spelling mistakes were manually changed. This study accepted one missing letter, one wrong additional letter and one incorrectly placed letter as correct. For example, ~~–a~~mire" changed into ~~–a~~dmire", ~~–i~~dl" changed into ~~–i~~dol" and ~~–i~~npiraatin" into ~~–i~~nspiration". Additionally, multi words items were likewise seperated. For example, ~~–m~~yaunt" changed into ~~–m~~y aunt" and ~~–j~~oggingevery" changed into ~~–j~~ogging every". These words were corrected and retained. Because it may change the data analysis. Misspelled words then will be calculated as offlist words. These principles for correcting spelling mistakes were based on the premise that the focus is to explore words.

Apart from that, Lexical richness canbe attempted with different methods, statistical measurement done, but for greater reliability, cross-checking with each other will be better. It was difficult to analyze data for the fewest number of participants that unable to write as indicated by the requirement. The participants did not write a minimum of 350 words, some participants could only write 150-300 words. This was difficult to analyze because the essay is not up to the need. The differences between word length have an impact on the findings, especially in lexical variation. TTR sensitive to word length, therefore findings showed the highest TTR for participants who did not fulfill the writing requirement. Participants who write according to requirementresults with low TTR. Apart from that, some participants exceeded the word limit by producing more than 450 words of text. It is therefore not possible to omit the exceeded words because the objective is to analyze the words. All in all, lexical variation, lexical density, lexical sophistication can be used to account for lexical knowledge of students. We can assume that lexical richness perceives a more noticeable advancement of second language learners.

University of Malaya

CHAPTER 2: LITERATURE REVIEW

2.1 Introduction

Vocabulary is known as a fundamental variable that affects the level of language proficiency of academic writing students. This chapter discusses the research framework in detail. This chapter includes what is vocabulary, education system in Malaysia, lexical richness, lexical variation, lexical density, and lexical sophistication.

2.2 Definition of word.

What‘s a word? Word is difficult to define as easily as people often assume. Read (2000) claims that word cannot be defined in literal terms. Daller et al (2007) defined word as types, tokens, running words, and lemma and word families in a context. Typically word is defined as a significant element with practical meaning. Lexis or words are pronounced, transmitted, processed and decoded in the right way and in the right context to enable others to understand and deliver the intended message. A claim by Zhang (2014) supports the notion of Bloomfield‘s Minimal Free Forms. Words are assumed to be an expression of speech that can stand alone. The word term known as token, type, lemma and word families considered as confusing. Bloomfield defined the word as the minimum meaningful unit of language working definition, orthographic word, has a certain intuitive validity. The idea is to stress the basic stability of a word. Words have the same form but different meanings. For example, bring, brings, brought and bringing. Another example, long, length and lengthens; good, better and best are separate words. Words have extended meanings and even embrace different grammatical categories. Knowing a word involves, understanding the different

University of Malaya

meanings carried by a single form. Bloomfield defined the word as the minimum meaningful unit of language. The idea is to stress the basic stability of a word.

The questions of the vocabulary term in applied linguistics are uncertain. It is essential to investigate the difference between some fundamental concepts. As for words of content, they may have different forms. According to Read (2000) the word defined as an inflected and base form of lemma; and it is similar to a component of speech. Besides that, Bauer and Nation (1993) explain the parts of speech which include inflections such as singular, plural, past tense, comparative, superlative and possessive. The action word of *work, works, working* and *worked* is normally considered as the same lemma with different forms. Additionally, derivatives of words may modify the word class and the meaning. The lemma is not larger than the word family. This is because the word family has a larger scope than lemma. Lemma is the base form under which the word is entered and assigned its place, typically stem or simplest form. Read (2000) defined lemma as the base and inflected form of the word. All items included in lemma considered the same part of speech. For example, *happy* is considered as inflected forms and derivate forms. The word family for *happy,* is also known as *happier, unhappy, happiness* and *happily.* Another example, the words lucky is connected to the same word family such as luckier, unlucky, and luckily through inflection and derivation. The connection of words and meaning shows the changes of meaning. According to Zhang (2014) the word family is a mixture of terms that share similar meanings. However, this depends on how a researcher wants to distinguish the concept of words according to their research purposes.

Lexicon or lexis can be defined as a branch of knowledge that includes meaningful words in a language. Words are so pervasive in our lives. We use it to write and communicate information every day. A claim by Lewis (1993) that vocabulary is known

University of Malaya

as the center of the language. Language primarily consists of a lexical nature. Lewis (1993) developed the Lexical Approach as language consists of grammaticalized lexis not lexicalized grammar. This refers that language is building blocks of lexis not grammar, communication or functions. Language fully concentrates on developing learners‘ proficiency with lexis, words and word combination. This is to stress more importance on vocabulary, not grammar. Vocabulary affects the language acquisition and language competency of a learner.

When considering words, it is important to master its meaning and linguistic form stipulated by Schmitt & McCarthy (1997, p4). To support further, adequate vocabulary size is important to use effective language. Vocabulary makes a learner speak, write, and express their ideas, opinions, and thoughts. The learners with poor lexical knowledge are unable to convey their knowledge through speaking and writing. Learning vocabulary is essential to language acquisition. First, learners should learn vocabulary to acquire a language and effectively produce it. Learning vocabulary is not just a temporary process, but a lifelong process. Vocabulary can draw on the notion of lexis and the entire compilation of language (Barcroft ;Sunderman & Schmitt, 2011).

2.3 Receptive and Productive vocabulary

Receptive and productive vocabulary is one dimension of vocabulary knowledge. Lexical richness categorized into 2 which are receptive and productive claims (Milton 2009). A claim by Nation (1990) receptive and productive vocabulary words encountered learned comprehends and accumulated in one‘s memory via language skills. Receptive vocabulary is known as passive and active with productive vocabulary. According to Alqahtani (2015) receptive vocabulary defined as the ability to capacity to acknowledge a term and recall its definition when it is encountered. First, the learner

University of Malaya

will know a word, and then the learner will recognize the word. The more the learner deals with the word, the more they become familiar with the particular word. The learner will have vocabulary knowledge in their mental lexicon and they can use the word. Learners can acquire receptive vocabulary knowledge through reading and listening. Gallego and Llach (2009) present that students tend to increase receptive vocabulary incrementally over some time. According to Schmitt (2010), automaticity and fluency are considered as a major role in vocabulary development.

Productive vocabulary requires more linguistic knowledge, where a learner has a sufficient amount of knowledge through explanation, practice, a stronger semantic association, and concern to use the word. Productive vocabulary enables memory to retrieve and recall a word and use it wisely in a correct context. To support more, Nation (2001) clarifies that productive vocabulary involves recalling receptive or productive knowledge and producing it in written or spoken discourse. Subsequently, this process involves the knowledge of pronunciation, correct written form, spelling and grammatical structures. Productive studies only take account of tokens and types. Nation (2000) stated that a receptive or productive vocabulary express an idea by using accurate vocabulary in a speech or writing. Students develop vocabulary knowledge by learning more receptive vocabulary compared to productive vocabulary. By considering the productive aspect of vocabulary, students can remember the word ‗impede' and can offer similar words to reflect the meaning. However, this might be impossible in writing (Laufer, 1998). Productive vocabulary knowledge assumed as the words that are understood and can be pronounced by the learners. Productive vocabulary can be regarded as a process of the active word because learners can generate words to express their thoughts and feelings which understood by others Webb (2005) and Laufer et al, (2004); Webb (2008) or to pass on the word as in the original learners' language. Laufer (1998) divides knowledge into productive vocabulary into a controlled and free

University of Malaya

vocabulary. Controlled productive vocabulary knowledge indicates the capacity to construct words when the cue is given while; free productive vocabulary knowledge is the ability to use words spontaneously and without specific encouragement to produce certain words, such as writing independently.

2.4 Lexical Knowledge

Nation (2001), mentioned that first English language speakers eventually know 20 000 word families or 70 000 words. Nevertheless, argued that English second language learners know less vocabulary compared to native speakers Laufer & Yano (2001). A strong Basic English vocabulary is important for language growth. Without adequate vocabulary size, learners cannot have effective communication. According to Kirchner (2013) second language learners‘ vocabulary size was perceived as an important aspect in evaluating their readiness to learn more of the English language especially for first-year tertiary students. They need to have sufficient basic and academic vocabulary repertoire to adapt to the university learning environment, especially when English used as the medium of instruction.

Vocabulary notion is crucial in language acquisition. Knowledge of vocabulary is a fundamental concept of a language that encompasses all aspects of words. Nation (2000) presented three categories to acquire words such as form, meaning and use. The form is on written, spoken and parts of the word. Secondly, the meaning is on concepts, referents and, associations. Third, use refers to the grammatical functions, constraints on use and collocation of words. Read (2000) stated that words have meaning, and words can generate a message. Words form the sentences, sentences compile into a paragraph and paragraphs compose a text.

University of Malaya

Vocabulary learning is crucial for language acquisition (Nation&Meara,2010).It established that lexical competence plays a significant role not only in the second language listening to a claim by (Chang, 2007) but also in L2 writing (Coxhead & Byrd, 2007). Evaluating a learner is easy by evaluating the amount of vocabulary that they know. However, Gallego & Llach (2009) claims that there is noticeable advancement between words and the quality of writing. To justify more, Vermeer (2001) justified that a significant connection found between words and vocabulary richness of a learner.

Vocabulary knowledge is known as a basic term covering all concepts of vocabulary knowledge (Siskova 2012).Words are the core of the language, according to Lewis (1993) and nothing can be transmitted without a word. Additionally, the communication process will not be successful without words. According to Chien (2015) vocabulary knowledge focuses on the meaning of a word and implies how the word fits into a text. However, vocabulary knowledge is a lifelong process where learners can expand it through language acquisition. Milton (2008) stated that it is possible to measure vocabulary knowledge of a learner. By measuring learners‘ vocabulary, it gives a good deal of language development. Another claim by Laufer & Goldstein (2004) lexical knowledge incorporates progressive knowledge where a learner is ready to use the word and use it correctly in writing. This research focuses on vocabulary knowledge of second language learners in a written narrative text.

University of Malaya

2.5 Importance of learning vocabulary

Words counted mainly by type and token in the present research because vocabulary is known in any language as the largest part. To deliver successful communication either spoken or written, vocabulary is needed. A learner is unable to express thoughts, ideas, meanings, without vocabulary (Hadi, 2017). A learner needs to acquire

vocabulary knowledge to have a good English proficiency. Vocabulary is known as demanding issues in second language vocabulary learning and teaching. This is because of large vocabulary size and learners may have difficulty acquiring the words of the language (Nation and Meara, 2002). In language learning, grammatical rules and language systems are more emphasized compared to learning vocabulary. Singleton (1994) stressed that vocabulary is central for language learning. Vocabulary makes up a language and marks the set of words that a particular person can use in communication. Pokrivcakova (2014) described in her article that personal vocabulary is not stable and it grows and evolves (or diminishes) throughout a person's lifetime. Vocabulary‘s role encompasses all the aspects of words in second language learning. A primarily lexical language affects language acquisition and language competency. Singleton (1994) mentioned that the main challenge of learning and using a language is the nitty-gritty of the lexicon. Similarly, Lewis (2008) argued that language learners are having trouble acquiring a larger number of vocabularies in a language. This is impossible to acquire all the words in a language where there is infinity. Laufer (1998) compared the number of words they acquired among foreign learners and native learners. It is important to master its meaning and linguistic form when considering words Schmitt & McCarthy, (1997).For further clarification, Milton (2013) clarifies that vocabulary knowledge shows a relationship in language performance in four key skills such as reading, writing, speaking and listening. This relationship clearly explained in just not focusing on words and highlighted that learners with solid vocabulary knowledge tend to achieve good academic performance. The learner must master good number of vocabulary so that they can express it in their academic language task. However, understanding a word has several levels to acquire the word‘s depth. A strong statement by Jensen & Duffelmeyel (1996) knowing a word demonstrates the knowledge of a learner‘s text. Account for this, Stahl (2005) describes vocabulary not only as a matter of knowing the definition,

University of Malaya

but show how that particular word fits into the text. Knowing vocabulary is not just a matter of having vocabulary knowledge, but a learner is capable of producing the right word in the right context. This is why second language learners must master vocabulary; related language teachers have to put a lot of emphasis on vocabulary teaching and learning among students. Vocabulary interrelated with vocabulary skills in speaking, listening, reading, and quality of writing. Quality of writing closely associated with vocabulary skills and knowledge results as an indicator of the writing standard. To make it clear, learners with a lack of vocabulary do not have enough vocabulary knowledge. However, learning all the vocabulary is not possible for a second language learner. Vocabulary learning is a relational process. Many new theories and insights by Nation & Meara (2010), Schmitt (2008), and Read (2000) have emerged that vocabulary has a direct impact. There are new trends and ideas to stick with the nautical metaphor.

2.6 Dimension of vocabulary

Second language learners acquire and remember one word effortlessly by focusing on sound, tone of voice, perception, tone, pitch and speed of enunciation. Also, the second language learner allocates the word‘s ability and breadth. Learning a word not only l focusing on the number of words, but it also reflects how the word fits into the academic text. At the same time, the number of words plays a significant role where the language learner must acquire any real gains in the overall size of the vocabulary. In account of vocabulary size, it varies between the learners, but it cannot increase vocabulary knowledge. According to Read (2000) vocabulary has different scopes. Vocabulary not only relates to the knowledge and meanings of words but also expands and deepens over time. In another way, learning vocabulary is known as typically a more conscious and demanding process. However, past research has addressed the need

University of Malaya

for vocabulary to be considered as knowledge and a cited model of the process of acquiring vocabulary knowledge. Another concept developed by Daller et al. (2007) on vocabulary knowledge is lexical space. There are three features of lexical space such as lexical breadth, lexical depth, and fluency.

2.7 Lexical breadth

Lexical breadth is about the vocabulary size of a learner. This vocabulary size refers to how many total words the particular learner knows; basically this dimension shows the quantity of vocabulary. In second language learning, acquiring all the words is not considered because there are too many numbers of words. It is difficult for a second language learner to acquire all the words by having only one time of vocabulary instruction. It definitely takes time to acquire vocabulary knowledge through instruction and some other language practices. The learner should engage in vocabulary learning to increase their vocabulary knowledge. Anderson and Freebody (1981) clarified vocabulary breadth as the quantity vocabulary a learner knows about the meaning of the words with minimum requirement. To be more detailed, lexical breadth is concerned with estimating the size of the vocabulary that a learner should know. This depends on the age range, language users like children and adults by knowing the total number of words they know. To support further, Daller et al. (2007), stated that individual variation has the least lexical context, especially in the most frequent 2000 words. This is considered to be learner‘s at a lower level. However, Milton (2007) argued that measuring vocabulary is different between learners where they may end up with compromises test validity. This argument also indicated a predictable connection between word frequency and vocabulary knowledge. Students with huge vocabulary knowledge have the excellent capacity in speaking, reading, listening and writing in

University of Malaya

language skills. English proficient learner can produce good academic quality work. However, it‘s vice versa for a lesser sized learner of vocabulary. Learners with a small size of lexical breadth eventually deal with difficulties in language competence. Vocabularies play an important role and contribute to almost all aspects of language skills. Brown (1996) further clarified the significant contribution of vocabulary interrelated with broad consensus with vocabulary growth. There are a few vocabulary tests to measure learner‘s lexical richness. For an examples, Vocabulary Level Test, Academic Word List, Vocabulary Size Test and so on. Lexical knowledge matters in language learning. According to Webb, Sasao, and Ballance, (2017), the test was divided into five segments such as 2000, 3000, 5000, university and 10000-word levels. The first level is measured at 2000 - 3000 high-frequency word levels, the second level, and another 1000 word level is used by a learner on rare words or low frequency. To add on, the Vocabulary Level Test is an indicator of vocabulary coverage in a text and an accurate measurement to find the level of learner‘s vocabulary knowledge. In another, Singleton (1994) mentioned the knowledge of a language determined by the use of lexicon in a text. Without words, learners are unable to understand a language and convey it in an academic text. As stated by Nation and Waring (1997) a language learner should have 3000 to 5000 word families. Whereas for Schmitt (2010) a range of 16000 – 20000 word families for native speakers appears to be fair. Lexical breadth connected with vocabulary size because it will increase the number of possible links between the words in a lexicon. The second dimension is on lexical depth. Anderson &Freebody (1981) explained that lexical depth is more about the quality and depth of the words. This refers to the depth of learner‘s vocabulary knowledge. We can assume that a learner with a strong lexical depth conveys excellent ideas in a text.

University of Malaya

2.8 Lexical depth

Read (2000) highlighted three categories of the depth of vocabulary development. First, having a limited and lack definition of a word. Second, vocabulary knowledge includes characteristic and comprehensive word knowledge. Thirdly, using the vocabulary from memory and use it in an appropriate sets. Developing the lexical depth of a learner involves increasing some connections between the words in the lexicon. Milton (2013) mentioned that there is a connection between lexical breadth and lexical depth; it connects through different combinations of words. Knowing a word is not just about vocabulary knowledge, but it is more about using this communicable knowledge and performs them in a variety of circumstances. A learner should be able to use vocabulary knowledge to vary and use their lexical depth in language skills to some extent. Language learning is not a short process, but it takes time to thoroughly acquire knowledge and depends on the learner. In the beginning, some learners tend to acquire vocabulary, while others are more likely to acquire it later. Zhong (2011) highlighted that learners should have complexity and multidimensional vocabulary knowledge. Lexical depth reflects a context that a learner is able and understands the definition of a word. For example, the learner knows a different range of language skills in spelling, pronunciation, register, phonology, morphology, syntax and semantic relationship with other words. Shen (2008) proves that vocabulary depth as an indicator of reading comprehension compared to lexical breadth. He proved in his study that there is a positive relationship between lexical depth and lexical breadth. His hypothesis was that the larger the quantity of words the greater the depth of learners‘ knowledge. Past researchers researched different methods with different measurements to measure the dimension of vocabulary knowledge. Vocabulary assessment like matching words with definition, translating words, finding word synonyms, discussing by Wesche and Paribakht (1996) checklist. Nation (1990)‘ Vocabulary Level Test known as a common

University of Malaya

measurement to calculate the lexical richness of students. This test particularly measures how much a learner knows about vocabulary. Measuring vocabulary knowledge is important because in language skills we can check the ability of learners to fulfill tasks. Measuring vocabulary, on the other hand, gives an insight into the learner‘s vocabulary growth in a wide range of vocabulary across a group of learners.

2.9 Fluency of vocabulary

The third dimension is fluency, considered an important concept of lexical competence is the ease with which second language learners acquire their vocabulary knowledge to use. Doczi and Kormos (2016) argue that fluency has to do with speed and automatic access to lexical items and their depth of word knowledge. The relationship between the quantity of vocabulary that a student knows and the range of vocabulary that a student knows is quintessential. This shows that a learner is ready to use vocabulary for fluent recognition and production that takes place. According to Laufer and Nation (2001) a learner can improve their lexical fluency through repetition of action by strengthening existing lexical items. Reading comprehension, oral practice tasks, regular free writing, silent reading, and vocabulary sources also play a significant role in learning vocabulary. Fluency doesn‘t come overnight, and the learner has to work hard and put effort to improve vocabulary fluency.

2.10 Vocabulary acquisition.

Vocabulary plays an important role in second language research. Without words, no ideas can be produced by a second language learner as mentioned earlier. They cannot produce any text without words. Vocabulary is definitely known as a complicated issue

University of Malaya

for students because a restricted vocabulary impedes good language transmission in a second language. According to Schmitt (2000) he emphasized that lexical knowledge is the key to communication and acquisition of a second language. While vocabulary is the main model for mastering a language, Krashen (1989) claims second language learner gives importance to words, not grammar. Wilkins (1972) stated, "Without grammar very little is conveyed, without vocabulary nothing can be conveyed". Vocabulary knowledge allows a learner to use language, conversely, and generate successful language skills in an academic context. Many language researchers have focused on vocabulary issues. Vocabulary has since been a trend in the field of language research. Vocabulary needs to be demonstrated daily among second language learners in an academic context. Lexical knowledge in the second language spectrum is now known as an absolutely crucial factor. Past researchers found that acquiring vocabulary is essential to complete written and spoken text in second language learning. Vocabulary acquisition helps a learner with an appropriate vocabulary to use the structures and functions for comprehensive communication. Research shows a major problem of poor vocabulary between second language learners. When a learner wants to produce a written text, the learner needs to have a vocabulary in their mental lexicon so that they can express their thoughts in a written text. The learner should have the required vocabulary size to express their meaning and concept. In a bird's eye, vocabulary has no rules and regulations that a learner should follow to acquire and develop their knowledge. Singleton (1999) clarified major learning challenges and using a first or second language is generally syntactic rules in the lexicon. According to Read (2000) language acquisition is an active process where a learner aware of vocabulary learning. Learning vocabulary is not just memorizing second language words; learning vocabulary depends on the dictionary as their language guidance. I fully agree that a learner should know the importance of vocabulary and the purpose of learning it and not

University of Malaya

just memorizing the words and claiming to have adequate vocabulary knowledge. Past studies literally emphasized learning vocabulary is a significant, vocabulary issue in second language learning, and ways to enhance the vocabulary of the learner. Schmitt (2000) demonstrates that vocabulary is not just related to linguistic units, but interrelated with access to language skills and communicative process. In addition, he argued that the context of the classroom focused primarily on the range of vocabulary learning among students. However, students face difficulty to acquire vocabulary. Simply put, vocabulary is far more different from phonology and syntax. There are no rules for learning vocabulary. Perhaps this is because learning vocabulary has no guidelines, such as what are the first 100 words a learner should know or a vocabulary list according to dialogs or conversations.

Teachers primarily focus on rushing with the English syllabus and not helping to acquire vocabulary in the language learning process. Teaching vocabulary is complicated because teachers do not have an idea on what to start in vocabulary teaching. On the other hand, language teachers should prepare and invent new methods for teaching vocabulary. Teaching and learning vocabulary is a critical process. Language teachers should able to teach by using the latest techniques, up to date materials, games and arts to create a successful environment.

University of Malaya

2.11 Lexical Richness

Read (2000) mentioned that lexical richness defined as measuring lexical use and characteristic of a learner. Word is the primary focus in lexical richness aims to otherwise, the lexical reservoir of the learner determines the uniqueness of one's vocabulary defines the variety in a text. Lexis plays a major role in writing composition, not only in deciding the quality of text but also to determine different types of a word

too. A claim by Malvern &Richards (2013) stated that lexical richness question about the quality of vocabulary helps the students to do well in their academic writing while lacking of lexical knowledge trouble the students in their writing. Wang (2005) mentioned that lexical richness is the main key in scoring good grades. According to Failasofah et al. (2018) vocabulary is the center of a language, whether a native language, second language or foreign language. This reflects the level of language. What we imply is that lexical richness is successful, but not an exclusive, predictor of skill. Laufer& Nation (1995) highlights the thought of an individual vocabulary‘s expansiveness and depth. Lexical richness helps non-native speakers to acquire vocabulary and improve their language skills (Crossely et al., 2010pg (47) &Jukneviciene, 2007).

To discuss further, the term lexical richness is wider and the term is lacking claims by past researchers. Nation and Webb (2011) defined lexical richness in the same construct with lexical variation, lexical density, lexical sophistication, lexical complexity, lexical diversity, and lexical individuality. Leonard (1986) defined the term lexical variation is the percentage of different words in a written document which traditionally measured by using Type Token Ratio. Lexical density is focused on lexical words, which are also known as content words. Content words are known as nouns, verbs, adjectives, and adverbs. Whereas lexical sophistication is defined as the percentage of low frequency or advanced words. Another dimension of Read (2000) vocabulary errors considered as an indicator of lexical richness. In a glimpse, Laufer and Nation (1995) mentioned, lexical richness is all about measuring a learner‘s vocabulary size and is known as a reliable measure to study learner‘s vocabulary knowledge. Reliable results can be obtained in the lexical richness that has also been produced over time by the same learner with different pieces of written text. Arnaud (1984) claims that lexical richness known as lexicometric and stylostatistical research.

University of Malaya

Nation and Webb (2011) quality of vocabulary knowledge defines lexical richness in a text. First, the vocabulary size used in a text such as tokens, types, lemmas, and families considered as a text variation. It is advisable to use different variations of vocabulary as a result of a text‘s richness in a written work. Engber (1995) and Malvern and Richard (2002) clarified that a text with a smaller size of words with too much repetition is not considered as lexical richness. Vice versa, a text with different variables and less word repetition considered as lexically rich text. Additionally, the text should have a lower frequency proportion of advanced words. Lower frequency or rare words are known as uncommon words writers write to describe the meaning in a precise and sophisticated way. Compared to a basic text, a text with some higher difficult and advanced words is lexically rich text. Moreover, the precision of form with less misspelled words is also included in the richness of the text. A text with a wider range of right spellings and pronunciation, results in lexical richness. Engber (1995) claimed that a higher proportion of semantic accuracy and specificity in meaning by providing more accurate meaning. A text that uses few words with exact meaning that makes the writing more understandable. The lexically rich text will have precise language consisting of clear and direct words with direct concepts related to particular subjects. Apart from that, the writer may use more semantic related words to display a richer text than less semantic related words. This is more likely to contain different sets of words such as synonyms, antonyms, coordinates, superordinates, and subordinates. Crossley and McNamara (2009) suggested that first language writers have a higher number of verbs compared to second language writers. A higher proportion of content words show a text‘s lexical richness. Therefore, grammatical accuracy is to engage in the richness of the text. However, this study did not focus on grammatical structures and lexical errors. Lexical errors are due to the use of a word that requires a formality standard in content. Errors can give a poor measure of lexical richness. A claim by Engber (1995) fewer errors in

University of Malaya

the text considered richer in lexical. At the same time, errors are also seen as a shred of evidence that the learner uses a definite language system in the development of language learning development. However, errors are not focused on this research. Another point about the lexical knowledge in a text is the relationship between time, speech and writing rate. For example, quantity of words produced within a minute. This indicates how knowledgeable a person is in lexical terms Apart from speaking and writing, reading rate is also applied by looking at greater variation in word types and a larger proportion of low frequency or difficult words may increase the time it takes to read a text.

Lexical richness focuses on the quantity of word in a text. Lexical richness measures vocabulary size in written production. According to Milton (2008) measuring learners _vocabulary knowledge can help provide insight into the scale of learning that takes place as much as possible with other language skills measures. Lexical richness can display the vocabulary knowledge in content (Jarvis 2002). It recognizes progressive and less capable learner‘s dialect. A claim by Engber (1993) and Linnarud (1986) measuring lexical richness makes it possible to discriminate learners with different proficiency levels. This claim makes sense because learners with higher vocabulary sizes will do well in the academic task. For example, the learner is capable to read a question, identify the task, draft the possible answer and compile it into a text. Without vocabulary, a learner cannot do or express his or her thoughts. Additionally, Laufer and Nation (1995) justified that lexical richness as the quality of lexical knowledge in the written text of a learner. The types of lexical richness known as lexical variation, lexical diversity, lexical complexity, lexical individuality, lexical density and lexical sophistication claim by Linnarud (1986). This research only focused on three lexical types such as lexical variation, lexical density, and lexical sophistication. Nation (2001) argues it is essential to measure the degree of words used by second language learners.

University of Malaya

According to Laufer (2005) and Laufer and Goldstein (2004) lexical richness is the words that learner stores in the mental lexicon and use it in real-time production. Lexical richness divided into three subcategories claims by Laufer and Nation (1995) lexical variation, lexical density, and lexical sophistication. Laufer & Nation (1995) define lexical variation is the different percentage of types in a written production and lexical density" as the percentage of lexical words, nouns, verbs, adjectives, and adverbs. Lexical sophistication is the percentage of low frequency or advanced words. Lexical richness used in this study to analyze the lexical knowledge of second language learners in written production. Laufer and Nation (1995) presented that lexical richness checks writers' lexical knowledge whether they have used a varied vocabulary.

A study by Signes and Arroitia (2015) focused on Spanish learners in writing work at the beginning and end of the semester on lexical density and diversity. This longitudinal study indicated that the linguistic evolution of the student has significant progress throughout the writing process and it is possible to rate the learner's proficiency level. By looking at statistical measures that show a high frequency of word families of lexical density and diversity it gives a clear picture of the student's language level analysis. However, low-frequency word is also known as an indicator of the vocabulary knowledge of a learner. Lemmouh (2008) examined grades of students at Swedish University by looking at the vocabulary in the written production. The main objective was to highlight the vocabulary richness as a predictor of the quality of the essay. He focused on measuring productive and receptive vocabulary knowledge using a lexical frequency profile within two groups of participants. However, as reflected by the ratings of college educators, there are no significant results between lexical richness and the overall performance of the essay. However, the essay's quality is mainly based on content and grammar characteristics and not lexical characteristics.

University of Malaya

Lexical Frequency Profile uses statistical significance to conclude that students who have used a high proportion of advanced words pass with distinction where the grade depends on the vocabulary usage by students. Another qualitative study by Azodi, Karimi, and Vaezi (2014) showed a significant improvement in their lexical knowledge from 2nd, 4th, 6th and 8th semesters students. Results have been justified according to Laufer and Nation (1995), different skill level learners have different frequency profiles, and changes in text profile show students‘ vocabulary growth. At the same time, participants have also been exposed to a variety of vocabulary instructions that make them express their vocabulary knowledge in their written work. Also, Signes and Arroitia (2015) found that reliable results can be obtained from two pieces of written text from the same learner. This research was conducted to analyze lexical words and a variety of words used by students at Valencia University. The data compared with the first and final semesters and the results showed good progress in the quality of writing. They concluded that it is reliable to discriminate the proficiency levels of learners. Another past study by Siskova (2012) was on Czech EFL learners‘ narrative of lexical richness by comparing the relationship between different measures. Participants wrote a picture-based story and these data were analyzed using computational methods. Results have shown an impact on lexical density and sophistication. By measuring lexical richness, a learner‘s lexical knowledge can be interpreted (Nation &Webb, 2011, p245).

2.12 Lexical variation

Lexical variation is the same term with lexical diversity and lexical range, defined by Nation and Webb (2011). Lexical variation measures a variety of vocabulary in a written text. This measurement track changes in the variety of vocabulary used by a learner in a computational method. Lexical variation assesses the ratio of token and

University of Malaya

types in the text using the Type Token Ratio (TTR). (T) considered as the word types ratio and (N) is the total words in a text (Templin, 1957). TTR is a common measurement for researchers to find out a text's lexical richness in second language learning. It is also used to analyze texts in stylistic, neuropathy, language acquisition and forensics (McCarthy& Jarvis, 2007). Types refer to different words and token refers to the total words in the text.TTR dividing the types (the different number of words) with tokens (the total number of words). TTR represents the size of the vocabulary; is sensitive to the range of text as the value decreases when the text becomes longer. TTR will decrease when word measurement increases. TTR is measured by a formula on the types and tokens in a text by dividing the number of types and token in a written text. Read (2000) mentioned that tokens and types are considered as an important unit for measuring the total words in a text. Tokens and types are not the same distinctions. The token is the total words in a text, and included with the repeated words and it is counted as a token. Vice versa types focus on the different words excluded by repeated words in written production. Repeated words aren't counted. Word repetition eventually increases the total token in a written work. For example, there are a totally 5-word forms in the sentence *the woman saw the baby,* but the word *the* appear twice, so this considered as 4different word forms. The above sentence has 5 tokens and 4 types. To support more, Zhai (2016) mentioned that higher lexical variation shows that a learner has a larger vocabulary and is capable of using synonyms and other related words and not using repeated words. TTR‘s disadvantages are that it is sensitive to an essay length. Type Token Ratio sensitive to total words written in an essay, the smaller amount of new or different types written reduced Type Token Ratio. A low ratio indicates that a student has repeated the same vocabulary many times in a text. Furthermore, this measure has limitations because of the sample size, as the ratio tends to decrease as the sample size increases (Arnaud et al., 1992). Past researchers have proposed alternative methods to

University of Malaya

resolve this conflict. First suggestion by Engber (1995) to use Mean segmental TTR (MSTTR) by making one way to improve TTR.MSTTR calculated by dividing a sample into successive segments of a given length and then calculating the average TTR of all segments. Some priority was given to the size of the length of the text so that TTR was measured with the same length of text.

In spite of the fact MSTTR decreases the issue of quantity words, Malvern et al. (2004) criticize that they cannot focus on the standard sample size. Based on my opinion, this critique is because it is not possible to conduct an activity asking the students to write just exact 100 words based on a topic. If the student produces 110, by using MSTTR only 100 words be measured. At the same time, the balance of 10 words considered a waste of data. It can give insight words on different types and researchers won't be able to analyze a data validly. Gijsel et al (2005) claim that a corpus with a short sample yields less clear results. It is not advisable to shrink the text or use only a limited number of words for research purposes, it will affect the result. Secondly, another solution by Broeder, Extra and Van Hout (1993) to replace TTR with Index of Guiraud, which calculates the number of types and the tokens by reducing the length of text. Below is the formula to calculate TTR:

Type Token Ratio = number of types / number of tokens * 100

Limitation of this tool is the inability to find richness of words in a short text. The solution for this is to fix a number of words in a composition by resolving this conflict. But distinguishing the quality of different words in a composition is still not reliable.

University of Malaya

2.13 Lexical density

Originally Ure (1971) defined lexical density is the calculation of content words in a written document. Lexical density separates words of function and words of content. Lexical words and content words are similar and words of function known as grammatical words. The relationship between types and token is measured by lexical density (LD). As proposed Laufer and Nation (1995) density categorized as nouns, adjectives, verbs, and adverbs. These categories of words primarily convey information and core subject matter; explained the text content and information in a text. Content words or function words express thoughts in a written work known as dense if the percentage of density is high relative to the number of words. Density is an indicator of lexical measurement by evaluating the lexical knowledge of the learner and their ability to use lexis in written production.

The calculation shows the range of several lexical words and content words in a text. Halliday (1985) and Ure (1971) discovered that density in written is higher compared to verbal communication. This is because Ure (1971) claims that there needs a lot of planning and communication level. Linnarud (1986) and Engber (1995) argued there is a low and insignificant correlation between lexical density and holistic ratings of L2 writing. Content words are needed to convey expression in a text. Lexical density differentiates the terms, ideas or characteristics of written and spoken. Content words express a context with the fullness of information, and higher density shows that the text is dense. Laufer and Nation (1995) demonstrated lexical density normally calculate a text‘s richness as it is related to structure and characteristic of written text. Lexical density is sensitive to function words that affect the validity of the text. Good writing composition assumed that there were few errors in the use of words. However, this study does not focus on lexical errors. From Linnarud (1986) and Read (2000) definitions of lexical richness, it can be seen that similar aspect of lexical richness is

University of Malaya

lexical variation, sophistication, and density. Yu (2009) said that the concept of lexical is not accurate because the concept is interrelated with other lexical types. Each researcher has their justification and term to identify lexical richness according to their studies. Researchers have different research settings, findings, and results. According to this study, below is the formula to measure lexical density:

Lexical density = number of lexical words / total number of tokens *100

Past research shows that there is no deep revelation or justification made in learner's speaking and the quality of lexical density. It depends on the researcher what are the words are categorized into parts of speech. For example, for lexical adverbs, O'Loughlin (1995) accepted all adverbs of time, manner, and place, and Engber (1995) accepted ~~–a~~dverbs with an adjective base, especially those with an -ly suffix". Ure (1971) critics those past researchers were not able to define lexical terms to identify the density of a text. In this study, lexical words defined as nouns, adjectives, verbs, and adjectives. Content words and lexical words measured as the ratio of a written text shown in the formula before.

2.14 Lexical sophistication

Read (2013) defined lexical sophistication as a rare word or advanced word utilized in written production. It is identified with how the words are used in content. Lexical sophistication is known as a rarity indication level of difficulty. An advanced word or rare words refer to a limited low-frequency vocabulary used in their text by a particular learner. According to Read (2000) the lexically rich text will have a greater word with

University of Malaya

semantic network, synonyms, coordinates, superordinates, and subordinates than less skilled learners. Less proficient learners have poor lexical knowledge where they will include basic words in their writing. Aviad and Laufer (2013), reviewed the degree of the frequency words is various words level written by a student in a text or the overall extent of vocabulary level according to frequency. Yi andLuo (2013), express those messages that have a smaller amount of fundamental words are rich. Content should contain a higher amount of advanced words contrasted with essential words. Lexical sophistication can distinguish lexical knowledge of the second language learner. It indicates the number of vocabulary that a learner can deliver in a written production by using the right word. A quality piece of text is known as ―a selection of words relevant to the subject or style of writing, and not just general, everyday vocabulary, enabling writers to express their meanings in a precise and sophisticated manner" (Read, 2000, p. 203). It actually refers to the number of low- frequency words that were considered difficult. Laufer and Nation (1995) investigated the reliability and validity of the frequency as a measure of written compositions of lexical sophistication. This was developed by Laufer and Nation (1995)which is ―the percentage of words that a learner uses in his or her writing at different vocabulary frequency levels" or ―the relative proportion of words from different frequency levels" (p.310).Read (2000) defined lexical sophistication as a rare word or advanced word used in a text. It is related to how words are used in a text. Vocabprofile give relevant information about the vocabulary frequency that is written by the first and final semester student. By looking at vocabulary frequency we are able to show student's lexical richness. Cobb & Horst (2004) listed first 1000 words are known as declaration, generalization, and claims related to justification. The second 1000 list is more on argument or discussion about a topic. Other than this, the Academic word list focuses on analysis, interpreting concept and demonstrating valid statements. Off list is to concede, assert and to premise a

University of Malaya

justification or ideas in a text. Nation and Webb (2011) stated that text that with well-known words is richer than fewer numbers of basic words. A text should contain more advanced words than basic words. Lexical sophistication can show the level of lexical knowledge that the second language learner has earned. It indicates the greatest vocabulary size that a learner can produce in a written production by using the right word.

Below is the formula to calculate lexical sophistication.

Lexical sophistication = (number of advanced tokens / number of lexical tokens) * 100

2.15 Conclusion

As a conclusion, Laufer (2005) said, there is no accurate measure to calculate lexical knowledge and lexical use of a learner. Past literature gave clear insights into different views of the concepts of vocabulary richness. However, the current study focuses on lexical richness in students‘ writing to check the vocabulary knowledge of students in written production. To emphasize more on vocabulary size and lexical richness, this study obtains a brief idea about lexical measures in second language learners. Past researchers conducted research using various vocabulary analyses on the validity in spoken and written in different settings, but the results are different. To add to this, there are very few studies in Malaysian settings on lexical richness. Thus, this study aims to own better understanding of lexical richness in second language learners in Malaysia.

University of Malaya

CHAPTER 3: METHODOLOGY

3.1 Introduction

This chapter will discuss the procedures of data collection that were conducted in this study. Below is a brief description and explanation of the participants, research instrument, data analysis and ethical consideration of this study.

3.2 Research Design

This study used a qualitative research written composition of students doing the foundation program at Windfield International College. The collected data (students' essay) was analyzed on *http://www.lextutor.ca/* which is a free software for the researcher. This software was used through a vocabulary profile program to analyze lexical variation, lexical density, and lexical sophistication. Data were analyzed to answer research questions.

3.3 Research Participants

This research conducted in aim to study the Lexical Richness of second language learner particularly from Windfield International College. There are several departments such as Business, Healthcare, Medical Science, Information Technology, Nursing, and many other departments. Participants for this study were Allied Science Faculty undergraduates enrolled in a Foundation Science course. Only students from the Department of Foundation in Science (FIS) were selected for this study. The Foundation course is usually about 12 months long, and this is considered a short course for the

University of Malaya

students. This course conducted in three semesters. Students take the English Language in the core academic subjects. Students choose further study areas and prepare for a smooth transition into science, engineering or information technology. For this program, Windfield International College uses English as a medium of instruction.

There are four main intakes for this course per year, which is in January, March, June, and September. Each Intake has 3 semesters with 14 weeks of lecture. The students take English Language A in the first semester and English Proficiency in the second semester and English Language B for the third semester. Based on this study, only students from two semesters were selected as research respondents. The participants of this study are taken from 2 intakes that were January intake first semester and September intake final semester. A total of 15 participants took part in the data collection from January Intake 2018 first semester, particularly in English Language A. On the other hand, another 15 participants were in English Language B from September Intake 2017 final semester. All participants were Malaysians aged between 17 to 18 years. The participants speak their own mother tongue as their First languages such as Malay, Mandarin, and Tamil. They practice English as a second language. Besides that, all participants have completed SPM and their English results have been collected. This is to ensure their proficiency and their background of the English language. Participants‘ age, race and their English SPM results attached in Table 3.1 for first semester participants and Table 3.2 for final semester participants. However, total participants were 30 students with a sample of 13 Indians, 9 Malays and 8 Chinese participants. For gender preferences, they were 19 females and 11 males. Details of the participants attached as follows.

University of Malaya

Table 3.1: Participants Details First semester English Language A - January Intake

STUDENT IDENTITY	AGE	RACE	SPM RESUTLS	ENGLISH AS FIRST LANGUAGE OR SECOND LANGUAGE
Student 001	18	Chinese	A	Second language
Student 002	18	Indian	C+	Second language
Student 003	18	Chinese	A-	Second language
Student 004	18	Malay	C	Second language
Student 005	18	Chinese	B-	Second language
Student 006	18	Chinese	B-	Second language
Student 007	18	Malay	C-	Second language
Student 008	18	Indian	C+	Second language
Student 009	18	Indian	C-	Second language
Student 010	18	Malay	C-	Second language
Students 011	18	Malay	D	Second language
Student 012	18	Malay	C-	Second language
Student 013	18	Malay	C+	Second language
Student 014	18	Malay	D	Second language
Student 015	18	Malay	A-	Second language

University of Malaya

Table 3.2 : Participants Details Final semester English Language B - September Intake

STUDENT IDENTITY	AGE	RACE	SPM RESUTLS	ENGLISH AS FIRST LANGUAGE OR SECOND LANGUAGE
Student FS 001	18	Indian	A	Second language
Student FS 002	18	Indian	A	Second language
Student FS 003	18	Indian	A-	Second language
Student FS 004	18	Chinese	B	Second language
Student FS 005	18	Chinese	B-	Second language
Student FS 006	18	Indian	B+	Second language
Student FS 007	18	Chinese	B-	Second language
Student FS 008	18	Indian	B-	Second language
Student FS 009	18	Indian	C	Second language
Student FS 010	18	Chinese	B+	Second language
Students FS 011	18	Indian	B-	Second language
Student FS 012	18	Indian	C	Second language
Student FS 013	18	Indian	A-	Second language
Student FS 014	18	Indian	B	Second language
Student FS 015	18	Indian	B	Second language

University of Malaya

3.4 Data Collection Procedure

This study used qualitative methods for data collection and analysis. The data collected from two sources; demographic survey questionnaires and student sample written production. Questionnaires distributed to students to find their language background. The respondents were required to answer all the questions. The first objective is to evaluate the lexical richness of the students in their written production. The *http://www.lextutor.ca/* fully used to distinguish students‘ lexical richness in terms of variation, density, content words, and function words. The second research aim is to compare the lexical richness in lexical variation, lexical density and lexical sophistication between the first semester students and final semester students. Lastly, student‘s language background has been taken into consideration in analyzing its effect on the quality of writing. A total of 30 written compositions obtained from Windfield International College to achieve the above mentioned goals.

3.5 Instruments

A demographic survey was distributed to students to find their language background. Particularly because the English subject has less specialized terminology than Biology or Chemistry, the English language B class has been invited to participate. Participants write an essay on the topic "The person I most admire." This topic was chosen because this particular topic is in the English Language A syllabus and the students are already familiar with the topic. This essay of description serves to define a role model as an example that inspires others to live meaningful lives. Participants were required to write an essay within one hour for at least 350 – 450 words and were not allowed to refer to a dictionary or discuss the topic. All participants had to write on the same topic to maintain the reliability of the study. Participants had to express their knowledge from

University of Malaya

different angles, not limited to any direction of writing. Participants can express their thoughts on their role models as to how they play an essential role in their positive development in life. At the same time, they can focus on the negative impact of role models and find out why they like and don‘t like the person. Since it‘s a general topic, the participants can write any points. Precisely, there is no restriction on writing this essay. This is to ensure that the students can write on the basis of their lexical knowledge without any need for the necessary expert knowledge of the subject. Participants were given no guidance to write the essay such as keywords, main points or quotes. They can write about any person they admire based on the related topic. They were given an hour and required to submit to the instructor upon the compilation of the essay. At the same time, they are not allowed to use dictionaries or verbal guidance from the instructor. The data typed and transcribed into Microsoft word documents without changing any grammar or sentence structure. The collected data (students‘ essay) were analyzed using online analysis software available on *http://www.lextutor.ca/* which is free software for a researcher. Lextutor is a good tool to predict and assess students‘ vocabulary knowledge.

Laufer and Nation (1995) suggested earlier that Vocabprofile can be a significant measurement in assessing learners‘ vocabulary production. Laufer & Nation (1999) claim that it is possible to get a reliable result through Language Frequency Profile in students writing. The study also claims that it evaluates students‘ proficiency level in English. Findings provide valid and reliable insight on vocabulary growth. Siskova (2012) used Vocabprofile to measure lexical richness focusing on lexical diversity, lexical sophistication and lexical density of Czech EFL learners. This study compares the results of different vocabulary measurements. The data was statistically analyzed using correlation analysis to confirm the study‘s assumptions on scales, independence, normality, and linearity. Findings showed that through this measurement, its abilityto

University of Malaya

demonstrate lexical types of lexical diversity (saying how many different words are used), lexical sophistication (saying how many advanced words are used) and lexical density (saying what is the proportion of content words in the text) of students' vocabulary knowledge. Results showed a significant correlation between lexical diversity and lexical sophistication and poor correlation between lexical diversity and density. A past study by Abduh &Rosmaladewi (2017) examined students' vocabulary levels in business English students through lextutor. This study able to measure students' ranges of vocabulary. The study assumes that the students are familiar with higher-level academic words, but the results have shown that participants only know with basic words, academic lexis, and terminologies. Another study by Nur (2015) using lextutor to analyze students' essays and academic textbooks, has successfully managed to show the students' vocabulary level. Findings showed that Vocabprofile is a significant tool that interprets percentage of vocabulary, usage of K2 and AWL words in their text. This method provided insight into the use of correct academic vocabulary in written text.

There are criteria that use the *http://www.lextutor.ca/*program to measure vocabulary for students. The Vocabprofile indicates the range of words according to the word frequency list. The frequency shows 2000 common words list West(1953), and Academic word list by Coxhead's (2000).This frequency list shows the list of words used by students in their written production. Offlist words include vocabulary which is not in the common and academic word list; unlike measures based on type-token ratios suitable to identify learners' vocabulary frequency and their vocabulary development. The Vocabprofile shows the common, academic and out of list words. Daller et al.(2003) claims that frequency list gives an insight of learner's ability of vocabulary usage in a text The Vocabprofile draws in a text on frequency lists such as _sophisticated' or _unusual' words in a text. This shows a significant picture if a variety

University of Malaya

of words used by the second language learners. Laufer and Paribakht (2000) study found that common frequency of a word is 1000 to 2000 and the rest included in academic and off list words. This justified as valid measurement and predictable to discriminate learners by looking at their academic text.

3.6 Data Analysis

Below is the quick procedure on how Lextutor works according to Szudarski (2018) guidelines.

Figure 3.1: Compleat Lexical Tutor webpage screenshot

Figure 3.1 above shows the process involved in the analysis of data. There are various types of programs in *http://www.lextutor.ca/*. For example, group lex, corpus correct, range, text tools, frequency and much more. Figure 1 shows the Vocabprofile used for this study. We have to click on vocabprofile and there will be 3 other sections on this part. There areVp-kids, Vp-classic, and Vp-compleat. For this study, Vp-classic

is used to analyze the data. Vp classic operates through Laufer & Nation's original four-way sorter. It operates two frequency indicators (K1 & K2) mainly from the General Service List. Additionally, it shows lexical indicators such as the Academic Word List (AWL). Vp Classic is for smaller files compare to Vpcompleat. This maximum word limit for this study was only below 600 words.

Figure 3.2 VocabProfile webpage screenshot

Figure 3.2 shows that the process of student's text has been inserted into vocabprofile. Next, copy and paste the entire text produced by the students into the box and submit. The analysis shows the results of the type-token ratio, the percentage of density and the frequency of words by rank. There is no complication in the size of the text since students' writing is not more than 500 words in this study.

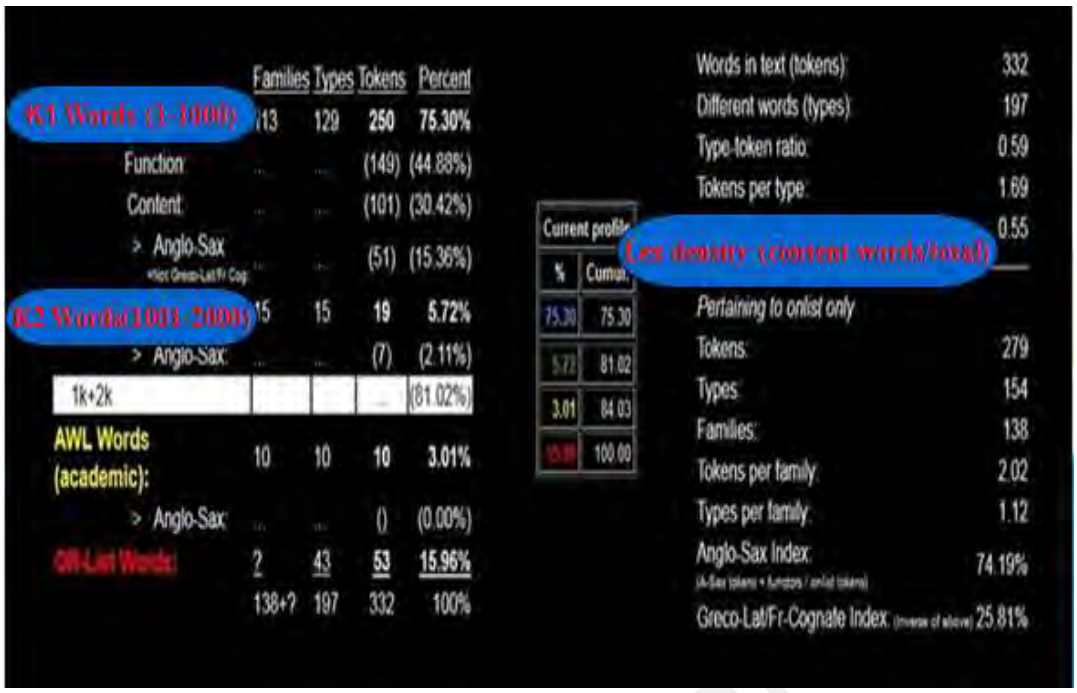


Figure 3.3: VocabProfile data analysis

Figure 3.3 shows a sample text result. Results show the type-token ratio, show total words and different words in the text of the student. Besides that, the analysis also shows the percentage of lexical density. Finally, the software also highlights the percentages for K1 words, K2 words, Academic Word List and Offlist words. A typical text is usually written texts with 70 percent of 1000 frequent word families, 10 percent of the second frequent word families and the rest of the less frequent families (Szudarski 2018). There is also a column showing the number of words included in families, types, and tokens. Another section shows the details of the words produced by the students in the following with different colors indicating the token list.

3.7 Ethical consideration

To conduct this research, ethical consideration been given extra care. A formal letter was given to college management, program director and coordinator of the particular faculty. The letter was given to ask permission to research Windfield International

University of Ma

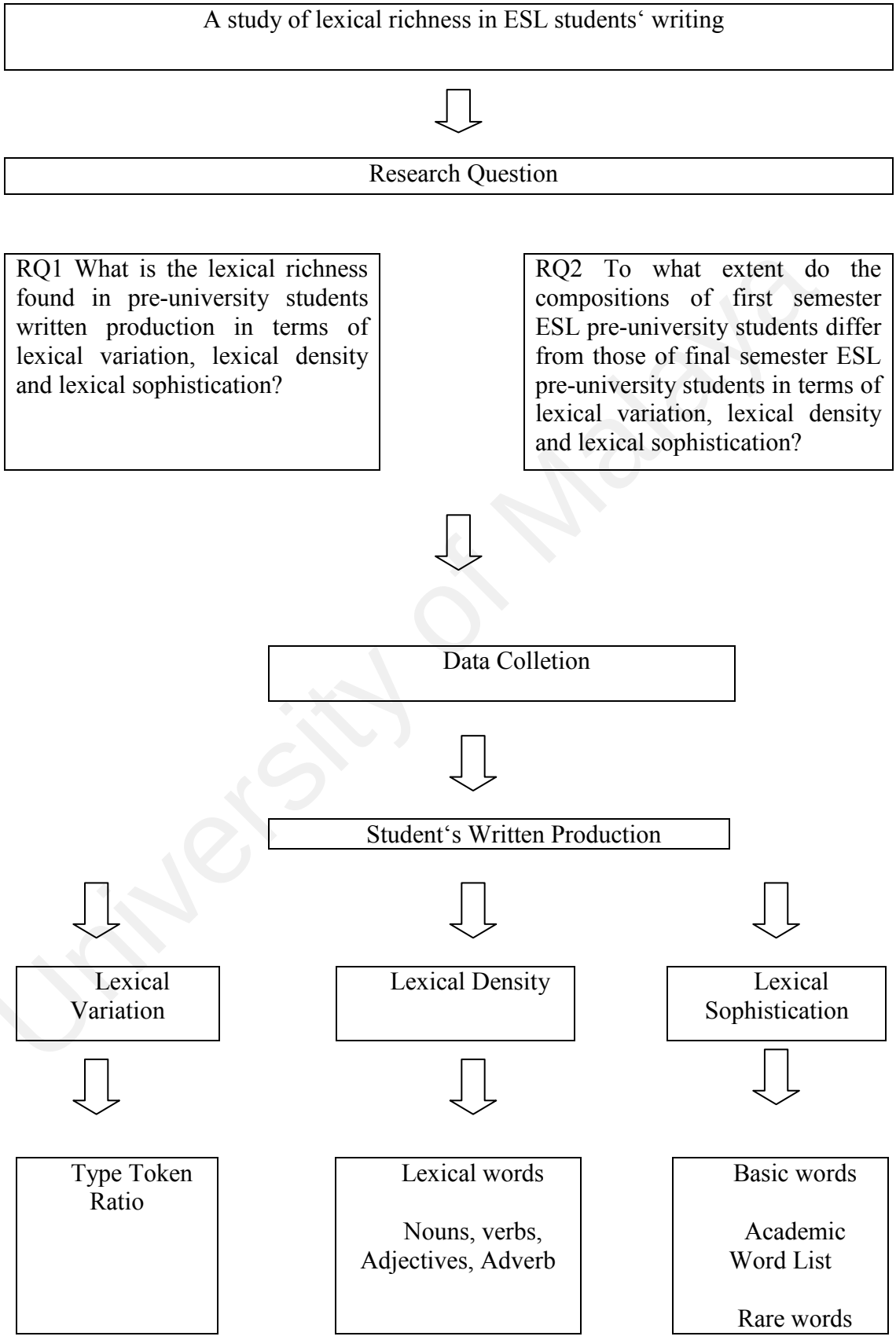
College by accessing their students. Once the college management approve, consent forms given to the participants‘ parents since some of the students are not yet 18 years old to obtain their permission for their children to involve in this research. The withdrawal of participants along the study process allowed as a way to respect participants‘ rights.

Also, the anonymity of the participants will be maintained by assigning members such as Student 001 for first semester students and Student FS 001 for the final semester instead of their real names and assigning pseudonyms for the transcription of written production. Small incentives have been given to all participants in this study as a form of appreciation.

University of Malaya

3.8 Conceptual Framework

The Conceptual Framework below shows a schematic outline of elements in this study.



University of Malaya

Overall, figure 3.7 show that this research carries a title as a study of lexical richness in ESL students‘ writing. This study aims to measure the lexical richness of second language learners‘ writings. In this study, lexical richness divided into 3 aspects such as lexical variation, lexical density, and lexical sophistication. Lexical variation measures a variety of vocabulary in a written text. The variation of words consists of more types compared to tokens. More types show different words in the text. A text is seen as good quality if a learner uses a larger vocabulary by not repeating words. Lexical variation is calculated by the TTR. Next Lexical density calculates content words in students‘ written text. Laufer and Nation (1995), density focuses on lexical words such as nouns, adverbs, verbs, and adjectives. These words mainly convey information. Lexical density calculates a text‘s richness because it shows the content and conveys information clearly. Besides that, lexical sophistication distinguishes the scope of lexical knowledge earned by a second language learner. According to Aviad and Laufer (2013) it is possible to evaluate a piece of written work by looking at the degree of the frequency words level in a text. Sophisticated words show the selection of words used by a student in text. It also shows the word categories in their written text ad help to explore the suitability of vocabulary profile features to capture sophistication. For example, sophistication shows basic words, academic word list and off-list words that produced by students in their written text. Laufer and Nation (1995) claim lexical sophistication is the percentage of words that a learner uses at different vocabulary frequency level. In conclusion, this study intends to measure the lexical richness of Pre- University students to have a clear understanding of lexical richness.

Next follow by data collection using qualitative methods. Written essay on the topic ―The person I admire the most" collected from first and final semester students. Research participants are already familiar with the topic so no problem getting the data from them. Collected data was typed and transcribed into Microsoft word documents

University of Malaya

without changing any grammar or sentence structure. Spelling errors were changed since errors will be then analyzed as advanced words in Lextutor. Collected data were analyzed into Lextutor to obtain statistical analysis on lexical variation, lexical density, and lexical sophistication.

3.9 Conclusion

Overall, this chapter discussed the brief description of participants who participated in this research. Additionally, an explanation of the research tool such as the question and how the text involved in data analysis was provided to those participants. It also clarified the research ethics and consideration of this study.

University of Malaya

CHAPTER 4: ANALYSIS OF DATA AND DISCUSSION OF FINDINGS

4.1 Introduction

This chapter explains the findings of this research. Results analyzed through http:/www.lextutor.ca/.To answer research questions 1 & 2, data analyzed and discussed lexical richness in three main areas; lexical variation, lexical density, and lexical sophistication.

4.2 Results of Lexical Variation

The lexical variation measures types which are known as different words and tokens on the ratio of different words (types) to the total number of words (tokens), as a type-token ratio (TTR) by Johansson (2009).Lexical variation assessed solely in terms of the words that produce in the text by the students. Findings show the lexical resources that the participants have used in the text. Milton & Freeman (1996) claim that learner has good proficiency when they can to produce a variety of words; this is known as a determinant of language proficiency. Results show that Lexical variation is an overview of written performance measurement of ranges of accuracy and fluency. Based on the result, the type token ratio ranges 38% to 55%. Three participants, student 009, 013 and 014 generated a ratio below 40%. The highest ratio was 55% produced by student 011, while student 003 was the second-highest with 54%. The remaining classified as a moderate range from 41% to 49%. This result demonstrates the measurement of lexical variation in written production. The highest variation produced by student 011 with 55% and the lowest variation by two participants, student 013 and student 014 with 38%. The highest TTR produced 153 token and 84 types. The lowest ratio wrote by student 013 with 326 tokens with 125 types. Student 014 wrote 537 tokens with 203

University of Malaya

types. The fact is that the low ratio students produced more types of words, due to the reason that they wrote more. Students wrote unnecessary elaboration out of the topic. The writings should be precise and clear. To support this, Wang (2014) claims writers with small lexical knowledge do not tend to produce precise text, their text seems to be repetitive. Due to repetition, their TTR seems to have a low ratio. Results show the highest variation produced by student 011 from the first semester. The results of lexical variation are shown in the table below

University of Malaya

4.2 Results of Lexical Variation

Table 4.2 Result of Lexical Variation for students in first semester

STUDENT IDENTITY	TYPES	TOKENS	RATIO/PERCENT
Student 001	222	493	45 %
Student 002	246	548	45 %
Student 003	124	229	54 %
Student 004	127	302	42 %
Student 005	185	395	47 %
Student 006	196	403	49 %
Student 007	188	455	41 %
Student 008	125	288	43 %
Student 009	101	261	39 %
Student 010	165	388	43 %
Student 011	84	153	55%
Student 012	109	243	45%
Student 013	125	326	38 %
Student 014	203	537	38 %
Student 015	183	441	41 %

University of Malaya

The writer demonstrated the highest TTR value, who was able to write the information related to the topic. Focusing on a bright site, the participant has vocabulary knowledge to convey the ideas into the writing without repeating the words. Student 011 was able to express ideas and tried his best, but he was unable to fulfill the task. This shows the poor writing standard. However, this cannot be concluded that student 011's text is rich just by looking into TTR. He was able to pass SPM in English. Having D grade indicates that the participant already has low English proficiency as English is the second language. To clarify this, Wang (2014) has shown that writers with small vocabulary knowledge can only generate written text with fewer vocabularies. The writing was repetitive and student 011 has the restricted vocabulary to accomplish the assignment, although TTR indicates the highest ratio. Below is an example. Types have been underlined.

Example Student 011 – highest variation

I *admired her for selfless and her unconditional motherly love for us and everyone that was close to her. She never expected anything for her attitude to be good to everyone always smiling and positive approach to life.*

Next, student 013 produced the lowest TTR. The participant wrote a lengthy introduction to the topic and revealed the admired person in paragraph 3. Although the information is long, the introduction includes the right declaration of a thesis statement. It showed that student 013 did adequate elaboration on the topic, but did not immediately introduce the admired person. There was necessary information related to a description of the person, for example incident and qualities that the participant likes. From the perspective of lexical variation, we can assume that student 013 cannot precisely transmit ideas. The vocabulary seems to be limited and repetitive. Student 013 received C+ in SPM and this demonstrates a moderate language background. However, this participant managed to write 326 words, which is almost up to the requirement of

University of Malaya

an essay although TTR shows the lowest range. TTR is not applicable as the lowest TTR student 013 texts are better than the highest TTR, student 011. Below is the sample text by student 013. Types have been underlined.

Example Student 013- Lowest Variation

In this world they have one person which is they always with you. When you are born until you become a younger person. The person is always sacrifice to do anything for you to see you are always happy. We can't imagine how much sacrifice, care, waste their time to you. In addition, the person that was I mentioned is very hardworking, intelligent, caring, kind person and the most important is she is very beautiful person to me.

Results showed the highest types logged by student 002 with 246 types with 548 tokens. The student was able to write appropriate content using different types of variation. The content has less repetition and variety compared to other texts. However, the student exceeded the number of the word. Within the requirement, the student failed to produce 350–450 words. Although student 002 did not produce text about the length of the text, he managed to produce it using the highest types with different words. This shows that the student has a good lexical variation not only by concentrating on a number of words but also by being able to write with a variety of vocabulary. A proficient learner capable writes a composition using different types of words. A claim by Williamson (2009) that a text with higher types indicates a greater lexical variety. Surprisingly student 002 got C+ in SPM and showed noticeable progress in this writing. This may be due to familiarity and a previous understanding of the topic. Below is an example. Types have been underlined.

Example Student 002- Highest Types

In the conclusion, I would like to say that my dad is my world. He is my inspiration, my role model, my idol, my backbone and everything. I will take care of him as the pupil of my eye. I would like to be a successful and good father like him.

University of Malaya

Besides, second-lowest types produced by student 009. The text below showed frequent repetition words and she was unable to use synonyms to produce the text. However, this is due to the student's limited vocabulary, receives C- in SPM. By having a limited vocabulary she faces difficulty in properly arranging points to fulfill the task. We can accept that the student did not master the essential vocabulary to accomplish the task, although the particular topic is from the English Language A syllabus. The student was unable to compose related to the topic, but a poor compilation of words causes the essay to have the lowest types. It also demonstrates that the student has poor ideas on the subject, and is not ready to give further explanation on the points and take the knowledge into the writing. Below is an example.

Example Student 009- Low Types

The second my admire is are my sister.my sister attitude as assertive for in family. She also always care for family. I like my sister because my sister always follow up on my request and my sister so very sporting. My sister always provide guidance on how to be disciplined, to be a good child.

Apart from focusing on the highest and lowest variation, Table 4.2 shows inconsistent results. The results were divided into three categories. The first student wrote 350 to 450 words as needed. Results show student 005, 006 and 010 produced text-based on requirement. 350 – 450 word limit known as the pre-university standard. These particular students have lexical knowledge to complete a task by adhering to the word limit. We can assume that these students planned their ideas, thoughts, knowledge for the topic and compiled those words into the text with the requirement. They created the text according to the standard of writing.

The second category of the result shows that seven participants did not write according to the requirement. Student 003, 004, 008, 009, 011, 012 and 013 generated text below 350 words. Student 011 produced just 153 words. We cannot expect much

University of Malaya

since student 011 only got D (pass) in SPM. Jarvis (2002), argues that lexical variation has a link with the context of the language. It‘s not surprising that learners with low background knowledge have difficulty producing words in an essay. This particular student only manages to write three paragraphs in total, but the Type Token Ratio shows that student 011 has a higher lexical variety. The Type Token Ratio results depending on the content connected with the sample size. If a text contains too many repetitions of words, TTR will decrease. However, this shows students with a low size of vocabulary does not have appropriate words to explain their ideas in a written context. Without words it is not feasible for a learner to express their ideas, it does not imply that they have an idea it is just that they have difficulty expressing it through words, because the topic given for this research is a familiar topic for the students. They already had classes, assignments; discussions based on the topic and certainly, know what the title of the essay is all about. It is also a free writing genre, so students can produce their understanding according to their lexical knowledge. At the same time, this demonstrates the student has a very low diversity of vocabulary. Below is the example of a paragraph that produced by student 011.

Student 011

As a child I looked at my mother as a super hero. I have always wondered how she took care of me and my sister. Working and being a housewife at the same time.

The third category of results refers to participants (students 001, 002, 007, 014 and 015) that exceeded the word limit. It was clearly stated in the task that the text should contain 350-450 words, yet the students did not produce text according to the requirement. Student 001 writes unnecessary elaboration in text based on the topic. This clearly shows the student‘s ability to write in detail, but not precisely. The main content is one, but there have been no appropriate supporting details. The idea was the same, but the student did not write precisely in the text and there were repeating words. A claim

University of Malaya

by Breeze (2008) declares that based on her studies stated whether the writing is repetitive and lexically adventurous, this can be due to familiarity with the words that she doesn‘t want to try using new words in her text. Furthermore, less proficient writers tend to be satisfied when communication is achieved, and are less concerned with the words. We can assume that students have difficulty writing accurate text by about the length of text or write unnecessary elaboration that affects the variation.

Below is the example of student 001‘s text.

Student 001

My dad cooks, so my mom doesn't. He is a great cook to me. He always prepares my breakfast and dinner everyday. Every morning he will wake up early and cook for us then wake us up to be ready for school. He prepares different kinds of breakfast every single day; he makes sure that we were well fed before school.

Overall, it is difficult to justify the quality of the text produced by first semester students due to the different words in a composition. Additional claims by Nation & Webb (2011) TTR are sensitive to sample size, making it difficult to check the quality of the text. However, the impact of text length on LD measures be avoided as much as possible because they generate construct-irrelevant variances and incorrect results. According to Engber (1995); Malvern and Richards (2002), text with less repetition and more different words is lexically rich. The findings showed students‘ lexical knowledge in their writing performance. Student 001 has an approximate idea on the topic which enables to describe the characteristic of the idol. Lexical variation has been correlated with the student‘s performance in writing. TTR is sensitive to text size, so it is hard to assess the quality of the content. In any case, content with less repetition and a higher number of various words is lexically rich, according to Spencer et al (2015). It shows that the highest TTR value shows the ability of the writer who can

University of Malaya

compose the essay correctly without repeating more than once. The results of the final semester lexical variation are shown in Table 4.3 below.

University of Malaya

4.3 Results of Lexical Variation

Table 4.3 Results of Lexical Variation for students in final semester

STUDENT IDENTITY	TYPES	TOKENS	RATIO/PERCENT
Student FS 001	186	369	50 %
Student FS 002	169	320	53 %
Student FS 003	135	285	47 %
Student FS 004	281	578	49 %
Student FS 005	167	396	42 %
Student FS 006	234	512	46 %
Student FS 007	191	407	47 %
Student FS 008	186	387	48 %
Student FS 009	172	363	47 %
Student FS 010	253	516	49 %
Student FS 011	215	500	43 %
Student FS 012	238	525	45 %
Student FS 013	193	485	40 %
Student FS 014	249	562	44 %
Student FS 015	157	318	49 %

University of Malaya

Table 4.3 shows the Type Token Ratio from the lowest 40% to the highest 53%. The outcome of this lexical variation is based on the sample size produced by the final semester students. However, all of the participants‘ text length is different. Results indicate different TTR shows the lexical knowledge of the particular students. Overall, three participants (student FS 002, FS 003 and FS 015) write below 350 words indicating that the written text was not up to the requirement. Surprisingly, student FS 002 and FS 003 got A and A- in SPM, while student FS 015 got B in SPM. These students have a great English background. However, they were unable to compose up to the writing standard yet there were not many variations in the complete amount of words. Student FS 002 writes 320 tokens, student FS 003 with 285 tokens and student FS with 318 tokens.

Next, only five participants (student FS 001, FS 005, FS 007, FS 008 and FS 009) manage to compose as required by the essay. The result indicates student FS 001‘s as second highest TTR with 50 % of TTR. Data shows student FS 001 has a variety of words to produce the text. Johansson (2008) claims, greater lexical diversity demonstrates a high variety of vocabulary in a text. Data shows student FS 001 write using different types in the text. The text has many different words and fewer repetitions that have already been used. We can say that student FS 001 has good lexical in the store because he can express ideas precisely in a text. The text was according to the writing standard. To add on to this, student FS 001 has A in SPM, which shows that he already has prior knowledge and, since he is in the final semester, he must also have writing experiences.

Apart from that, seven participants (student FS 004, FS 006, FS010, FS 011, FS 012, FS 013, and FS 014) exceeded the word limits. Results show that there were quite a number of differences with the additional tokens like 35 words to 128 words. However, the students attempted to do their best to write the text. According to Lintunen, Peltola

University of Malaya

&Varila (2014) claims that learners tend to create longer texts because of their writing experiences. Since these students have produced longer text, we can claim that they have writing experiences which makes them express ideas based on the topic. Particularly these students have B grade in SPM except student FS 012 (grade C). It shows that the students might have good background knowledge in English where they can compile the words and create them into a text. Additionally, we can also argue that students write using a variety of words because of their familiarity with the topic and their understanding of vocabulary throughout the course. A claim by Koizumi & Innami (2012) vocabulary use related to a comprehensive understanding of vocabulary knowledge. Without knowing a word the participants would not be able to express their views in the text. Another insight that can be drawn from this result is, some participants wrote lengthy elaboration on points which could be a reason for this. Lexical variation referred to as ~~r~~ange of expression". According to Crossley & Louwerse (2007) the longer text is usually known as lexically less diverse because participants elaborate their points as more words are introduced to a text. However, TTR greatly influenced by the sample size.

The students were acquainted with the topic so they had plenty of ideas shared in their writing and eventually their responses are well concentrated and related to the topic. Results show that these students are final semester students, so they already had enough instruction and practice to fulfill the task, unlike fresh first semester students. However, this is just an assumption.

The highest TTR logged by student FS 002 with 53%, 320 tokens, and 169 types. The highest TTR has distinct lexical knowledge and has provided content with less word repetition. Student FS 002 can write with limited tokens but varies with types. The idea was concise and precise with a limited quantity of words. This indicates the student

University of Malaya

has a good diversity of lexical knowledge. Additionally, student FS 002 got an A in English SPM shows; he has a high proficiency level compared to other students. We can claim that student FS 002, may have a lot of writing experience or good vocabulary knowledge through reading books. He composes his text with a variety of types that make the essay up to the standard. His writing output shows that he can deliver the content and compile it into text by using different types in a sentence. The example is below. Types have been underlined.

Student FS 002 – Highest TTR

Her name is Dhiya and she is the love of my life. The whole neighborhood knows not to mess with her as she has big brother to deal with if anything happens to her. Despite that, she has now grown up with a face of fathers but hair like mothers. Smartest, gorgeous and wisdom girl I have ever known in my entire life.

Student FS 013 generated the lowest TTR of 485 tokens and 193 types. Student 013 has a short idea of the topic, but the ideas were not written focusing on the text length. A statement by Read (2000) a student with low vocabulary knowledge will use more repetition of words; eventually increase the tokens, not the types. Student FS 013 introduces new points with new words that lead to repetition. Simultaneously, word repetition decreases the type and generates low TTR. The two lowest TTR of the student FS 013 and FS 005 show repetitive and lexically inattentive writing. Below is an example.

Student 013 – Lowest TTR

First to all, Datuk Lee Chong Wei is not a quitter when he was young he said he wanted to become an athlete by playing badminton for the nation .His passion for badminton grew day by day as he wanted the athletes playing in matches on televisyen.

University of Malaya

According to Mwinlaaru & Xuan (2016) it is sensible to think that less skilled L2 writers, such as students in these studies attain coherence by repeating significant vocabulary, rather than using a variety of pronouns or synonyms, or by deploying a broader variety of syntactic structures. This obviously indicates a student‘s ability to checkpoints of interest, but not precisely. Moreover, when the composition is achieved, less proficient writers tend to achieved and less worried about words. Breeze (2008) further clarified this is because L2writers prefer to use the vocabulary they recognize and examine and not venture into new scope. This clearly indicates the capacity of the student to write in detail by offering appropriate explanations and examples; but not exactly about the word limits.

To answer research question 2, the results of the first semester and final semester students were compared. The findings indicate no significant difference between the first semester and the final semester. The idea was an expectation to find out significant differences in the final semester by looking at two different semesters. This is because the final semester students have already had already passed 2 semesters of English courses, while the first semester students have two more semesters to go. Below is the example of comparison between first and final semester students.

Highest TTR – First semester (Student 011)	Highest TTR -Final semester (student FS 002)
He loves gardening and exercising. In spite of his age, he stills go for jungle trekking and jogging every week. He spend his day in a productive way by maintaining the condition of the farm in a good shape and condition. He will get trees or plants to be planted in our back yard. He takes care of them very well.	As soon the doctor present the infant, a sense of purity and innocence clouds the mind. Nothing but love sets between me and her. The negativity dissipates in my memory and from then on, my life has changed since then.

University of Malaya

Based on the example above, student 011 has 55% while student FS 002 has 53%. Surprisingly, student 011 received D in SPM and managed to produce the highest TTR in a text, while student FS 002 received A in SPM. Both students have different levels of English proficiency and have logged the highest TTR. Additionally, student 011 is a first-semester student who is still new to the course and studying English. Besides student, FS 002 is already acquainted with the English course in the final semester. The TTR indicates only 2 % differences between the first and final semesters, but we can claim this as significant differences in the highest lexical variation between first and final semesters due to their language background. Next is the lowest variation produced by student 013 and 014 with 38% and student FS 013 with 40%. The TTR shows minimal differences between the first and final semesters. It seems that the lowest variation from student 013 and student 014 is acceptable because they have moderate results in SPM that are C+ and D. We can claim that these students have difficulty building their ideas and writing them concerning text length. By writing a lot of elaboration and failing to produce precise text, student 014 produced text with a total of 537 tokens. Student FS 013 produced text with 485 tokens. Surprisingly student FS 013 got A- in SPM. Looking into student FS 013 text, it seems that the content is related to the topic and there are not many additional tokens produced with the requirement. Since the length of the text is not the same, we cannot conclude the variety at this point.

Lowest TTR – First semester student 014	Lowest TTR – Final semester student FS 013
Moreover, she also the one that very particular about their children health even we already teenage and adult. She will take care and always remind us to take medicine or go to hospital to get a treatment.	In a nutshell, Datuk Lee Chong Wei is a person that I admire the most because of all his good values in him. Not only me but many people take him as a role model to start a new life. I hope that one day when I become successful in life someone would admire me

University of Malaya

Overall, the results demonstrated a positive relationship between lexical knowledge and students‘ writing. Lexical variation has corresponded to the student‘s performance in writing.The TTR, however, was unbalanced for short texts influenced by long contrasts. The type token ratio for the first semester demonstrated an inconsistency change from the lowest 38% to the highest 55% where, in the final semester, the percentage showed from 40% to 53%, highest TTR created by student FS 002, the token delivered was 320 and types 169.Nevertheless, there was not much critical difference between lexical variation between the first and final semesters. In any case, slight contrasts are seen throughout the final semester contrasted with the first semester writings of the students. This is linked to the students‘ English proficiency level and their academic performance in SPM. We can ensure that the student conveyed the meaning and expressed the title of the essay about to the length of the text.

Kennedy (2014); Read,(2013); Ngangbam (2016), expressed that text with less repetition and higher extent of various words is lexically rich. Student 011 has the highest score with 153 tokens and 84 types for the first semester. This demonstrates that student 011 has the highest variation in contrast with other students. As more words are created in a text, it turns out that the text is longer and new tokens/words are reduced. This will lead to a drop in TTR. TTR cannot be used for texts of different lengths.

According to Hedgcock & Ferris (2013)the lexically poor language will include fewer different words with more repetition of those words. Beaugrande (2014)said knowing a word involves something other than recognizing the words and producing them in writing. It is essential to know how the words used to convey a good idea depending on the context setting. Based on data collection, students created a small proportion of lexical variation. In any case, students have done their best to describe their views according to their vocabulary dimension and also to express their thoughts.

University of Malaya

A claim by Wang (2014) lexical variation only relates only to the range of vocabulary in a text that demonstrates the breadth of the writer‘s vocabulary. We cannot conclude students‘ lexical knowledge by merely focusing on vocabulary size. Lexical variation can be an element of student lexical knowledge to evaluate the variety of their writing vocabulary. This study not only focuses on lexical variation to evaluate the lexical richness of the written text but also on other types of lexical density and lexical sophistication. Apart from that, students‘ lexical knowledge cannot be concluded by looking at lexical variation, the density of words should also be included.

4.4 Lexical Density

Lexical density refers to content words and their percentage of lexical words in a text such as nouns, verbs, adjectives, and adverbs to the total number of words in a text. A popular 'minor variant' is to calculate the density of the noun, the number of nouns divided by the total number of tokens in the text. Introducing the notion of lexical density, Ure (1971) distinguishes between words with lexical properties, and those without them. According to Ure (1971), items that do not have lexical characteristics defined as "purely in terms of grammar" (p. 445), meaning that such words have a more grammatical-syntactic function than the lexical items. The table below shows the results of the lexical density of first semester students.

University of Malaya

4.4 Results of Lexical Density

Table 4.4 Result of Lexical Density for students in first semester

Students Identity	Content words	Total words	Density percentage %
Student 001	207	493	42%
Student 002	242	548	44%
Student 003	103	229	45%
Student 004	119	302	39%
Student 005	209	395	53%
Student 006	192	463	48%
Student 007	222	455	49%
Student 008	124	288	43%
Student 009	113	261	43%
Student 010	176	388	45%
Student 011	59	153	39%
Student 012	106	243	44%
Student 013	127	326	39%
Student 014	228	537	42%
Student 015	190	441	43%

University of Malaya

Table 4.4 shows that the lexical density found based on the data ranges from 39%to 53%. All participants in this research are second language learners. Overall the highest density produced by student 005 with 209 lexical words and 186 function words. The density percentages indicate words of nouns, adjectives, adverb and verb.

Student 005 created a text with high content words that demonstrated high density. Student 005 has B- in SPM, which shows he already has a good background in English. Results show student 005 got 53% density with 129 content words and 186 function words. Student 005used more content words than function words. The idea was brief by listing the characteristics and inspiration the student had on his idol. Student 005 described idol‘s life including interesting facts and details of the idol. Student 005‘s text shows that the participant is knowledgeable on the topic and the text was informative.

A text which is higher in content words conveys the meaning or core content of an idea. Johansson (2008) stated that the higher the content words, the higher the lexical density. Additionally, when the density is high, the reader can understand and grasp the meaning. For example, we can disregard all function words in the text and just focus on content words, yet we can understand the writer‘s idea. We can argue that student 005 should have good lexical knowledge that he can create a text by using more content words that convey the meaning. However, Leedham (2014) indicated that a high percentage of lexical density with 50 percent above no doubt composed of a first language learner and great writing has a high percentage of lexical density. To et al. (2013) said a text in an English textbook with high lexical density can be easy to read. Lexical density plays a major role in evaluating vocabulary knowledge. The large majority of the written texts have 40% or higher lexical density (Johansson, 2008). The matter of lexical considered having nouns, verbs, and adjectives to show the content of a text.

University of Malaya

Result shows (Table 4.4) that student 005 successfully created density text above 50 percent, and this suggests the student's lexical knowledge. This is due to the familiarity of student 005 on the topic. Although student 005 practices English as a second language, he manages to write quality text. This shows that he should practice more in English and enhance his vocabulary knowledge. Below is the example of student 005.

Student 005 – Highest density

*The **person** that I **admire** the most is a **well known scientist** and a **legendary character** in **physics career.** He is an **extraordinary scientist** because he had **discovered** a lot of **physics theories** and he was a **noble prize winner**. He was **born** from a **poor family background** and he has been **suspected** to be **mentally retarded** at his **young age**. But he had **proof himself** after **years** of **hard work.***

The example above shows the second-highest lexical density with 49% produced by student 007. Student 007 writes 147 content words and 233 function words. There aren't many differences between student 005 and student 007. Student 007 got C- in SPM, however, and this shows progress in her writing where she able to produce the second-highest lexical density with 49%. Student 007 tries to closely express her ideas using content words. The ideas were clear with less use of function words. Below is an example.

Student 007 – Second highest of lexical density

*I really **admire** my **mother**. She is **beautiful girl** with **intelligent brain**. Her **childhood memories** is not **like other kids**. She **used** to **live** in a **poor condition** when she in her **16 years** old. She is **responsible** to **work** to **help** her **parent** to **maintaining** their **financial.** She is **strongwoman** who can **endure** all the **pain** and **keep trying** to be a **good daughter** to her **parent**.*

Apart from that, the findings show the lowest density generated by three students with a similar density of 39 %. Student 004 logged 119 content words and 183 function words, student 011 logged 59 content words and 94 function words, student 013 with

University of Malaya

127 content and 199 function words in their text. Student 004 produced 183 function words and 91 content words. Student 004 has associated points but unable to convey them by using more content words. She used so many functional words such as pronouns, conjunctions, and determiners. These functional words decrease the text's density. Student 004 has the only C in SPM. It shows that she has moderate language proficiency. However, based on the above example, we can claim that she is attempting to convey her ideas by providing too much elaboration using function words. Johansson (2008) stated that text shows minimal information with a large proportion of function words such as prepositions, determiners, pronouns, and conjunction. She has difficulty writing her ideas into the content. If she reduces the use of function words and focuses more on writing content words, she will be able to create a quality text with lexical properties. Student 004 needs to work with vocabulary knowledge to create text with more specific meanings because it is important to create a successful text because it shows the standard of writing. Below is the example of the lowest density by student 011. The following table shows examples of the lowest density.

Student 004 – Lowest Density

My ***mom always cook*** *my* ***favorite foods*** *and my* ***siblings favorite****. My mom is* ***fully housewife*** *and she* ***never complained*** *to* ***do*** *all the* ***homework*** *to their* ***child*** *or to my* ***father****. My* ***mom always responsible*** *what she* ***do*** *and* ***always responsible*** *to my* ***siblings*** *and my* ***father*** *for our* ***foods****, our* ***clothes****,* ***take care*** *of my* ***younger brother*** *at* ***home*** *and* ***everyday before*** *we* ***want*** *go to* ***school*** *or* ***work*** *my* ***mother*** *will* ***prepare*** *the* ***breakfast*** *in* ***every morning****.*

Student 011 also produced the lowest density of 39%. Student 011 was only able to produce 39 content words and 94 function words in his text. His essay was not up to the writing requirement and he only wrote a total of 153 words. This shows that he is unable to fulfill this task. Besides, student 011 only got D in SPM. Due to a lack of vocabulary knowledge, he could not use content words to create his writing more

University of Malaya

adequate. Apart from that, student 011 claims to be the highest lexical variation among the rest. Focusing on lexical variation, the findings show that student 011 generated the highest variation. We cannot conclude this statement because TTR is sensitive to sample size. But when concentrating on lexical density, it shows that student 011 is unable to write a quality text using lexical properties. The example below shows that student 011 expressed her ideas, but that's not enough for writing. She should use more content words to convey the information in a text. Below is an example.

Student 011 – Lowest density

I am ***lucky*** *to* ***have*** *her as a* ***mother****. She is my* ***hero*** *and* ***role model****. I* ***appreciate everything*** *she* ***ever*** *did and did not* ***either****. I will* ***always owe*** *you in my* ***life*** *and* ***success*** *in* ***everything*** *I* ***do****.*

Student 013 has an idea about the individual she admires but didnot write it using content words. Lack of content words confuses the reader in terms of understanding the original idea of what the writer is attempting to convey. Similarly, this shows low text quality, with low density generally creating text with an excessive number of word repetitions and not being able to produce text precisely. This can be due to background knowledge of student 013, where she only got C+ in SPM. She might be struggling to write her ideas using content words. That's why she unnecessarily used a lot of function words to convey her ideas. The next example shows a part of a text written by student 013.

Student 013 – Lowest Density

Finally*, the* ***person*** *the* ***most*** *I* ***admire*** *is my* ***mother****, because she is* ***like angle*** *in my* ***family. We*** *all* ***love*** *him so much, because of she I am in this* ***world*** *and I am here to be a* ***good person.***

University of Malaya

Furthermore, the outcomes show that student 002 produced the highest content words with 181 content words in his text. Student 002 got C+ in SPM. This shows that student 002 has moderate English proficiency yet he manages to produce the highest density words in his text. He used a lot of content words to describe his idol as inspiration, backbone, successful, learned, and positive and many more. These keywords expressed his ideas on the subject. A claim by Zafar (2016), lexical density is capable of concluding and distinguishing a composed text by showing the quality-centered around content words. Student 002 produced quality content according to the writing standard. Below is the student 002‘ example.

Student 002 - Highest content words

He is my ***inspiration,*** *my* ***role model****, my* ***idol****, my backbone and* ***everything****. I will* ***take care*** *of him as the* ***pupil*** *of my* ***eye****. I would* ***like*** *to be a* ***successful*** *and* ***good father like*** *him. I have* ***learned*** *many* ***good things*** *from him and also* ***positive thoughts****.*

Whereas the lowest content words were produced by student 011 with only 39 content words among the rest. This shows poor performance because he wasn‘t even able to write according to the essay requirement. The total words he wrote were only 94. Lexical variation analysis shows that student 011 logged the highest TTR. However, TTR is sensitive to sample size. Student 011 only manages to pass in SPM. The result indicates student 011 can write ideas linked to the topic. However, he cannot convey his ideas in detail by writing examples and experiences to support his point of view. Statistically, it shows poor performance by student 011, but when analyzing his text it shows that he has some vocabulary knowledge to write his ideas. We can claim that student 011 requires more English writing practices to improve his writing skills. He might be a slow learner because he regarded English as a second language. Additional

University of Malaya

instructions, assessments, and vocabulary will help him improve his proficiency level where he will be helped to write with more content words. Below is an example.

Student 011 – Lowest content words

I am ***lucky*** *to* ***have*** *her as a* ***mother****. She is my* ***hero*** *and* ***role model****. I* ***appreciate everything*** *she* ***ever*** *did and did not* ***either****. I will* ***always owe*** *you in my* ***life*** *and* ***success*** *in* ***everything*** *I* ***do****.*

Based on Table 4.4, lexical density shows how rich text interrelated with content words. By researching this, we receive information packaging in a composition. It gives us how lexical words are a significant component of a piece of text. Johansson (2008) claims that a text with a high proportion of content words has more information than a text with a high proportion of function words (prepositions, interjections, pronouns, conjunctions and count words).A piece of text can be judged by analyzing the lexical density in a text.

Apart from that, Table 4.4 shows that student 002 has the highest content words that have been logged among the rest. Even though student 002 got C+ in SPM, he was able to write the highest density in his text. However, it is not possible to judge that this is the best among the rest, as all the students have a different total number of words in their composition. The lowest content words produced by student 011 with 59 together with 94 function words. This results in the lowest density. Since the length of the text is short, we cannot expect more from the student. This can be due to student 011 English background knowledge. Student 011 only passes in SPM and this shows that he is already struggling to accomplish in English. Next, the highest function words logged with 309 and 228 content words by student 014. By focusing on the number of function words, we can conclude that student 014 performed poorly among first semester students. To support more, student 014 only passes in SPM. This shows that she has

University of Malaya

difficulty writing text where she used a lot of repetition by using function words. Student 014 performed well when the percentage correlated with content and function words.

Besides that, Ure and Ure & Ellis (1997), emphasized the significance of lexical when analyzing the notion of lexical density. Traditionally, nouns, verbs, and adjectives are the three-word classes that are considered to have lexical properties. Lexical words looked at the basic building of a language. Without content vocabulary, the text will not have any specific information, context and ideas expressed. Olinghouse& Wilson (2013) concludes that a large majority of the spoken texts have a lexical density of less than 40%, while written texts have a lexical density of 40% or higher. The degree of content words in composition features prominence in evaluating a piece of work.

University of Malaya

4.5 Results of Lexical Density

Table 4.5 Results of Lexical Density for students in final semester.

Students Identity	Content words	Total words	Density Percentage%
Student FS 001	174	369	47%
Student FS 002	133	320	42%
Student FS 003	134	285	47%
Student FS 004	277	578	48%
Student FS 005	152	396	38%
Student FS 006	265	512	52%
Student FS 007	174	407	43%
Student FS 008	182	387	47%
Student FS 009	162	363	45%
Student FS 010	281	516	54%
Student FS 011	209	500	42%
Student FS 012	227	525	43%
Student FS 013	219	485	45%
Student FS 014	251	562	45%
Student FS 015	137	318	43%

University of Malaya

Other than the first semester, Table 4.5 shows that the final semester results demonstrated satisfactory results range from 38% to 54%. These findings correlate with content words and function words produced by students in their composition. The highest density produced by student FS 010. The lowest density produced by student FS 005 when examining the writing performance in this study. According to Signes &Arroitia (2015) a higher density demonstrates that the information set forward is concentrated. The example above shows student FS 010 managed to create precise text with higher content words compare to function words. The content words logged by student 005 show how concentrated a text is according to the topic. Apart from that, by using verbs student 005 successfully highlights the main part of the topic. A reader is still able to understand the content by simply reading the lexical words omitting functional words from the text. Student FS 010 effectively presents ideas without elaborating too much and demands the reader's attention. Density words convey the meaning of the text. The example below shows the second highest density in the final semester.

Student FS 010 – Highest Density

Elon Musk *as a* ***whole*** *has* ***helped*** *me* ***realized*** *that* ***true scale*** *of* ***self improvement*** *and has* ***helped*** *me* ***build up*** *my****self confidence*** *to* ***chase*** *my* ***dreams*** *with every bit of* ***hard work***, ***determination*** *and* ***energy*** *that I have. His* ***achievement*** *has* ***trumped all previous technological revolutionists*** *by* ***miles*** *and* ***helped*** *the* ***world realize*** *that there is more than* ***meets*** *the* ***eye*** *when it* ***comes*** *to* ***achieving*** *such* ***feats****.*

Student FS 006 manages to produce the second-highest density with 180 content words and 247 function words. There were only 2% differences between student FS 010 and student FS 006. There are no enormous differences we can claim in this matter. However, student FS 006 got B + in SPM which indicates that she already has good English proficiency. Student FS 006 can write her ideas using more content words, resulting in a higher density among other participants. The example above shows that

University of Malaya

she can describe her idol by using verbs and adjectives. Verbs show actions and adjectives modify the set of words that make the text fine. There are very few functional words that she used to describe her idol. Below is an example.

Student FS 006 – Second highest lexical density

Ellen is also **known** for **being** a **modest** and **funny person**. This is because; I **believe** that **humor plays** an **important role** in **everyone's life**. She **always comes up** with **jokes** and **stories** to **get everyone's attention**.

Student FS 005 produced the lowest density with only 38% in the final semester. This is the lowest density since the remaining density varies from 42% and above. The student wrote an unnecessary elaboration in the introduction. Student FS 005 could have included more nouns, verbs to show the feature of the topic rather than using repetitive functional words. According to writing, functional words only indicate the structural features of the sentences. This makes the text look low in density and affect the quality of a text. This also causes a lack of cohesion in writing. Liu (2000) claims that content lexical connections play a major role in the cohesion of writing. Student FS 005 produced incoherent ideas and insufficient functional words. The lack of cohesion results in low density in the text. Student FS 005 should improve using more lexical words such as, nouns, adjectives, verb, and adverbs to strengthen text quality. Student FS 005 could describe mom's character by writing proverbs, quotes and describe the personalities using lexical items. For example, my mom is my source of inspiration, she is kind, simple, and kind-hearted person. Below is an example.

Student FS 005 –Lowest Density

Every each of us is sure to have someone that we admire the most. I had started to admire this pretty woman since I was born, she is my one and only heroine at all time. The person I admire the most in my life was my lovely mother.

University of Malaya

Student FS 014 produced 211 content words and 311 function words resulting in 45% of lexical density. Student FS 014 got a B in SPM. The example above shows how structured her ideas are. She can describe the qualities of her idol and what makes her admire the person using content words. Content words convey the information in the text. For example, admire, strongest, highly, educated, intelligent, stern, decision and etc. She is able to use verbs, adverb, and adjectives in order to convey her reason for liking her idol. The next example shows the highest content words produced by student FS 014.

StudentFS 014 – Highest content words

The reason I admire my mother the most its because she is the strongest woman..My mom went through life more than I can ever imagine. My mom is not highly educated but she is very intelligent... still stern with my decision and I have confident in myself, I can succeed in anything with my own hard work.

It is quite surprising that student FS 002 produced the lowest content words, among the other participants. He manages to write 92 content words and 187 function words. Nevertheless, we cannot claim this as a significant result in the final semester because the length of the text is not the same. At the same time, lexical variation analysis shows that student FS 002 produced the highest TTR. The example above shows that student FS 002 has impressive vocabulary knowledge because his sense of idea is quite clearly related to the topic. Additionally, he got an A in SPM where we can claim that he has excellent English proficiency. Even though he produced the lowest content words but his lexical density percentage shows he able to log 42%. The next example shows the lowest content words produced by student FS 002.

Student FS 002 – Lowest content words

As soon the doctor present the infant, a sense of purity and innocence clouds the mind. Nothing but love sets between me and her. The negativity dissipates in my memory and from then on, my life has changed since then.

University of Malaya

To answer research question 2, a brief comparison has been made. The findings of the first semester students and the final semester students are quite similar. The range between both semesters starts from 38% to 54%. The lowest density produced by three students, student 004, 011 and 014 from the first semester with 38% and student FS 005 with 38% in the final semester. There is no difference between the first and final semester students‘ results. The highest 53% is from the first semester and 54% from the final semester student. Unfortunately, there are no huge differences between the two semesters. Therefore, it is not possible to conclude whether first or final semester students have higher lexical density compared to other semesters.

Whereas the highest content words produced by the final semester student FS 014 compared to the first semester student 002 with 242. But results show that student 002 has a higher lexical density compared to student FS 014. So we cannot prove a text by just looking at the content words, because lexical density correlates with function words and the total number of tokens in a text. From a different perspective, Gilquin &Paquot (2008) concluded that lexical density above 50% was considered a text is written by the first language learner. Based on results shows student 005 and student FS 006 and student FS 010 successfully produced a higher lexical density with 50% above. We can assume that the particular students have a good sense of writing or expressing cohesive ideas in a text.

Apart from the highest and lowest density, all participants manage to fulfill the task and everyone have different length of text. Since their vocabulary size is not the same with all students, however they able to write according to their vocabulary knowledge. Lexical density considered a major part of distinguishing the higher density of a text. A text with higher density has more concentrated ideas compared to low density.

University of Malaya

It also demonstrates how lexical density plays an important role in an academic text. By analyzing lexical density, it is possible to analyze second language learners text. However, there were no significant differences between the two semesters. But focusing on the bright side, we can clearly see how lexical items show significant progress and result in student‘s writing. Overall, three students produced more than 50 %, which shows that the text is moderately high. This study claims that lexical density is definable according to the written output of the students.

4.6 Lexical Sophistication

Results analyzed through Vocabulary Profile Program by (Cobb, 2002), calculated the percentages for the word frequency level. There are four levels of vocabulary frequency. The first level is considered to be the one thousand most common English words. The second level is the next thousand most common words, 1001 – 2000 thousand words. The third level is on the pre-computed vocabulary list, such as the Academic Word List (AWL) with Coxhead‘s (1998), associated word frequency distribution. The last level is Off list words, also known as the rareness of words in a text.

Based on Table 4.6 results shown according to the level of the vocabulary frequency band which is the first column shows the first 1000 words, the second column on 2000 words and third column on academic word list and the fourth column on off list words. According to Nation & Webb (2011) sophisticated words are known as the proportion of low frequency or advanced words in a text. The degree of lexical sophistication is measured by low frequency words. Low frequency words recognized as very difficult and more sophisticated words in a text. According to Lindqvist et al. (2013) lexical sophistication is defined as the percentage of sophisticated or advanced words in a text.

University of Malaya

However, there is a various definition of sophistication by past researchers. It measures the level of vocabulary.

University of Malaya

4.6 Results of Lexical Sophistication

Table 4.6 Result of Lexical Sophistication by Students First Semester

Students Identity	First 1000 words %	Second 1000 words %	Academic word list %	Off list %
Student 001	90.4	5.48	0.81	3.25
Student 002	88.87	4.56	1.82	4.74
Student 003	88.21	2.62	2.62	6.55
Student 004	90.73	6.29	0.33	2.65
Student 005	79.75	3.29	3.29	13.67
Student 006	86.35	8.44	0.25	4.96
Student 007	83.52	5.49	3.74	7.25
Student 008	90.62	5.21	0.35	3.82
Student 009	84.29	5.75	0.38	9.58
Student 010	88.40	6.19	1.55	3.87
Student 011	86.93	3.27	3.92	5.88
Student 012	89.71	4.94	1.65	3.70
Student 013	91.41	3.99	1.23	3.37
Student 014	89.20	6.15	2.42	2.23
Student 015	88.89	4.76	0.45	5.90

University of Malaya

Table 4.6 result shows that the participants manage to compose text with different frequency word lists. Results were analyzed through VocabProfile and it shows the participants‘ vocabulary response in the text. Table 4.5 shows the First 1000 words, the Second 1000 words, Academic word list and off list words which comprise a total of 100 percent. Overall results vary from 79% to 92% for the First 1000 words. This range shows good vocabulary knowledge among students because of K1 words known as an indicator of vocabulary. A good writer can be assessed not only by learning vocabulary but also by writing familiar words into text showing the vocabulary already known to the participants.

The highest K1 words written by student 013 with 91.41%. This shows good vocabulary knowledge that the student has. This also indicates strong relationship between the proficiency levels of the learner. Besides that, student 013 has good vocabulary knowledge where she can produce the first 1000 common words in her writing. Results show that student 013 produced 3.99 % of 2K words and 1% of AWL and 3% of rare words. This is due to English exposure. However, the particular student has good prior knowledge of the topic. This can also be a good vocabulary development for the student because she only got a C in SPM. She still uses the highest frequency K1 words in her text. We can claim that student 013 has a good storage of vocabulary through language skills and applies it in writing. Below is an example.

Student 013 – K1 words

In this world they have one person which is they always with you. When you are born until you become a younger person. The person is always sacrifice *to do anything for you to see you are always happy.*

The second highest K1 words produced by three students; the range was 90%. Student 001 got 90.4%, student 004 got 90.73% and student 008 got 90.62%. There

University of Malaya

were not many differences between the participants. However, the words they used show how well they know the word. The total word produced by student 001 was 493, which shows that she exceeded the writing requirement. Whereas student 004, and student 008 did not write up to the requirement. However, these participants were able to write the most 1000 common words in their text. This also indicates that they have good exposure to vocabulary. Laufer (1995) stated that learners should aware of the number of words used because they only include well-known words in their text. This indicates detailed compositions and how they differ significantly at a different level. Participants used most 1000 common words to express their ideas related to the topic. Student 001 got A in SPM, so she able to write her content by adding basic words into her text. However, student 004 and student 008 got C which shows that they have moderate language proficiency. But based on the results, it shows how they able to produce basic words in their text. We can assume that these participants learned new vocabulary through lectures, reading, writing and language input. Below is the example of K1 words that the participants used.

Student 001	person, she , beautiful, father, mother, world, anyone, loving, support
Student 004	favorite, fully, together, caring, concerned, children, respectful
Student 008	family, best, getting, married, morning, prepare, she, is , everything

Furthermore, results indicate that almost every first semester participant has logged 90% and above for basic 2000 words. Only two participants logged below 90%, which is student 005 with 83% and student 007 with 89%. This result shows the range of word frequency based on the text. This related to learners‘ lexical size in the results of writing. According to Nation and Waring (1997) a learner acquires 80% of vocabulary knowledge if they are able to write 2000 common words in their text. Results indicate that all the first semester participants manage to produce above 80% of the most

University of Malaya

frequent 2000 words. This can be an indication that first semester students have good vocabulary knowledge. This might vary for a different written document. Besides that, Hirsch and Nation (1992) claim that 95% of the vocabulary is sufficient for a basic understanding of a text. The participants‘ familiar with the frequent words and this indicates that they know the most words in the highest frequency band. Results show that students can write the most common words in English even though they are second language learners. Additionally, Daller et al. (2007) claims three dimensions of vocabulary knowledge. The first dimension is about the vocabulary size, that learner simply using the word without considering how well they understand the word. A learner uses too many words to describe the situation. Data showed that student 006 used off list words to express her idea. The example below shows, student 006 repeating the same content and the information is not precise. She tried to conclude her statement however it was not relatable. Below is an example.

Student 006

She is the best mother for me in this world. She gave birth to me. After study of science, I know that giving birth need a lot of braveness, perseverance and love. I am proud of my mother who gave birth to four children. I love her forever. She will always be that person that I admire the most.

The second dimension is on lexical depth, which shows that a learner should know the words very well to use them in writing. Student 002 used more sophisticated words to show that she has vocabulary exposure to use the words in a text. The example below shows that he can use common words to explain his admired person and express his feelings towards them. He produced 93% of K1 and K2 words in total. Student 013 has expertise with the meaning that he knows what to write in a text.

University of Malaya

Student 002

I would like to say that my dad is my world. He is my inspiration, my role model, my idol, my backbone and everything. I will take care of him as the pupil of my eye. I would like to be a successful and good father like him.

The third dimension is known as fluency. This reflects on learners‘ ability to retrieve the form or meaning from memory and use it correctly in a text. Breeze (2008) claims that it is possible to analyze the range of the vocabulary by looking at 2000 most frequent word families in English. Types and tokens are included in 2000 common words in the text. Advanced level of students is generally expected to have acquired a minimum active 2000 words of English vocabulary. The result shows that all participants produced 83 % above for K1 and K2 words indicating that they can easily produce an academic text. Student 001, 003, and 015 only got an A for SPM. This is not surprising that they are considered as an advanced level student. But most students got C and D in their SPM. These participants are first semester students with ongoing instructions yet they able to write the text using 2000 most common words in English. They can use academic vocabulary to fulfill the task. A claim by Usman &Abdullahi (2018) these categories of four levels represent the frequency of words which indicate the use of vocabulary as the word list as a predictor of vocabulary knowledge. Below is the example of student 010. Example shows a part of text written by student 010 that she only used K1 and K2 words without AWL and off list words. This is quite surprising because the student does not have an excellent English proficiency. Student 010 got C- but somehow manages to fulfill the task. Another example shows student 011‘text. Student 011 only got D in SPM. Example shows he only wrote 2 off list words (underlined) in a part of the text. This shows an advancement of his vocabulary knowledge that he acquires common words in English in a short time. He used so many repeated words in his text rather than using synonyms but we can claim that with proper

University of Malaya

and additional vocabulary instruction the student can improve his vocabulary knowledge. Apart from that, Yamamoto (2011) claims that learners use frequent repetition of words because they have learned well. Below are the examples.

Student 010 *There is also my mother who always teach me how to be a good daughter and person. A lot of things that she teach me from baby until now. She willing get hungry but she will never ever let her children feel hungry.*
Student 011 *I appreciate everything she ever did and did not either. I will always owe you in my life and success in everything I do. I can never pay her even anything portion of the struggle and hardship she went through for me always love her.*

Overall first semester results seem interesting because the text is higher with basic words and by having basic words the students can get a good grade or evaluation from the teachers. Student expands their vocabulary and basic lexical representations. Good quality of writing shows the lexical richness of students. Academically, they should focus more on vocabulary learning to improve and get a good grade where they can write a quality essay according to the pre-university standard. Next, the table below shows the final semester results.

University of Malaya

4.7 Results of Lexical Sophistication

Table 4.7 Result of Lexical Sophistication by Students Final Semester

Students Identity	First 1000 words %	Second 1000 words %	Academic word list %	Offlist %
Student FS 001	83.47	5.15	3.25	8.13
Student FS 002	87.19	4.38	0.62	7.81
Student FS 003	86.32	5.96	3.16	4.56
Student FS 004	80.45	6.40	4.50	8.65
Student FS 005	89.14	6.57	0.76	3.54
Student FS 006	83.40	4.49	1.37	10.74
Student FS 007	86.24	5.41	1.47	6.88
Student FS 008	88.11	4.65	1.81	5.43
Student FS 009	85.40	6.06	2.48	6.08
Student FS 010	74.42	4.84	7.17	13.57
Student FS 011	90.60	5.80	0.40	3.20
Student FS 012	90.67	3.62	2.10	3.62
Student FS 013	84.54	4.74	0.82	9.90
Student FS 014	91.40	4.27	1.42	3.20
Student FS 015	87.11	7.55	1.26	4.09

University of Malaya

Table 4.7 shows results ranging from 79% to 96% of K1 and K2 words. Results indicate that final semester participants show how well they used their vocabulary knowledge in their writing. All the participants fulfill the task according to their vocabulary knowledge and language proficiency. Results show that all the participants successfully produced 2000 most common words with 80% and above in their text. A claim by Cobb (2010) and Nation (2001) learners express 80% of their vocabulary knowledge with most frequent 2000 words in a text and this is considered informative. Final semester participants‘ K1 words range from 74% to 91% and K2 words range from 4% to 7%. This indicates that the final semester participants acquired the most frequent words in English and were able to express them through written production. We can claim that the final semester participants understand the topic and use relevant words to express their ideas. Students clearly stated characteristics, incidents, experiences, physical appearances to support their arguments. To answer, they need a well-known vocabulary to make the information brief and clear. To support more, the final semester students already had instructions and they will complete the course shortly. This shows that they are already experiencing and familiar with the task. The example below shows the words used by some of the final semester participants.

Student FS 002	brother, habit, deal, wisdom, life, funny, responsibilities, happily
Student FS 005	pretty, woman, lovely, sacrifice, learning, working, memorize
Student FS 014	strongest, single, educated, confident, arguments , struggled

Student FS 011 produced the highest sophisticated words with 96% by producing 90% of K1 words and 5.8% of K2 words. The Example below shows that student FS011 expressed her idea by using most 2000 common words and one off list words (underlined). We can claim that she is a typical learner who used more frequent words compare to less frequent words. She able to write her text contains the most words in the highest frequency band. To support further, student FS 011 got B- in SPM so she

University of Malaya

already seems to have the good background knowledge to accomplish this task. She understands well the words she used and makes her essay looks compiling with sophisticated words. Below is an example.

Student FS 011

When it come to friends, there is a term which is only a girl can understand another girls feelings". I totally disagree on this. As a boy, he understands me more than anyone around me. I think I am truly blessed to have a best friend like him. Having the ability to understand someone in and out is one of the character that I admire the mostly in him.

Next, the lowest sophisticated words produced by student FS 010 with 79% by producing 74% of K1 words and 4% of K2 words. The example below shows that student FS 010 did not use K2 words at all. He used 7% of AWL and 14 % of Off list words. We cannot claim that he has low lexical knowledge due to his lowest percentage because he may not able to call it up from memory and use it in the text. Student FS 010 got B+ in SPM and this shows that he can fulfill the task with his vocabulary knowledge. The example shows that he understands the proper noun and used it in his text. However, this can be least comprehensible for the reader to grasp the actual expression of the writer.

Student FS 010

Elon Musk, a South African born entrepreneur. Raised in America and inspired by the 90's technological revolution, he is hailed as the Steve Jobs of the modern age. He is the man that gives me goose bumps every time I mention his name.

Overall final semester students performed well in this task. They manage to produce text with higher frequency words that show that they are well using the words. Good quality of writing shows the lexical richness of the students because it is not necessary for students to use less frequent words or rare words to show their lexical knowledge.

University of Malaya

The students can use experienced words that they already comprehend to compile impressive text.

To answer research question 2, a comparison was made between first semester students and final semester students. The first semester shows a range of 83% to 97%. The final semester range is 79% to 96%. We can assume that first semester students used words in a sophisticated manner compares to final semester students. First semester students considered new students for the course, whereas final semester students are going to complete their course. The final semester student had more English instructions, assessments, writing practice compare to the first semester students.

The first comparison made between the first and final semesters with the highest sophistication. The first semester shows the highest sophistication by student 004 with 97% compared to the final semester by student FS 011 with 96%. Student 004 produced 90% of K1 words and 6 % of K2 words and student FS 011 produced 90% of K1 words and 5% of K2 words. Both texts contain the same percentage of K1 words and only 1 % difference for K2 words. However, student 004 describes her idol with K2 words, while student FS 011 describes using K1 and a few K2 words. We can claim that in a short time student 004 developed vocabulary knowledge because she only got C in SPM. Whereas we can assume that student FS 011 vocabulary knowledge showed no significant difference and remains the same. This is because she had the same percentage as student 004, even though she had all the training in the course. To add on, student FS 011 got B- in SPM shows that she already has good English proficiency. Conversely, we can claim that student 004 could be a fast learner to practice reading or speaking and acquire vocabulary in her daily routine. However, this is just an assumption. The example is below.

University of Malaya

Student 004 *Mother, birth, sad, angry, calm, complained*	*Student FS 011* *Friends, praise, understand, hurt, experience*

Next comparison made with the lowest sophisticated percentages between first and final semesters. The first semester shows the lowest sophistication by student 005 with 83%, while the final semester by student FS 010 with 74%.Student 005 produced 79% of K1 words and 3 % of K2 words. Student FS 010 produced 74% of K1 words and 4% ofK2 words. The computational analysis demonstrates that student 005 used more frequent words compared to student FS 010. 5% Differences does not claim significant differences. However, these differences seem moderate because student 005 got B- and student FS 010 got B+ in SPM. Another point is the total words in the text. Student 005 logged 395 words and student FS 010 with 516 words. Student 005 produced text according to the essay requirement, whereas student FS 010 exceeded the word limit. Both students have a good language background. But they produced many off list words (underlined) in their text compare to common words. The readers could not grasp the exact meaning or core content of the ideas in a text. However, we can claim that both students should improve using sophisticated words in text so that they can produce quality text. Nation (2001) claims that acquiring a word is not just simply using in a text but to know the word and meaning and produces it in the right context. Using words into writing is important rather than just writing it for the sake of completing an empty page. It is to create a good idea according to the suitability of the context. Below is an example.

Student 005 *He completed his secondary schooling at Switzerland and he enrolled in a four years mathematics and physics Diploma program at Zurich Polytechnic.*	*Student FS 010* *Steve jobs and Bill Gates which helped revolutionized the computer age, in the limelight of the revolutionary giants like Microsoft Facebook and Apple.*

University of Malaya

Besides, first semester students show poor language proficiency compare to final semester students in terms of language background. In the first semester, 3 students got A in SPM which indicates that they are outstanding in English proficiency. 2 students got B and 8 students got C grade in SPM. Whereas in final semester 4 students got A, 9 students got B and 2 students got C grade. But the findings show students with moderate English proficiency performed well in this task. First semester student 002, 004, 007, 008, 009, 010,012 and 013 got C in SPM but they able to expand the vocabulary that they produced 80% above in the text. This can be significant progress in their vocabulary development. Final semester student FS 009 and FS 012 got C in SPM and we still able to produce 85% and 90% of common words in text respectively. This can be accepted as they are final semester students who already had various learning materials to enhance their vocabulary.

Overall, lexical sophistication shows the frequency of vocabulary used by students. It indicates that frequent vocabulary is acquired earlier and how it was used in a text. A claim by Doczi& Kormos (2016) learners uses more common words they use in their daily lives. Using more frequent words indicates more refined text. According to Aviad & Laufer (2013) lexical sophistication can analyze the learner's proficiency levels. Analyzing lexical sophistication evaluates students‘ vocabulary knowledge and help them to produce a quality essay. We can assume that all participants used more frequent words to show that they have huge vocabulary knowledge. It is quite difficult to discriminate learners with different English proficiency because some students have developed their vocabulary knowledge in a short time. To conclude, students need proper learning materials to enhance their vocabulary knowledge to polish their writing.

University of Malaya

4.8 Conclusion

Overall, this chapter discussed the results of the lexical richness of text produced by the first and final semester students. The findings showed an overview of the written performance of participants. Lexical richness shows the check of variation, density and vocabulary word list of the language proficiency. The results for the first semester and final semester were not similar. We can see some vocabulary knowledge of the participants related to language background. Participants‘ lexical richness reflected on their writing. To conclude, there were no significant differences or strong justification made through data as expected. But there were noticeable lexical types such as lexical variation, lexical density and lexical sophistication in the writing. Overall, all participants have their own style of writing and richness in different types of lexical.

University of Malaya

CHAPTER 5: CONCLUSION

5.1 Introduction

This chapter summarizes and concludes the findings of this study. Insights gained from this study will be discussed as well. Additionally, suggestions are also offered for future studies in the Malaysian context.

5.2 Summary of findings

The results of the findings provide some insight into the lexical richness of students' writing. This study not only provides an insight into qualitative results but also provides applicable analysis and gave a clear picture of the lexical richness of vocabulary in terms of lexical variation, lexical density, and lexical sophistication. It also clarifies the problems in students writing and lexical richness, such as exceeding total words, less use of content words and low sophisticated words in a text. Based on this research, research question 1 was answered based on this study by highlighting the role of lexical richness in the quality of the essay produced by the student. The results of this qualitative data (student's writing) also indicate that relevant vocabulary input and using the vocabulary in the right context is the most significant factor to provide a text according to the writing standard. For research question 2, there were no significant differences found between first and final semester students. The research provided findings of the result, and show some advancement and possibilities that showed expected and unexpected results. The results of lexical richness is a useful analysis for teachers to determine what are the focus that they should be given to their students to achieve their language goal.

University of Malaya

5.3 Pedagogical Implication

From a pedagogical point of view, *http://www.lextutor.ca/*is to identify lexical richness that can measure second language learners vocabulary. Lexical richness serves as an important tool to identify student‘s vocabulary knowledge and to classify students writing according to proficiency level. It is also possible to analyze lexical knowledge and student‘s lexical richness in a particular task. Lexical richness has proven to be a reliable tool to evaluate students who are less skilled in writing based on their written production.

For a second language learner, it is difficult to acquire a language and it will take more practice to have good language skills in a particular language. Language learning takes time and more dedication is needed to acquire a language. Students should focus more on language skills such as reading, listening, speaking and writing. For example, reading storybooks, newspapers, magazines to learn new words. If they don‘t understand any words they can use a dictionary to search for meaning and acquire words. They can improve their English language skills if they practice more on language skills.

The role of the language teacher is also applicable to improve the vocabulary knowledge of the students. Language teachers can implement a variety kind of vocabulary assessment, discussion, games, movie reviews, contextual clues, connecting words, using mnemonics, playing word games and many more English language activities. The students will attempt to find the word meaning in a dictionary or take initiative to ask more questions about their subject to their teachers. Teachers ' role is important as their language input will affect student performance in academics. Teachers should prepare materials related to writing skills in their lessons and evaluations rather than focusing too much on the syllabus. The teacher can also plan an

University of Malaya

English program, according to the students‘ needs. Additional English exercises and assignments will help students score well in their academics.

5.4 Suggestion for Future Research

Future research should be possible using a longitudinal study to investigate solid progress among students. This study only focused only one time by looking at first semester students and final semester students. Future research can be done by focusing on 2 or 3 written texts from the same learner to identify significant differences or lexical development. In addition, future research can be done with different measures to identify lexical richness in student‘s writing.

5.5 Conclusion

To conclude, this study has gives some views based on the student‘s vocabulary knowledge that reflected in their writing. Overall, words play a major role in language. Without words, no one can acquire language. It is not easy to learn words, at the same time it is not impossible to learn so. Never stop learning and always make an effort to learn. So that one can be knowledgeable and perform well in academics. However, there‘s still a lot to explore and investigate in this field. More future research on this field might give a clear understanding of vocabulary among second language learners.

University of Malaya

REFERENCES

Abduh, A., & Rosmaladewi, R. (2017). Taking the Lextutor on-line tool to examine students‘ vocabulary level in business English students. *World Transactions on Engineering and Technology Education*, *15*(03), 283-286.

Aliakbari, M. (2002). Writing in a Foreign Language: A Writing Problem or a Language Problem?. *Journal of Pan-Pacific Association of Applied Linguistics*, *6*(2), 157-68.

Ali, M. S. (2003). English language teaching in primary schools: Policy and implementation concerns. *IPBA E-Journal*, 1-14.

Alqahtani, M. (2015). The importance of vocabulary in language learning and how to be taught. *International journal of teaching and education*, *3*(3), 21-34.

Arnaud, P. J. (1984). The Lexical Richness of L2 Written Productions and the Validity of Vocabulary Tests.

Arnaud, P. J. (1992). Objective lexical and grammatical characteristics of L2 written compositions and the validity of separate-component tests. In *Vocabulary and applied linguistics* (pp. 133-145). Palgrave Macmillan, London.

Azodi, N., Karimi, F., & Vaezi, R. (2014). Measuring the Lexical Richness of Productive Vocabulary in Iranian EFL University Students' Writing Performance. *Theory & Practice in Language Studies*, *4*(9).

Barcroft, J., Sunderman, G., & Schmitt, N. (2011). Lexis. In *The Routledge handbook of applied linguistics* (pp. 591-603). Routledge.

University of Malaya

Bauer, L., & Nation, P. (1993). Word families. *International journal of Lexicography*, *6*(4), 253-279.

Beatty, K. (2013). *Teaching & researching: Computer-assisted language learning*. Routledge.

Breeze, R. (2008). Researching simplicity and sophistication in student writing. *International Journal of English Studies*, *8*(1), 51-66.

Brown, G., Malmkjær, K., & Williams, J. (Eds.). (1996). *Performance and competence in second language acquisition*. Cambridge university press.

Cobb, T. (2002). Web Vocabprofile [accessed 13 June 2016 from http://www. lextutor. ca/vp/], an adaptation of Heatley. *Nation, & Coxhead's*.

Cobb, T., & Horst, M. (2004). Is there room for an academic word list in French. *Vocabulary in a second language*, *23*, 15-38.

Cobb, T. (2009). Web vocabprofile: An adaptation of Heatley & Nation's (1994) Range. *Retrieved August*, *1*, 2009.

Cobb, T. (2010). Learning about Language and Learners from Computer Programs. *Reading in a Foreign Language*, *22*(1), 181-200.

Coxhead, A. (1998). *An academic word list* (Vol. 18). School of Linguistics and Applied Language Studies.

Coxhead, A. (2000). A new academic word list. *TESOL quarterly*, *34*(2), 213-238.

University of Malaya

Coxhead, A., & Byrd, P. (2007). Preparing writing teachers to teach the vocabulary and grammar of academic prose. *Journal of second language writing*, *16*(3), 129-147.

Chang, A. C. S. (2007). The impact of vocabulary preparation on L2 listening comprehension, confidence and strategy use. *System*, *35*(4), 534-550.

Chien, C. W. (2015). Analysis the Effectiveness of Three Online Vocabulary Flashcard Websites on L2 Learners' Level of Lexical Knowledge. *English Language Teaching*, *8*(5), 111-121.

Crossley, S. A., Louwerse, M. M., McCarthy, P. M., & McNamara, D. S. (2007). A linguistic analysis of simplified and authentic texts. *The Modern Language Journal*, *91*(1), 15-30.

Crossley, S. A., & McNamara, D. S. (2009). Computational assessment of lexical differences in L1 and L2 writing. *Journal of Second Language Writing*, *18*(2), 119-135.

Daller, H., Milton, J., & Treffers-Daller, J. (Eds.). (2007). *Modelling and assessing vocabulary knowledge.*(P.150). Cambridge: Cambridge University Press.

De Beaugrande, R. (2014). *Linguistic theory: The discourse of fundamental works*. Routledge.

Dóczi, B., & Kormos, J. (2016). *Longitudinal developments in vocabulary knowledge and lexical organization* (pp. vii+-222). Oxford: Oxford University Press.

Douglas, S. R. (2010). *Non-native English speaking students at university: Lexical richness and academic success* (Doctoral dissertation, University of Calgary).

University of Malaya

Engber, C. (1993). The relationship of lexis to quality in L2 compositions. *Paper delivered at TESOL*.

Engber, C. A. (1995). The relationship of lexical proficiency to the quality of ESL compositions. *Journal of second language writing*, *4*(2), 139-155.

Failasofah, F., & Alkhrisheh, H. T. D. (2018). *Measuring indonesian students' lexical diversity and lexical sophistication. Indonesian Research Journal in Education| IRJE|*, 97-107.

Freebody, P., & Anderson, R. C. (1981). Effects of Vocabulary Difficulty, Text Cohesion, and Schema Availability on Reading Comprehension. Technical Report No. 225.

Gallego, M. T., & Llach, M. D. P. A. (2009). Exploring the increase of receptive vocabulary knowledge in the foreign language: A longitudinal study. *International Journal of English Studies*, *9*(1).

Gijsel, M. A., van Bon, W. H., & Bosman, A. M. (2004). Assessing reading skills by means of paper-and-pencil lexical decision: Issues of reliability, repetition, and word-pseudoword ratio. *Reading and Writing*, *17*(5), 517-536.

Gilquin, G., & Paquot, M. (2008). Too chatty: Learner academic writing and register variation. *English Text Construction*, *1*(1), 41-61.

Gregori-Signes, C., & Clavel-Arroitia, B. (2015). Analysing lexical density and lexical diversity in university students' written discourse. *Procedia-Social and Behavioral Sciences*, *198*, 546-556.

University of Malaya

Hadi, A. S. A. (2017). Significance of Vocabulary in Achieving Efficient Learning. *American Scientific Research Journal for Engineering, Technology, and Sciences (ASRJETS)*, *29*(1), 271-285.

Hamzah, M. S. G., & Abdullah, S. K. (2009). Teachers' teaching status and achievements of students of teaching and learning of mathematics and science in English (PPSMI) in primary and rural secondary schools. *European Journal of Social Sciences*, *10*(1), 143-161.

Ferris, D. R., Hedgcock, J., & Hedgcock, J. S. (2013). *Teaching ESL composition: Purpose, process, and practice*. Routledge.

He, L. (2018). Szudarski, P.(2018). Corpus Linguistics for Vocabulary: A Guide for Research. *International Journal of Corpus Linguistics*, *23*(1), 114-118.

Hirsch, E. D. (2003). Reading comprehension requires knowledge—of words and the world. *American Educator*, *27*(1), 10-13.

Jarvis, S. (2002). Short texts, best-fitting curves and new measures of lexical diversity. *Language Testing*, *19*(1), 57-84.

Jensen, S. J., & Duffelmeyer, F. A. (1996). Enhancing possible sentences through cooperative learning. *Journal of Adolescent & Adult Literacy*, *39*(8), 658.

Johansson, V. (2008). Lexical diversity and lexical density in speech and writing. *Lund University, Department of Linguistics and Phonetics*.

Johansson, V. (2009). Lexical diversity and lexical density in speech and writing: A developmental perspective. *Lund Working Papers in Linguistics*, *53*, 61-79.

University of Malaya

Jukneviciene, R. (2007). Analyzing topic-specific vocabulary in EFL student writing. *Retrieved from web. vu. lt/flf/r. jukneviciene/.../topic-specificvocabulary-20071. pdf*.

Krashen, S. (1989). We acquire vocabulary and spelling by reading: Additional evidence for the input hypothesis. *The modern language journal*, *73*(4), 440-464.

Kennedy, G. (2014). *An introduction to corpus linguistics*. Routledge.

Kirchner, R. E. (2004). The results of a vocabulary levels test: Version 1 conducted with foundation year chinese students. *Nation*, 2.

Koizumi, R.,& In'nami, Y. (2012). Effects of text length on lexical diversity measures: Using short texts with less than 200 tokens. *System*, *40*(4), 522-532.

Krashen, S. (1989). We acquire vocabulary and spelling by reading: Additional evidence for the input hypothesis. *The modern language journal*, *73*(4), 440-464.

Laufer, B. (1998). The development of passive and active vocabulary in a second language: Same or different?. *Applied linguistics*, *19*(2), 255-271.

Laufer, B., & Paribakht, T. S. (1998). The relationship between passive and active vocabularies: Effects of languagelearning context. *Language learning*, *48*(3), 365-391.

Laufer, B., & Nation, P. (1995). Vocabulary size and use: Lexical richness in L2 written production. *Applied linguistics*, *16*(3), 307-322.

University of Malaya

Laufer, B., & Nation, P. (1999). A vocabulary-size test of controlled productive ability. *Language testing*, *16*(1), 33-51.

Laufer, B., & Nation, P. (2001). Passive vocabulary size and speed of meaning recognition: Are they related?. *Eurosla Yearbook*, *1*(1), 7-28.

Laufer, B., Elder, C., Hill, K., & Congdon, P. (2004). Size and strength: do we need both to measure vocabulary knowledge?. *Language testing*, *21*(2), 202-226.

Laufer, B., & Goldstein, Z. (2004). Testing vocabulary knowledge: Size, strength, and computer adaptiveness. *Language learning*, *54*(3), 399-436.

Laufer, B. (2005). Focus on form in second language vocabulary learning. *Eurosla yearbook*, *5*(1), 223-250.

Laufer, B., & Yano, Y. (2001). Understanding unfamiliar words in a text: Do L2 learners understand how much they don't understand?. *Reading in a Foreign Language*, *13*(2), 549-566.

Lee, H. S. (2013). *The correlation between matriculation students' lexical richness and their writing scores/Lee Hui Shein* (Doctoral dissertation, University of Malaya).

Leedham, M. (2014). *Chinese students' writing in English: implications from a corpus-driven study*. Routledge.

Lemmouh, Z. (2008). The relationship between grades and the lexical richness of student essays. *Nordic Journal of English Studies*, *7*(3), 163-180.

Levitzky-Aviad, T., & Laufer, B. (2013). Lexical properties in the writing of foreign language learners over eight years of study: Single words and collocations. *C. Bardel, C. Lindqvist, & B. Laufer (Eds.) L*, *2*, 127-148.

University of Malaya

Lewis, M. (1993). *The lexical approach* (Vol. 1, p. 993). Hove: Language teaching publications.

Lindqvist, C., Gudmundson, A., & Bardel, C. (2013). A new approach to measuring lexical sophistication in L2 oral production.

Linnarud, M. (1976). Lexical density and lexical variation-an analysis of the lexical texture of Swedish students' written work.

Lintunen, P., Peltola, M. S., & Varila, M. L. (2014). Esipuhe-Preface. *AFinLA-e: Soveltavan kielitieteen tutkimuksia*, (6), 3-6.

Liu, M., & Braine, G. (2005). Cohesive features in argumentative writing produced by Chinese undergraduates. *System*, *33*(4), 623-636.

Maskor, Z. M., & Baharudin, H. (2016). Receptive Vocabulary Knowledge or Productive Vocabulary Knowledge in Writing Skill, Which One Important?. *International Journal of Academic Research in Business and Social Sciences*, *6*(11), 2222-6990.

Malvern, D., & Richards, B. (2002). Investigating accommodation in language proficiency interviews using a new measure of lexical diversity. *Language testing*, *19*(1), 85-104.

Malvern, D., Richards, B., Chipere, N., & Durán, P. (2004). *Lexical diversity and language development*. New York: Palgrave Macmillan.

Malvern, D., & Richards, B. (2012). Measures of lexical richness. *The Encyclopedia of applied linguistics.*

University of Malaya

McCarthy, P. M., & Jarvis, S. (2007). vocd: A theoretical and empirical evaluation. *Language Testing*, *24*(4), 459-488.

McNamara, D. S., Crossley, S. A., & McCarthy, P. M. (2010). Linguistic features of writing quality. *Written communication*, *27*(1), 57-86.

Milton, J. (2008). Vocabulary uptake from informal learning tasks. *Language Learning Journal*, *36*(2), 227-237.

Milton, J. (2009). *Measuring second language vocabulary acquisition* (Vol. 45). Multilingual Matters.

Milton, J., & Freeman, R. (1996). Lexical variation in the writing of Chinese learners of English. *Language and computers*, *16*, 121-132.

Milton, J. (2013). Measuring the contribution of vocabulary knowledge to proficiency in the four skills. *C. Bardel, C. Lindqvist, & B. Laufer (Eds.) L*, *2*, 57-78.

Mokhtar, A. A. (2017). English Lexical Acquisition of Adult Learners in Instructional Settings: Issue on Lexical Input. *Proceedings of the ICECRS*, *1*(1).

Mwinlaaru, I. N., & Xuan, W. W. (2016). A survey of studies in systemic functional language description and typology. *Functional Linguistics*, *3*(1), 8.

Nation, I. S. P. (1990). 1990: Teaching and learning vocabulary. New York: Newbury House.

Nation, I. S. (2001). *Learning vocabulary in another language*. New York: Cambridge University Press.

University of Malaya

Nation, I. S. P. (2005). Teaching and learning vocabulary. In *Handbook of research in second language teaching and learning* (pp. 605-620). Routledge.

Nation, I. S. P., &Meara, P.(2002). Vocabulary. *An introduction to applied linguistics*, 35-54.

Nation, I. S. P., & Meara, P. (2010). Vocabulary'in Schmitt, N. *An introduction to applied linguistics*, 35-54.

Nation, P., & Waring, R. (1997). Vocabulary size, text coverage and word lists. *Vocabulary: Description, acquisition and pedagogy*, *14*, 6-19.

Nation, I. S., & Webb, S. A. (2011). *Researching and analyzing vocabulary*. Boston, MA: Heinle, Cengage Learning.

Ngangbam, H. (2016). An analysis of syntactic errors committed by students of English language class in the written composition of Mutah University: A case study. *European Journal of English Language, Linguistics and Literature*, *3*(1), 1-13.

Nur, N. A. (2015). Analyzing vocabulary level of the students‘ essay writing and academic textbook utilizing lextutor. Eternal (*English, Teaching, Learning, and Research Journal),* 1(1), 15-27.

O'Loughlin, K. (1995). Lexical density in candidate output on direct and semi-direct versions of an oral proficiency test. *Language testing*, *12*(2), 217-237.

Olinghouse, N. G., & Wilson, J. (2013). The relationship between vocabulary and writing quality in three genres. *Reading and Writing*, *26*(1), 45-65.

University of Malaya

Pokrivčáková, S., Babocká, M., Bereczky, K., Bodorík, M., Bozdoğan, D., Dombeva, L.& Leung, P. (2015). CALL and Foreign Language Education: e-textbook for foreign language teachers. *Nitra: Constantine the Philosopher University*.

Read, J. (2013). Validating a test to measure depth of vocabulary knowledge. In *Validation in language assessment* (pp. 55-74). Routledge.

Read, J. (2000). *Assessing vocabulary* (pp. 1-85). Cambridge: Cambridge university press.

Richards, B. (1987). Type/token ratios: What do they really tell us?. *Journal of child language*, *14*(2), 201-209.

Saadiyah Darus & Kaladevi Subramaniam. (2009). Error analysis of the written English essays of secondary school students in Malaysia: A case study. *European journal of social sciences,* 8(3), 483-495.

Saadiyah Darus & Kaladevi Subramaniam. (2009). Error analysis of the written English essays of secondary school students in Malaysia: A case study. European journal of social sciences, 8(3), 483-495.

Schmitt, N. (2000). *Vocabulary in language teaching*. Ernst Klett Sprachen.

Schmitt, N. (2008). Instructed second language vocabulary learning. *Language teaching research*, *12*(3), 329-363.

Schmitt, N. (2010). *Researching vocabulary: A vocabulary research manual*. Springer.

Schmitt, N., & McCarthy, M. (1997). *Vocabulary: Description, acquisition and pedagogy*. Cambridge university press.

University of Malaya

Siskova, Z. (2012). Lexical richness in EFL students‘ narratives. *Language Studies Working Papers*, *4*, 26-36.

Singleton, D. M. (1999). *Exploring the second language mental lexicon*. Ernst Klett Sprachen.

Shen, Z. (2008). The roles of depth and breadth of vocabulary knowledge in EFL reading performance. *Asian Social Science*, *4*(12), 135-137.

Spencer, M., Muse, A., Wagner, R. K., Foorman, B., Petscher, Y., Schatschneider, C. Bishop, M. D. (2015). Examining the underlying dimensions of morphological awareness and vocabulary knowledge. *Reading and writing, 28*(7), 959-988.

Stahl, S. A. (2005). Four problems with teaching word meanings. *Teaching and learning vocabulary: Bringing research to practice*, 95-114.

To, V., Fan, S., & Thomas, D. (2013). Lexical density and readability: A case study of English textbooks. *Internet Journal of Language, Culture and Society*, (37), 61-71.

Ure, J. (1971). Lexical density and register differentiation. *Applications of linguistics*, 443-452.

Ure, J., & Ellis, J. (1977). Register in descriptive linguistics and linguistic sociology. *Issues in sociolinguistics*, 197-243.

Usman, A., & Abdullahi, D. M. (2018). Productive Vocabulary Knowledge of ESL Learners. *Asian Journal of Interdisciplinary Research*, *1*(1), 32-41.

Van Hout, R., & Vermeer, A. (2007). Comparing measures of lexical richness. *Modelling and assessing vocabulary knowledge,* 93-115.

University of Malaya

Vermeer, A. (1992). 12 Exploring the Second Language Learner Lexicon. *The construct of language proficiency: Applications of psychological models to language assessment*, 147.

Vermeer, A. (2001). Breadth and depth of vocabulary in relation to L1/L2 acquisition and frequency of input. *Applied psycholinguistics*, *22*(2), 217-234.

Wang, J., Cipolla, R., & Zha, H. (2005, April). Vision-based global localization using a visual vocabulary. In *Proceedings of the 2005 IEEE International Conference on Robotics and Automation* (pp. 4230-4235). IEEE.

Wang, H. (2014). *Writing and the ancient state: Early China in comparative perspective*. Cambridge University Press.

Waring, R., & Nation, I. S. P. (2004). Second language reading and incidental vocabulary learning. *Angles on the English speaking world*, *4*, 97-110.

Wilkins, D. A. (1972). *Linguistics in language teaching*. E. Arnold, 1973.

Webb, S. (2005). Receptive and productive vocabulary learning: The effects of reading and writing on word knowledge. *Studies in Second Language Acquisition*, *27*(1), 33-52.

Webb, S. (2008). Receptive and productive vocabulary sizes of L2 learners. *Studies in Second language acquisition*, *30*(1), 79-95.

Webb, S., Sasao, Y., & Ballance, O. (2017). The updated vocabulary levels test. *ITL-International Journal of Applied Linguistics*, *168*(1), 33-69.

University of Malaya

Wesche, M., & Paribakht, T. S. (1996). Assessing second language vocabulary knowledge: Depth versus breadth. *Canadian Modern Language Review*, *53*(1), 13-40.

Williamson, D. M. (2009, April). A framework for implementing automated scoring. In *Annual Meeting of the American Educational Research Association and the National Council on Measurement in Education, San Diego, CA*.

Yamamoto, H., & Haryu, E. (2011). Do Japanese 24-month-olds utilize lexical pitch accent for word recognition. In *Poster presented at the 2011 Biennial Meeting of the Society for Research in Child Development, Montreal, Canada*.

Yu, G. (2009). Lexical diversity in writing and speaking task performances. *Applied linguistics*, *31*(2), 236-259.

Yi, B., & Luo, S. (2013). Working memory and lexical knowledge in L2 argumentative writing. *Asian Journal of English Language Teaching*, *23*(1), 83-102.

Zafar, A. (2016). Error analysis: a tool to improve English skills of undergraduate students. *Procedia-Social and Behavioral Sciences*, *217*, 697-705.

Zhai, L. (2016). A study on Chinese EFL learners‘ vocabulary usage in writing. *Journal of Language Teaching and Research*, *7*(4), 752-759.

Zhai, L. (2016). A study on Chinese EFL learners‘ vocabulary usage in writing. *Journal of Language Teaching and Research*, *7*(4), 752-759.

Zhang, J. (2014). *Lexical richness and accommodation in oral English examinations with Chinese examiners* (Doctoral dissertation, University of the West of England).

University of Malaya

Zhong, H. (2011). Learning a word: From receptive to productive vocabulary use. In *The Asian conference on language learning: Official conference proceedings* (pp. 116-126).

University of Malaya