CHAPTER 1: INTRODUCTION

1.0 Introduction

The role of English in Malaysia, particularly in schools, continues to gain attention among educationists, academic scholars, politicians and parents. In the light of the importance of English, this current study seeks to examine Malaysian students' writing ability in English, with focus on pre-university learners in composing essays.

This chapter will discuss the importance of English language in a multi-lingual country and the effects of the international language on employability and spell out the objectives of the research in investigating the Form Six students' lexical richness.

1.1 Introduction to the Study

Malaysia is a multi-lingual country which comprises 140 languages spoken by different ethnic groups (Grimes, 2000). After Malaysia gained independence in 1957, Malay language was accepted as the national language with the aim to foster national unity (David, 2008). However, the English language which was first introduced by the British who ruled Malaya continued to thrive and became the main language in major sectors such as business, technology, and education especially international schools and higher institutions.

English language in Malaysia serves as a medium of instruction for English subject and also plays a significant role in both international and intranational communication (Ooi, 2001). Hence, the demand for competent users of English among students is increasing

due to the higher requirements by top companies in the working world. Fresh graduates are constantly struggling to secure a place in the competitive working world besides having the need to meet the stringent requirements to work with top-notch companies. Thus, the language has emerged as an important tool to measure students' success in both academic and career later on. Knowing the importance of the language, the Ministry of Education in this country has made it compulsory for students to take English Language paper in public examinations such as *Penilaian Menengah Rendah* (PMR) and *Sijil Pelajaran Malaysia* (SPM) before pursuing to different types of higher institutions of learning (IHL).

The education system in this country is structured as such to provide students a basis to further their studies in Form Six, Matriculation College, Community College, Polytechnic or Teacher Training. At this level, English is the main focus for pre-university students as they are obliged to take up the Malaysian University English Test (MUET) as a performance assessment on their English language before they enter the university. (Teacher Education Division, Ministry of Education, 2006 as cited in Hamzah & Abdullah, 2009).

1.1.1 Malaysian University English Test (MUET)

Malaysian Examinations Council (2001) describes MUET as a compulsory examination of English language proficiency. The test is set and run by Malaysia Examination Council and it is taken by Form Six students, Matriculation or pre-university students for the purpose of admission into the tertiary education. Rethinasamy & Kee (2011) state that MUET is similar to standardised English proficiency tests such as IELTS and TOEFL which aim for admission and placement purposes. Knowing its importance as a

placement test, educational institutions offer courses to prepare students for this high stakes test. Hence, it serves as a language improvement course for students learning English as a second language (ESL) to enhance their language proficiency before entering university.

MUET also aims to limit the broad gap of language needs between secondary and tertiary education, improve students' language competence and build up critical thinking skills. It covers four broad components: listening, speaking, reading comprehension and writing.

According to Malaysia Examination Council test description, listening skill develops students to understand different types of oral communication in social and academic situations and these students are expected to be able to critically analyse and evaluate information in English texts. Speaking skill is important for students to have active interaction and join discussions by managing and initiating conversations using social conventions. This productive skill engages students in communicating effectively and efficiently in the language.

Next, reading comprehension guides students to comprehend linear and non-linear texts. Reading skills such as skimming and scanning for important ideas in a complex text, using contextual clues to find out the meaning of a particular word, differentiating the main topic sentence and supporting details in a text and others are considered as important in language acquisition.

The writing component in MUET develops students to think critically by generating ideas and writing them down in a systematic order. It requires students to produce

various types of writing such as summaries, compositions and reports which involve higher order thinking skills and demands students to display a wide range of lexis. Through testing, students will be tested on the ability to react critically and aptly to information displayed in linear and non-linear texts.

Students are tested for their language proficiency in MUET where the test scores are allocated differently. The listening and speaking components in MUET carry 45 marks each, the writing component is given 75 marks while reading comprehension is given 135 marks. MUET gauges students in the four specified components and their language proficiency are evaluated and graded based on a cumulative score range of zero to three hundred. Next, the students are categorisedinto six levels of achievement namely Band 1 being the lowest to Band 6 being the highest (refer to appendix B) based on their cumulative scores from the four tested skills in MUET.

The scores depict the students' abilities in communicating effectively, understanding the language and performing well on a selected task. For example, Band 6 students are proficient language users who have shown extremely good command of English and uses language appropriately with hardly any grammatical errors. As a comparison, students in Band 1 level are considered as extremely limited English users who possess limited ability to function in the language.

1.2 Background of the Study

In the past, researchers (Li, 1997; Taiwo, 2004) focused on grammatical, phonological and orthographical aspects but little has been done in the field of lexis. However, several studies conducted by prominent researchers such as Laufer and Nation (1995),

Nation (2001) and Meara (1996) have been regarded as the basis of all vocabulary knowledge studies. These researchers investigated the use of lexis or words and their effects on the students' writing scores.

The importance of vocabulary knowledge in determining a successful communication in English as a second or foreign language is undeniable (McNeill, 1994; Lemmouh, 2008). In fact, Wilkins (cited in Wu, 2009) stresses that the absence of vocabulary in communication will eventually cause communication breakdown. Wu (2009) expresses that vocabulary is a vital medium to convey one's thoughts, expressions, translation and communication. Wu quotes Wilkins (2002); the renowned researcher in the field of linguistics who stresses on the role of vocabulary in communication that the absence of vocabulary will eventually cause communication breakdown.

Vocabulary acquisition is the focus of teaching and learning in Second Language Acquisition (SLA) context due to its importance in acquiring the listening, speaking, reading and writing skills (Lewis, 1993; Sánchez, A., & Rosa M., 2007; Mehring, 2005).

Several studies have shown a positive relationship between the students' usage of advanced vocabulary (lexical richness) and the quality of an essay as a whole (Jarvis, 2002 as cited in Lemmouh, 2008). Thus, this study aims to examine the relationship between lexical richness and scores of students' written essays in the Malaysian context, to ascertain if lexical richness has an impact on the quality of written essays.

1.3 Statement of Problem

The importance of English to ensure effective communication in academic, social, career and mass media is undeniable due to its status as the international language (Shafie & Nayan, 2010, Kitao, 2006). A study conducted by Wan, Shafinah and Azhari (2007) found that most research and development departments in the state of Kedah are seeking employees who are able to express their ideas orally in English (90%), deliver a convincing presentation using English (90%), write a report in English (60%), speak English effortlessly (60%), use English without grammatical errors (30%) and to have good persuasive skills in English (30%). Inevitably a strong foundation of language, especially having good vocabulary knowledge would be of great benefit to students who face difficulties in understanding specific terms in their studies, apart from just being able to use the language for successful communication with one another.

Low proficiency in English is one of the reasons why students find it hard to cope with their studies. Hamzah & Abdullah (2009) comment that poor language proficiency among students in Malaysia is mainly due to their poor foundation. One-third of students who graduated from public universities have very low English language proficiency and the Ministry of Higher Education (MOHE) stated that approximately 33% of students pursuing their tertiary education graduated with poor English language proficiency (The Star Online, 23 January 2007).

Consequently, they do not speak fluently due to limited exposure and insufficient practice outside the language classroom. Instead, mother tongue is used as the medium of communication in the home environment and in school. They feel comfortable and able to speak confidently to their friends and families using their mother tongue. This, in

turn will be resulted in weak performances in listening, speaking, reading and listening. Maros, Tan & Salehuddin (2007) state this problem is the most obvious when it comes to examinations and might persisteven when they are in the tertiary education.

Douglas (2010) in a study with NNS at a foreign university found that the level of the students' English language proficiency affects their overall academic success. Where writing is concerned, apart from accurate grammatical structures and relevant ideas, vocabulary is a vital element to measure the students' ability. In most academic contexts, written output has been accepted as a standard assessment on students' performance in education and thus vocabulary knowledge is extremely vital in securing good scores. Writing is one of the important skills to learn, it is regarded as one of the most challenging skills for NNS' students to convey their ideas efficiently (Darus & Khor, 2009).

According to Cummins (2009), various language skills are needed in language production. In timed writing examinations, students need to be able to retrieve their lexical knowledge to write efficiently. Writing involves higher metacognitive abilities which involve using accurate language and appropriate lexis in a given context. It also expects learners to connect ideas effectively in a sentence, ensure the meaning of a sentence is clear to combining meaningful sentences in order to produce coherent paragraphs. Among all the various elements required in writing, lexis plays an important role in producing, expanding and demonstrating one's ideas (Grabe, 1985; Engber, 1995; Raimes, 1983; Raimes, 1985 as cited in Douglas, 2010). Students from non-native English speaking background might face difficulties in constructing a sentence. NNS have a tendency to code-switch and this will cause confusion and complexity in structure and meaning when they are writing (Darus & Khor, 2009).

Apart from knowing how to produce and use lexis in an appropriate context, lexical richness has a strong correlation with scholastic achievement. Students who are able to write well correlate positively with their academic success (Douglas, 2010 ; Daller et al as cited in Crossley, Salsbury, McNamara, & Jarvis, 2010). Students who write well usually obtain good grades in their academic studies. Thus, this research aims to investigate the correlation between students' lexical richness and their scores in writing. To my knowledge, there have not been any such studies done in the Malaysian context with pre-university students.

1.4 Research Objectives

The objective of this research is to shed light on matriculation students' lexical richness in terms of low and high frequency vocabulary as measured by an online programme known as Range programme. This research also investigates the lexical richness by exploring the relationship between advanced vocabulary or low frequency words and the holistic scores given by teachers. Lexical richness is considered as an important component in writing and thus to a certain extent, might affect the judgment of teachers in grading students' writing. The Range programme works as the main tool to explore the correlation between lexical richness and writing scores in this research, therefore the researcher also intends to look into the effectiveness and reliability of this programme.

1.5 Research Questions

This study is guided by three research questions and they are as below:

1) What are the students' lexical richness in terms of low and high frequency vocabulary as measured by the Range programme?

2) How does lexical richness in composition writing correlate with writing scores in MUET?

3) How effective is Lexical Frequency Profile (LFP) in measuring lexical richness in writing?

The above questions will be examined and discussed further in relation to the objectives of the study.

1.6 Hypothesis

In response to the research questions that frame this research, the null hypothesis (H0) depicts that the use of advanced words in the students' essays will have no effects on their writing scores. Meanwhile, alternative hypothesis (H1) hypothesises that the presence of advanced vocabulary in the students' essays will be positively correlated to their writing scores. Hence, it shows a direct relationship between the two variables and consequently shows the impact of using advanced words in shaping teachers' perceptions in grading written essays. Therefore, if $p>0.05$, H0 will be rejected and H1 will be accepted and it shows that there is a relationship between the two variables.

1.7 Scope of the Study

In this study, 100 graded essays from the matriculation students at the Centre for Foundation Studies in Science (PASUM), University of Malaya were obtained. The essays were randomly selected from the matriculation centre after obtaining permission from the relevant department, while 50 students and 5 teachers were given

questionnaires to answer. The findings from these questionnaires are used to support the findings in this study.

1.8 Overview of the Dissertation

The first chapter introduces the background of the study, provides information about MUET and also the objectives of the study. Hypothesis and research questions are formed to meet the aims of this study. Chapter 2 provides the conceptual framework and review of relevant literature for the study. The choice of words in the academic context serves as a foundation of the study. A detailed investigation of how lexis is used and the effects on scores are also the focus inthis chapter. Next, chapter 3 further explains the processes involved in obtaining and analysing the relevant data. This study has generated a considerably large corpus of written essays that is used to achieve the objectives of the study. Chapter 4 presents the data analysis and findings of the study. Lastly, Chapter 5 amalgamates the findings in the previous chapter and provides suggestions that can help improve the pre-university students writing skills and also provide some suggestions for future research.

CHAPTER 2: LITERATURE REVIEW

2.0 Introduction

This study focuses on the relationship between students' lexical richness in writing and the holistic scores given by the examiners. Vocabulary use appears as a fundamental variable that affects student academic writing and their eventual academic outcomes. This chapter displays the conceptual framework that encompasses the variables of this study and surveys the relevant literature in each of these elements. The conceptual framework includes the research questions and the variables in this study. How social, cultural and language backgrounds affect writing proficiency, lexical knowledge and Range programme are further explained before exploring the common measures of lexical richness and previous studies on lexical richness.

2.1 Writing in an Academic Context

As discussed in the previous chapter, writing is one of the most challenging yet important skills to acquire. What elements are required for students to produce a good piece of academic writing? Writing needs sustainable input such as appropriate lexis in a particular context to produce comprehensible and meaningful sentences. Bereiter (1980) as cited in Shafie et.al (2010) state:

> "Writing proficiency develops over time. It begins as an association of ideas, growing knowledge of stylistic conventions and the use of processes for planning, evaluating and revising. Writing becomes more unified as

> writers write for an audience and transform experiences into knowledge."(p.59)

Ying (2009) states writing as one of the important skills to acquire in the process of learning English language. Writing an academic essay is deemed important as it is always carried out as a form of assessment or classroom practice to measure the students' understanding of a particular topic (Todd, Khongput & Darasawang, 2007). In a further explanation by Hinkel (2004), students' thinking skills in analysing, synthesising and reasoning are developed in the language learning process. Chase (2011) mentions that students' thinking skills and planning strategies can be enhanced through connecting new ideas with familiar ones, examining the possibility of implications, delineating information, and reinforcing theoretical frameworks.

Alamargot and Chanquoy (2001) in *Through the Models of Writing* state that the level of maturity and sufficient practice are the two most important elements that enable students to write better. On top of that, the capacity and expanding linguistic resources are also crucial in helping them to write effectively. The capacities of students' mental lexicon are affected by how well they know the issue being discussed and later activating their proper linguistic resources and rhetorical strategies to complete their assigned task of writing essays. When the students are familiar with the particular topic, they will be able to effortlessly select ideas from their long-term memory and sorting out their ideas into writing.

Meanwhile, students who have rich linguistic resources can write better and faster when they are able to choose lexis and syntactical structures more automatically. Ikah (2006)

states that students might face challenges due to the high requirement of lexis in writing an essay and thus a good grasp of vocabulary is required to be able to write effectively. At tertiary level, the significance of writing in measuring students' language proficiency cannot be denied. There are various types of assessments to examine the students' progress in writing skill. Hale et.al (1996) as cited in Hinkel (2004) carried out an in-depth research on academic writing and found that major assignments of writing 5 to 10 pages long essays are likely to be given to the students in the humanities course as out-of-class assignments.

Medium length essays of 1 to 5 pages are more frequently conducted as in-class assignment. Short writing tasks or expanded answers are also found in many written in-class and out-of-class tests, laboratory reports and case studies. Other types of popular writing assignments include rhetorical writing of exposition, cause-effect interpretation, classification, compare and contrast, analysis and argumentation are found in-class and out-of class assignments. Writing assignments such as expanded definition, process analysis, fact based exemplification and narration writing are least preferred by educators.Narration writing measures lower level thinking skills only and probably could not be used to measure writing proficiency among students in universities and colleges.

Horowitz (1986) as cited in Hinkel (2004) states the rationale behind writing assignments in academic context is to require students to display their understanding and familiarity with the course material. According to Hinkel (2004), undergraduate students are required to produce a dozen written assignments per semester which is more than what they are required to do when they are in high school. The quality of these students' writing product, such as essays and term papers need to be maintained

because it directly influences their course grades. This type of assessment can be graded or ungraded; and it can be utilized at the end of the class to measure students' understanding of the particular topic. It can enhance the students' ability in thinking and having them to write for the audience is an effective way to learn. Academic writing can be tested through several in-class or out-of-class assignments and thus, the writing proficiency of students can be measured effectively.

Writing in English as the second language has emerged in most educational programmes in non-English speaking countries. The number of newly developed writing courses in education have been growing and it is not a surprising fact as the statistics on enrolment of students into college and universities are increasing (Leki, 2009). Therefore, these learning institutions need to provide conducive environments to boost students' writing skills for academic context and workplace.

On the roles of universities in ensuring the language development of students, Erling and Richardson (2010) state that educational institutions nowadays are concerned with the students' language learning processas a part of their general academic development. Universities encourage students to have more practice in writing tasks which are related to their field of studies. Therefore, there has been a transformation towards an approach of connecting field experts with writing to form a specific field of writing (Benesch, 2001; & Johns, 2001 as cited in Erling & Richardson, 2010).

According to Nakamaru (2010), students need to seek advice and assistance from experienced and trained tutors or peers to overcome their writing difficulties. Nevertheless, writing centres and professional guidance in local universities are limited. The lack of tutors who are experts in the writing field indirectly affects the students'

ability to write well because they do not have a place to turn to when they encounter problems in writing. Moreover, most NNS students enrol in English as a Second Language (ESL) or Intensive English programmes (IEPs) to enhance their English language proficiency. These programmes are particularly useful for them to further improve their academic writing. This is an important fact to acknowledge as NNS students display numerous problems in writing. Most students are not taught how to write academically even in their first language.

It is crucial that an academic writing component is present in a course or programme to prepare and guide NNS to write academically. They need to gradually build up their skills and proficiency in writing for the academic writing course to exhibit good academic knowledge as expected in the course instead of becoming a good narrator of self-experiences or personal stories (Hinkel, 2004).

However, the difficulty to produce written work academically influences the perceptions of professionals in the education field. These perceptions arise and affect the students negatively as they gradually become frustrated and lose interest in learning the language. Some of them seemed to think that academic institutions are irrational in keeping up the standards of their language and are ignoring their efforts which they put in their writings (John, 1997 as cited in Hinkel, 2004).

Having considered the importance of writing and the high demand of it in academic institutions, this research is done to investigate the effects of several important elements such as lexis that works as a foundation of a good essay.

2.2 The Writing Processes

According to Hinkel (2004), writing can be divided into two types: product-oriented and process-oriented. Traditional approach in teaching writing emphasised on product-oriented writing, with an end result being the completed composition of a student. Most educators in the 1960's and 1970's who assigned topics for students to write either as an in-class or homework assignment tend to emphasise on the outcome of the task instead of focusing on the process of writing (Hinkel, 2004). As a consequence, students are inclined to ignore the process of writing and tend to have difficulty in producing good quality writing. However, the emergence of early reformers managed to shift the focus of writing to process-oriented where educators focus on students' thoughts and complex skills during the process of writing, and not just focus on accuracy in form and content (Zeng, 2005).

Nor Azni Abdullah (1993) claims that the emphasis on writing instruction has indeed shifted from a focus on form to a greater concern for the writer and the writing process. She asserts that the change in this strand is backed by the view that writing may be perceived as having a liberating influence on the writer and a means for exploration of self. Since the 1990s, Malaysian ESL teachers have also tried to make this shift and have begun to move away from a product to a more process based approach to writing. This is done in view of the call by the former Prime Minister of Malaysia, Tun Sri Dr.Mahahtir Mohamad in his Vision 2020 mandate, that our citizens should be able to not just consume information rapidly in the fast-changing information era, but to make sense and meaning in their communication, and writing is one way they can do this.

The process of writing and how it takes place in one's mental lexicon is indeed complex. Olive (2003) views writing as the most complicated cognitive process that one needs to achieve because writing requires several main cognitive components to function at different stages of representation. Based on his explanation of the mental processes at the semantic stage, planning processes involve putting up a pre-verbal message that relates to the ideas a writer needs to convey. These ideas are usually extracted from the writer's long term memory and re-organisation of the ideas will then be carried out if necessary.

The writer's long term memory or background knowledge is found to be crucial in helping readers to comprehend a text. Background knowledge comes from several perspectives such as world knowledge, cultural knowledge, subject-matter knowledge and linguistic knowledge. However, world knowledge is rather subjective because the knowledge varies from countries, regions and cultures (Pang et. al, 2003). Nakamaru (2010) observes that students' language proficiency in English language might be impeded due to their language backgrounds. NNSs may have various potencies and needs based on their previous knowledge in writing in the targeted language.

It is noted that reading and writing skills are interrelated in the academic setting. Fitzgerald & Shanahan (2000) and Farahzad & Emam (2010) emphasise that students need a large amount of input garnered from reading to determine the quality of writing. Besides that, Durukan (2011) also states that there is a relationship between reading and writing, and reiterates that writing and reading are the first skills to be learnt by students. Further to support this point on the correlation between the two skills, Kessler, cited in Farahzad & Emam (2010) said that "good writers are good readers… good reading is the key to becoming a good writer".

While some researchers have shown the connection between reading and writing skills, Graham and Perin (2007) view both these skills as important aspects of literacy which require different instructions, a finding from their study on difficulties faced by teenagers in writing. In a further explanation, they state that writing not only plays the role as the basic requirement for involvement in civic life and the global economy, it also serves as a necessity in the academic field. Though writing is crucial, high school graduates are unable to display good writing skills even at the basic levels as required by institutions or employers. They do not have the foundation of literacy skills of reading and writing to meet the growing demands of high school curriculum (Pang et. al, 2003; Snow & Biancarosa, 2003). Thus, poor writing skill should be treated as one of the factors of national literacy crisis.

Students who enter tertiary education without a solid foundation of English language have difficulty in understanding and using the language especially in producing written output. They have problems in meeting the basic requirements set by colleges, universities and work places (Graham & Perin, 2007). McCabe (2003) as cited in Chase (2011) states that an estimation of one third students who failed to reach the standard scores for English writing courses plan to enrol in a college or university. Consequently when they do eventually succeed to continue their studies in universities they still fail miserably in their writing.

According to Llach (2010), writing is one of the most difficult skills to manage in studying English as L2 due to the high subjectivity of the nature of writing. The obstacles in writing can affect the English development of NSs and NNSs. (Bonanno & Jones, 2007; Paton, 2007).

Weigle (2004) conducted various studies on writing difficulties faced by students under examination contexts and found that the NNSs do not perform well. Another study by Ruetten (1994) found that 62% of NNSs have difficulties in writing compared to only 30% of NSs who could not pass the second semester. On the perception of lecturers on NNSs writing ability, Nakamaru (2010) found that most lecturers in academic institutions labelled NNSs as students who need guidance with English language. These findings tend to reflect that lecturers and tutors in writing centres come to a conclusion that NNSs are weak in expressing themselves using English language as English is not their first language (Leki, 2009).

2.3 Impact of student's Social, Cultural and Language Backgrounds on Writing Proficiency

Hinkel (2004) asserts that learning to write in the target language is different from learning to write in the first language. While NSs have acquired and developed their English language proficiency as their native language, NNSs have to devote years of learning it as a second language. NNSs have to put more effort in learning the language compared to NSs. They have to overcome the obstacles in learning English in their studies at tertiary level education. According to Chen (2007), NNSs perceive writing in English as a difficult skill to acquire although the importance of writing has always played a crucial role in the development of their language proficiency. Yet, the difficulties of writing academically not only apply to NSs but NNSs as well.

There are differences in terms of the specific challenges of writing to both the NNSs and NSs. Based on English for Academic Purpose (EAP) needs analysis, Berman and Cheng (2010) reveal that the language proficiency varies between NSs and NNSs in the

academic field. The needs analysis looked into the perceptions of undergraduate NNSs on the most difficult language skill to acquire and its' comparison of perceptions with the NSs. The results displayed that both groups of students perceived productive skill of writing as the most challenging skill to acquire compared to receptive skills. Writing skills need to be taught explicitly in classroom by including academic writing tasks to students because they need help in improving their language proficiency especially in written production.

Another study conducted by The University of Hawai'i M~noa Writing Program on students' language proficiency, the results showed that NNSs show interest in how languages work but due to the different languages and cultural backgrounds of these students, they have distinct needs and skills compared to NSs. NNSs tend to have the perception of learning English as an obstacle in their study and their main concern is the inability to express ideas and specific concepts in English. Their writing often ends up confusing readers especially experts in the education field. Therefore, in comparison to NNSs, NSs are often regarded as outstanding students with high motivation to excel in their studies, while NNSs gets de-motivated and some tend to give up and discontinue their studies.

Fuentes (2009) states that the written product of NNSs shows certain characteristics which reflects of their incompetence. Jafarpur (1996) notes that the lexical grammatical knowledge of NNSs and NSs' writing performance, in terms of lexical precision, differ greatly. He found that NNSs' scores in lexical knowledge are lower than NSs. However, NNSs written production are not necessary worse than NSs as a further investigation found that NNSs' written product seems to have better content knowledge compared to

NSs who excel only in linguistic command. Hence, it may be assumed that NNSs prioritise the content of their writing rather than grammar and lexis.

This is a completely different view for NSs as they show more concern on the use of grammar rather than the content in writing. NNSs are able to keep pace with NSs but they only differ in terms of grammar and structure use in writing (Burrough-Boenisch, 2003). NNSs are able to maintain their academic scores on par with NSs but their process of learning English as L2 depends largely on various complex factors. Bialystok (2001) argues that NNSs might not be able to use and speak the language fluently like the NSs even after dedicating years of learning the language in a non-native speaking environment. At times, they might also feel pressured with their outcomes in writing even after countless efforts in learning the language by sacrificing time to improve their fluency in English language (Severino, 2009).

As a consequence, NNSs are often regarded as students who face challenges in improving their writing skills (Bacha, 2002). There are many colleges or universities which use English as a medium of instruction and learning, and these institutions have negative views on NNSs' abilities in producing particular types of writing such as reports, summaries and thesis. Hinkel (2004) and Lillis & Scott (2007) claim that advanced NNSs demonstrate several grammatical problems and underperform in their general academic studies because of the emphasis on writing as an important assessment mode in universities. Identifying and utilising appropriate grammar and structures in writing are seen as the main problems in fully acquiring English among the NNSs. Their poor language proficiency such as the lack of lexical and grammatical skills has huge implications in their academic writing.

As mentioned earlier, there are distinct differences between NSs and NNSs. The differences might be due to cultural, social and language backgrounds, all of which play an important role in determining their English language proficiency. Their language backgrounds such as their mother tongue use and the lack of exposure to the target language may be some of the hindrances for NNSs to master the English language to produce accurate and coherent pieces of writing. On the cultural aspect, Pandian (2000), in his study on readership behaviour among multi-ethnic Malaysian students, found that they have poor readership skills. He termed this as "readership behaviour" which refers to the lack of interest in reading practices regardless of the availability of the different forms of media such as newspaper, magazines and radio as reading materials.

Malaysian students are literate but they are not avid readers, according to a survey on the reading behaviour among Malaysians. The Ministry of Education found that only 20 percent of Malaysians are active readers and this includes students who read for examinations. Fitzgerald & Shanahan (2000) and Farahzad & Emam (2010) research on the link between the exposure of reading materials and vocabulary knowledge reveal that reading determines how well a person writes. From the readership behaviour conducted on NNSs in Malaysia, it is found that most of the students are not proficient readers and this has affected the outcomes of their writing.

Apart from that, the social background of an individual also has a great impact on the language production in English language. Choy & Troudi (2006) investigate the changing perceptions and behaviours of Malaysian students in learning English as a second language in a local college. They found that students rarely use the English language outside their academic institutions. They are exposed to the language in classes, specifically the English language as a subject and also Maths and Science in

schools. However the English language is regarded as a foreign language rather than a second language. The limited exposure to the English language has apparently influenced the language ability of these students negatively.

Social and cultural backgrounds are proven to bring effects on the ability of students to read and speak either positively or negatively. According to Komolafe & Yara (2010) from this language background perspective, most students who come from a multi-lingual family tend to have more problems in writing in English. This is mainly due to the reason that NNSs reside in a society where the mother tongue is the most dominant language. It is undeniable that the effect of the mother tongue has long been accepted as a major determinant and it has negative interference on the students in learning English as a second language. According to Yong, Tan and Yong (2012), "In a country that is full of diversity in terms of race and culture, it is only natural for Malaysians to be influenced by their first language when using English as their second language" (p. 19).

The students' language proficiency in English might also be affected by their language backgrounds. The students who come from a non-English speaking background may have an impact on their English language proficiency. Different language backgrounds such as using their mother tongue and the amount of exposure to the English language may be the obstacles for them to master the language. Giridharan and Conlan (2003) found that the amount of input in the target language influenced the outcome in one's writing and speaking skills.

The review on NNSs social, cultural and language backgrounds as given above are also reflective of the situation among Malaysian students who are also learning English as a second language. The current study is set in a Malaysian matriculation centre and the

subjects are ESL learners who also face many problems with their writing skills. The limited exposure to the language due to negative readership behaviour, their different upbringing, and the use of their mother tongue (namely Bahasa Malaysia) as the dominant language in their daily lives appear to impact the writing outcome and indirectly affect their course grades.

2.4 Lexical Richness in Writing

As mentioned earlier in Section 2.3 on the impacts of NNSs social, cultural and language backgrounds on writing proficiency, these aspects are usually associated with limited vocabulary knowledge which may affect the students writing ability. Wang (2005) states that NNSs tend to have smaller vocabulary size compared to NSs. Lexis is known as the most fundamental feature in linguistic and academic progress, and plays a significant role in the quality of the compositions among students who are learning English as a second language (Llach & Gallego, 2009). Thus, a well-written composition consists of a wide range of words used appropriately in a context (Engber, 1995; Laufer & Nation, 1995).

Vocabulary or lexis is one of the most important aspects in linguistic competences which have been found to be well-correlated with general language competency in learning English as a second or foreign language (Jukneviciene, 2007). In the educational field, individuals who are qualified to evaluate students' writing found that NNS students' writing have the most errors on lexis and this is considered a serious issue (Santos, 1988 as cited in Hinkel, 2004).

Lexical richness might be helpful when students are required to write for academic purposes at tertiary level studies. Insufficient lexical knowledge often results in difficulties among students when they are required to write an essay (Tercanlioglu, 2004). Andrews (2009) asserts that using a varied vocabulary in writing is tied to academic success and Wang (2005) also views lexical richness as a key determination in securing good grades in writing. With this in view, using only a limited range of lexical items in writing may bring negative consequences on the quality of writing. Most college students are restrained in utilising active vocabulary such as advanced or difficult words in their compositions due to their inadequate active lexis.

> "College students are confined to a rather limited selection of active vocabulary: they always avoid picking out comparatively higher-level or advanced vocabulary (e.g. college English Band 4 and Band 6 vocabulary) and tend to employ comparatively lower-level or basic words; there is comparatively high repetition of words in the same composition." (Feng, 2008, p.111)

This difficulty will arise when they are required to use more active vocabulary in producing written output. Feng (2008) on the recurrence of words in a written composition, states that it has a negative effect on learners' writing, besides diminishing their development in vocabulary variation and sophistication. Lexical richness is not only limited to developing students' linguistic competence but also plays a major role in NNSs to acquire a language (Crossley et al., 2010, Jukneviciene, 2007).

An increase in the use of richer lexis may put students in a better position in terms of scores in writing compared to others who have inadequate lexis. Looking at this relationship between lexical richness and scores in writing, NSs have a greater advantage because their lexical richness enables them to write better than NNSs. Laufer and Nation (1995) concur that limited lexical resources appear to reduce writers' possibilities in expressing their ideas.

Students might exert higher frequency words when producing written works and this often results in unsatisfactory grades. (Feng, 2008; Lemmouh, 2008). Llach (2010) states the importance of selecting the right vocabulary because this may in turn determine the quality of students' writing. Educators might give students better scores if they utilize more low frequency words in their writing, as it reflects their ability to use advanced lexical knowledge and have a better idea of the writing topic.

2.5 Lexical Knowledge

> "To the Anglo-Saxons a vocabulary was a 'word board', to be owned and treasured; to the Chinese, a sea of words to be fished." (Morgan 1986:4)

> "No text comprehension is possible, either in one's native language or in a foreign language, without understanding the text's vocabulary." (Laufer, 1997:20)

The quotes best describe the importance of words or lexis regardless of learners'backgrounds or cultures. However, the question that still persists is how do

humans store words in mental lexicon? Mental lexicon is a complex structure where words are stored and organized according to the linguistic aspects such as phonology, semantics, syntactic, as well as other non-linguistic aspects (Tamariz, 2004; Sripada, 2008 and Elman, 2004). Hrabincova (2002) comments that this remains vague and would continue to be an issue that needs to be further researched on.

Pustejovsky (1996) has labelled mental lexicon as the 'sense enumeration model'. In his view, words are entered into the mental lexicon as a list of information. However, the kind of information that goes into the mental lexicon may be deliberated. According to Elman (2004), a crucial aspect of knowing a language is to know the words of the particular language. Therefore, mental lexicon is regarded as a form of mental dictionary, where words act as stimuli on mental states. Tamariz (2004) states that mental lexicon will be utilized actively to produce words during language production such as speaking and writing.

Lexis or words need to be able to be pronounced, transmitted, processed and decoded in such a way so that one can understand and deliver the intended message. In mental lexicon, words need to have certain representations to be triggered to enable one to retrieve the word with its linguistic aspects. On the structure of the mental lexicon, Tamariz (2004) describes it as a flexible and robust component that is able to sustain adaptive pressures on it. The lexicon is constantly developing and adapting to changes, and this is referred to as 'homeostasis' because of its ability to juggle all the pressures and at the same time differentiate between independent elements.

Words are used daily with attention to the selection and use of appropriate word in a context. To use the right word within a wide range of vocabulary knowledge, language

production such as writing and speaking place high demand on mental lexicon. Thus, it makes the selection of appropriate words and the formation of words in a string of meaningful sentence which will allow one to communicate effectively, clearly and accordingly. Segler (2001) in his research on Second Language Vocabulary Acquisition and Learning Strategies in ICALL Environments identified lexical knowledge as word knowledge that can be looked at from different perspectives, that is from the quantitative and qualitative perspectives.

The quantitative perspective can be defined as the number of words that a student knows but is uncertain of, whilst the latter relates to the quality of words that a learner knows such as using a particular word in the right context, knowing the meaning and producing it appropriately in written form. The quantitative perspective is always related to the vocabulary size of a learner.

Vocabulary size is the total number of words or lexis that a learner knows and in which he/she has some understanding of the meaning (Llach & Gallego, 2009; Crossley & et al., 2010; Mehrpour, Razmjoo, & Kian, 2011). Llach and Gallego (2009) deduce that there is a close and positive relationship between vocabulary size and reading comprehension, and a significant correlation between lexis and the quality of writing. Not only that, Vermeer (2001) also proves that there is a relationship between language proficiency and vocabulary size or lexical richness.

Vocabulary size plays a crucial part in institutional placement whereby its assessment covers an extension of words families (Laufer & Nation, 1999). Nadarajan (2008) comments that there is a standard amount of stored words or vocabulary size for students in a particular level. Various researchers (Bauer & Nation 1993; Nation, 1990

andMeara, 1996) have pointed out the challenge in defining what knowing a word means. Nation (2001) compiled a list of what knowing a word means in Table 1.

Table 1: What is involved in knowing a word?

Form	spoken	receptive	What does the word sound like?
		productive	How is the word pronounced?
	written	receptive	What does the word look like?
		productive	How is the word written and spelled?
	word parts	receptive	What parts are recognisable in this word?
		productive	What word parts are needed to express the meaning?
Meaning	form and meaning	receptive	What meaning does this word form signal?
		productive	What word form can be used to express this meaning?
	concept and referents	receptive	What is included in the concept?
		productive	What items can the concept refer to?
	associations	receptive	What other words does this make us think of?
		productive	What other words could we use instead of this one?
Use	grammatical functions	receptive	In what patterns does the word occur?
		productive	In what patterns must we use this word?
	collocations	receptive	What words or types of words occur with this one?
		productive	What words or types of words must we use with this one?
	constraints on use (register, frequency …)	receptive	Where, when, and how often would we expect to meet this word?
		productive	Where, when, and how often can we use this word?

(Nation, 2001:27)

The number of words increases in demand when students need to succeed in receptive skills such as reading authentic materials. A vocabulary size which consists of 3000-5000 word families is deemed as ideal for them to carry out their actions (Waring & Nation, 1997). A vocabulary size of 10,000 word families is required to accomplish a higher level of reading activity such as reading university text books. NNSs need to have a deep breadth of vocabulary size to read academic books due to the presence of

specialized terms (Hazenberg & Hulstjin, 1996). To enable NNSs to communicate in English language, they should have at least 2000 words (Schmitt, 2000 and Meara, 1996).

On another note, Waring and Nation (1997) and Douglas (2010) label NS as someone who has an extensive breadth and depth of vocabulary size. A NS preschool child has a large vocabulary size of 4000 to 5000 word families, while an average university student has 17,000 and university graduates should have about 20,000. These researchers conclude that NS vocabulary size is more extensive due to the constant use of the language.

The wide gap between the vocabulary size of NSs and NNSs exists because the latter only manage to acquire less than 500 words after they dedicated several years to learning English as L2 while the NSs can simply add about 1000 word families every year. Nevertheless, an adult NNS can still achieve a stable growth of vocabulary size in the target language. Vocabulary size can be divided into quantitative and qualitative perspective. The quantitative perspective is always related to the breadth of vocabulary whereby the qualitative perspective is linked to the depth of vocabulary knowledge.

2.5.1 Breadth and Depth of Vocabulary

Vocabulary or lexical knowledge can be defined as the main focus in language acquisition and it comprises at least two dimensions which are vocabulary breadth and depth (Alderson, 2000; Schmitt, 2000 and Nation, 1997). There is a dispute in defining the difference between vocabulary breadth and depth due to their close relationship in vocabulary knowledge.

Kuang (2011) refers to the breadth of vocabulary knowledge as vocabulary size. According to Zimmerman (2004), vocabulary breadth deals with the number of word families a student knows and the surface meaning of the words. It is also known as a discrete - point approach which assesses students' receptive knowledge based on recall and recognition of words besides dealing with the number of words that they know (Johansson, 2008).

Vocabulary depth refers to how well a person understands a word which includes synonymy, polysemy and collocation (Zimmerman, 2004). Johansson (2008) termed vocabulary depth as the assessment of quality of vocabulary or lexical knowledge which measures the learners' understanding of the meaning of words and its uses in context. Depth is not viewed as an independent construct but as a whole in its role in natural communication.

Thus, students who produce and use words or lexis accurately in communication and writing are those who have receptive-productive knowledge (Read, 2000; Zareva, 2005). Qian (1999) who carried out a vocabulary depth study on 77 Korean and Chinese students found that their knowledge of primary words comprise of understanding the synonymy, polysemy and collocation. Qian conclude that the participants in the study will have a better understanding if the particular word has the elements which are similar to their first language.

2.5.2 Active and Passive Vocabulary

Apart from the breadth and depth of vocabulary knowledge, knowledge of active and passive vocabulary also plays an important role in acquiring a language. Passive vocabulary or receptive vocabulary involves a cognitive process in understanding a word form and acquiring its meaning when reading or listening. Active or productive vocabulary is related to the desire of expressing a meaning and producing the right word either in spoken or written forms (Nation, 2000).

Students acquired more receptive vocabulary compared to productive vocabulary and thus increasing the size of vocabulary knowledge (Nation 2001; Read 2000; Zhou 2010; Zimmerman 2004). By considering the productive vocabulary aspect, a student may remember the word '*impede'*and is able to provide an identical word with the same meaning when the word is spotted within a text. However, the student might not be able to utilise the word while producing productive knowledge such as in writing and speaking (Laufer, 1998).

Therefore, the process of transforming receptive vocabulary into productive vocabulary requires a lot of effort and this change is deemed as a challenging task by most students. They mainly acquired receptive vocabulary through subconscious acquisition of passive words from listening and reading, and perceive the newly encounter word form and retrieve its meaning before utilising it as an active word. Zhou (2010) states vocabulary knowledge as a developmental process in which a word builds up from passive to active word level.

2.6 Previous Studies on Lexical Richness

Previous studies by Goodfellow et al. (2002), Nadarajan (2011), Mokhtar (2010) and Engber (1995) found significant correlation between the students' lexical quality and their language production, specifically in writing skill. Goodfellow et al. (2002) examine the writing of a group of learners of French as a foreign language course by utilising lexical frequency. A total of 36 students from a Level 1 French course participated in the study. The students' essays which had been submitted and graded by their tutor were transcribed and analysed using the French LFP programme. The researchers found that there is a significant correlation between students' lexical knowledge in written texts and the grades awarded by their tutor.

Nadarajan (2011) examines the relationship between advanced words and holistic scores of L2 students' compositions before comparing their lexical richness with L1 students. A total of 387 sample compositions which were analysed using the Range programme revealed that there was no relationship between advanced words and holistic scores. However, the study revealed contradictory results between lexical richness, holistic scores and teachers' evaluation. Evaluation of teachers on NSs and NNSs writings provided an insight into NNSs' abilities in learning and using the words similarly to their NS peers. The teachers are found to award better grades to writings that contained more advanced words or low frequency words.

Mokhtar (2010) conducted a study to examine the students' lexical knowledge by utilising four vocabulary assessments. The assessments are the Vocabulary Level Test (VLT), Passive Vocabulary Test (PVT), Controlled Active Vocabulary Test (CAVT) and Free Active Vocabulary Test (FAVT). These tests were conducted with 360

university students. The students were asked to compose an essay of 300-400 words on the title "University education should be made free for all Malaysians. Do you agree?" Then, the researcher analysed the essays using the Range programme which compares the words in the essay with the Lexical Frequency Profile (LFP) word list. It was found that most of the students have restricted lexical knowledge and could not utilize the wordsappropriately in their writing.

Engber (1995) also investigated the relationship between samples of timed essays written by NNSs and grades given by examiners. He found that students who use rich lexical choice correctly in writing tend to have better grades. Similarly, Astika (1993) also found a similar finding where lexical richness is proven to be the strongest predictor in the students' writing proficiency grades.

In a study carried out by Teoh (2009), compositions were obtained from young adult ESL students in a selected private university college in Malaysia. The objectives of the study were to measure the quality of compositions written by ESL students and to investigate whether exposing students to vocabulary learning strategies would improve their vocabulary knowledge. Written compositions were collected from experimental and control groups before using the Range programme to categorise the words produced in the compositions into four levels of word frequency. The compositions were also holistically scored for vocabulary and data analysis showed that there was a weak relationship between the holistic vocabulary scores and the LFP.

However, not all the research produced the findings given below. Laufer (1998) conducted an extensive research to investigate students' lexical richness in writing by analysing the vocabulary used in a composition of about 300 - 400 words. The

participants were given an opinion-based essay to write on, entitled "Should a government limit the number of children in families?" Laufer allocated a specific duration of time (90 minutes) for the students to complete the writing and the compositions were later collected and analysed using the LFP software.

The scores were calculated and tabulated based on three categories of scores: the passive vocabulary score, the controlled active score and the LFP. The results were not expected as Laufer found that the students' free active vocabulary did not have a clear relationship with passive and the controlled active vocabulary knowledge. Thus, the research proved that students who are exposed to a wide range of knowledge and comprehend more lexis compared to other students were not equivalent to the notion that they are able to utilise more low frequency vocabulary in their writing.

Lemmouh (2008) carried out a study to explore the connection between lexical richness and the holistic scores of 37 advanced Swedish university students using LFP. He measured the proportion of low frequency words in the essays and categorised them to three distinct variables: essay grade, course grade and vocabulary knowledge. To further validate the findings of his research, a 14-item questionnaire was directed to the teachers at the English department to prompt responses on the marking criteria of a written composition.

Nevertheless, his study found out that there was no noticeable relationship between students' lexical richness and the writing quality as mirrored by the teachers' ratings. Therefore, it may be concluded that this might be due to the typical marking criteria used by teachers in the English department which focused more on the content and grammatical aspects rather than just looking at lexical features alone.

To sum up, it is observed that over the years, lexical richness and writing are interrelated, and this has created increasing interest of prominent researchers to carry out studies to examine the relationship between the two variables. Similar studies have been conducted on various learners' different language backgrounds to prove the existence of the relationship between the impact of vocabulary and writing quality.

While on one hand, some researchers have found that there is a close relationship between advanced vocabulary and the quality of written output, others on the other hand have shown contradictory results. Thus, it has prompted a similar study like this to be conducted to find out the relationship between students' lexical richness and their writing scores in the Malaysian context.

2.7 Lexical Frequency Profile/Range Programme

Looking at the importance of vocabulary in written composition, Lexical Frequency Profile (LFP) or the Range programme was developed by Laufer and Nation in year 1995 to fulfil its original purpose of examining whether a particular assessment is suitable for learners with a specified level of proficiency.Building on this work, they later extended the purpose of LFP to measure the lexical richness of NNSs in their writing (Meara, 2005).

In a further explanation by Meara (2005), he states that Lauren and Nation claimed LFP as a practical and analytical tool in measuring students' lexical richness in writing, providing identical findings of different written work done by the same person, distinguishing students of different levels in language proficiency and also enabling the correlation between writing with an autonomous measure of lexical knowledge.

Pokorny (2009) and Bogaards & Laufer (2004) explain that LFP framework evaluates students' lexical knowledge in writing by using Range programme. LFP is a programme which is suitable for usin examining the lexical richness in writing among students learning English as a foreign language or second language. Laufer and Nation categorized LFP into four frequency lists.

Band 1 covers the 1000 most frequent words in English, Band 2 covers the next 1000 most common words, Academic Word List (AWL) includes 3,100 words of 570 word families in scholastic texts and the fourth category Not In the Lists (NIL) covers the less frequent words (Smith, 2005). On another note, Zhou (2010) clarifies that there are two most frequently cited word lists. The first word list is University Word List (UWL) by Xue and Nation (1984) and the second word list is Academic Word List (AWL) by Coxhead (2000).

AWL covers a higher percentage of the 3.5 million-word corpus of scholastic texts and composed of four major disciplines of scholastic texts such as arts, commerce, law and science. In a further explanation, Coxhead (2000) describes that AWL has advantages over UWL because UWL has lower reliability of selection values and have flaws in previous works.

Based on the earlier explanation of the crucial role of the types of vocabulary in Sections 2.5.1 and 2.5.2, one cannot deny the importance of active and passive vocabulary in learning and acquiring a language. An adequate knowledge of active and passive words is important to ensure successful communication. The use of appropriate vocabulary or lexis is not only important in communicating more effectively with one another, but also crucial in determining the quality of a piece of writing. According to

Read (2000), the lexical knowledge of a learner can determine the quality of a written work and thus the assessment of lexical knowledge has become the focus in academic writing besides grammar.

Meara & Bell (2001) take intrinsic and extrinsic measures of lexical variety into consideration when determining productive vocabulary in writing. Intrinsic measures denote the types and tokens in a text while extrinsic measures signify the consultation of additional information about the words being used beyond a given context. Nation (2001) and Read (2000) further elaborate on the definition of types as running words in a context and tokens as the sum of *different* word forms in a context. For example, the sentence "The girl likes the cookies that her mum bakes" consists of nine *tokens* while the sentence has only eight different *types* because the article 'the' is a single type repeated twice.

To determine writing quality, there are several electronic tools that the educators can utilise to identify a student's lexical knowledge. There are some other tools developed between 1980s and 1990s such as Lexical Originality (LO), Lexical Density (LD), Lexical Sophistication (LS), Lexical Variation (LV), Semantic Variation and Lexical Quality. However while each of these have its own strengths, they are not without flaws. Nonetheless these are tools which are able to offer some amount of objectivity to such research and have been accepted as reliable and sensitive lexical measurement tools.

Firstly, Lexical Originality is the percentage of words used by a particular writer and no one else in the group. In other words, this tool assesses a learner's performance relative to the group in which the composition was composed.

$$LO = \frac{\text{Number of tokens unique to one writer x 100}}{\text{Total number of tokens}}$$

Based on the formula above, the index will change when the group changes. Thus, LO of a particular composition is unstable because it is characterised by both the composition question and the group factor.

Next, Lexical Density is the percentage of lexical words in a composition, such as nouns, verbs, adjectives and adverbs.

$$LD = \frac{\text{Number of lexical tokens x 100}}{\text{Total number of tokens}}$$

Looking at the importance of lexical words in a composition which primarily display information, it is considered ‘dense’ if a particular composition consists of too many lexical words compared to the total number of words. As this tool is affected by the number of words, researchers doubt the validity of this tool to measure lexical richness.

The third tool is Lexical Sophistication which measures the percentage of low frequency words/advanced words in a composition.

$$LS = \frac{\text{Number of advanced tokens x 100}}{\text{Total number of lexical tokens in a word}}$$

The researcher needs to pick out words that are deemed as advanced words in a learner’s composition and thus, it is important to measure LS using the particular

learner's level as a basis. Nevertheless, this tool has its disadvantages due to the different perceptions of advanced words among researchers. Besides, learners from other countries who have different educational backgrounds might make the measure unstable.

Another analytical tool which has been used is Lexical Variation which is the type/token ratio between the different words in the text and the total number of running words.

$$LV = \frac{\text{Number of types x } 100}{\text{Number of tokens}}$$

The limitation of this tool is the inability to measure a learner's lexical richness in a short text. Researchers might solve this problem by having a fixed number of words in a composition, but the problem still exists because it fails to distinguish the quality of different advanced words in a composition.

Apart from the tools described above, the more recent electronic and commonly used tools which are being used to guide educators in measuring students' lexical knowledge are P_Lex, V_Size and Lexical Frequency Profile (LFP). P_Lex is a probing tool for educators or researchers to identify the lexical difficulties in short texts but this may only be used with caution given as it has not been thoroughly tested. It has similar characteristics like LFP where the outcome sorts out words in the text into a frequency list.

The other tool known as V_Size deals with NNSs' productive vocabulary in writing and it is based on a supposition that "texts generated from a vocabulary size will tend to have a characteristics shape" (Meara & Miralpeix 2007, 1). V_Size is able to generate a lexical frequency profile and also information of lexical knowledge of the particular learner who wrote the text (Meara & Alcoy, 2008).

Lexical Frequency Profile (LFP) is a common tool that educators and researchers use to examine NNSs' lexis in free writing (Laufer, 1998; Laufer & Nation, 1995; Meara & Bell, 2001; Muncie, 2002). This educational tool was developed by Laufer & Nation in year 1995 and its' purpose was to ensure a particular text is suitable for learners of a specified proficiency. Their original intention then shifted to applying LFP to measure NNS productive vocabulary. Laufer & Nation (1995) state that LFP works by calculating the total word tokens in a text before categorizing the word tokens into a four frequency band list. The four frequency lists comprise of Band 1, Band 2, Academic Word List (AWL) and Not in the lists (NIL).

Band 1 features the most common 1,000 words in English, Band 2 lists out the next most common words in English, AWL consists of 3.3 million words found in 570 word families and NIL covers the less frequent words. AWL by Coxhead (2000) is an updated list of the previous University Word List (UWL) which covers an extensive range of academic vocabulary found in a wide range of educational textbooks compared to UWL.

However, based on a recent study by Hyland and Tse (2007), they found that there are downsides of AWL despite its coverage of 10.6% of the corpus. Discrete words or lexis in the AWL list carry diverse meanings in terms of range, frequency, collocation and

meaning across the academic disciplines. Nevertheless, the AWL has been utilised in the field of teaching and practices over the years and it is found to be useful in helping instructors to develop an overall vocabulary learning goals for English courses. Lastly, LFP or Range programme generates NIL which features words that are neither in Band 1 nor Band 2 (Laufer & Nation, 1995). An example of LFP analysis is displayed in Table 2:

Table 2: Example of Lexical Frequency Profile (LFP) Analysis/Range Programme

Word List	Tokens (%)	Types (%)	Families (%)
One	216 (75.3)	104 (71.1)	84
Two	30 (10.5)	14 (9.9)	10
Three	18 (6.3)	15 (10.6)	14
NIL	23(8.0)	12 (8.5)	?
Total	287	142	108

(Meara, 2005)

The table shows that LFP categorised the four frequency band list as Band One, Band Two, Band Three or AWL and NIL into numbers and percentages of word tokens, word types and word families. On the definition of "tokens", "types" and "word families", Nation (2001) defines tokens as every single word that occurs in a text. Therefore, the total number of words in a written essay will be counted as tokens regardless of the times of occurrence.

Meanwhile, "types" is defined as words that only occur once will be counted. Thus, if a particular word occurs twice in a page, only the first instance of the word will be counted. In sum, types are referred to the total number of different words in a page. A word family covers all the words including inflected and derived forms of a word. The numbers and percentages in the LFP analysis are generated based on three ready-made base word lists of Baseword 1 (the most frequent 1000 words of English), Baseword 2

(the second 1000 most frequent words of English) and Baseword 3 (words that are frequently appear in academic textbooks). Words that are not found in these three base word lists will be categorised in NIL.

To ensure that LFP generates near-to-accurate analysis of learners' productive vocabulary, several conditions need to be considered. To ensure an accurate analysis on the types of words used in the corpus of essays, proper names must be omitted, spelling errors are corrected when inputting a text to be analyzed by LFP and semantically incorrect such as erroneous meaning or incorrect collocation should be omitted, because all these could not be assumed as known knowledge by the students (Laufer, 1998).

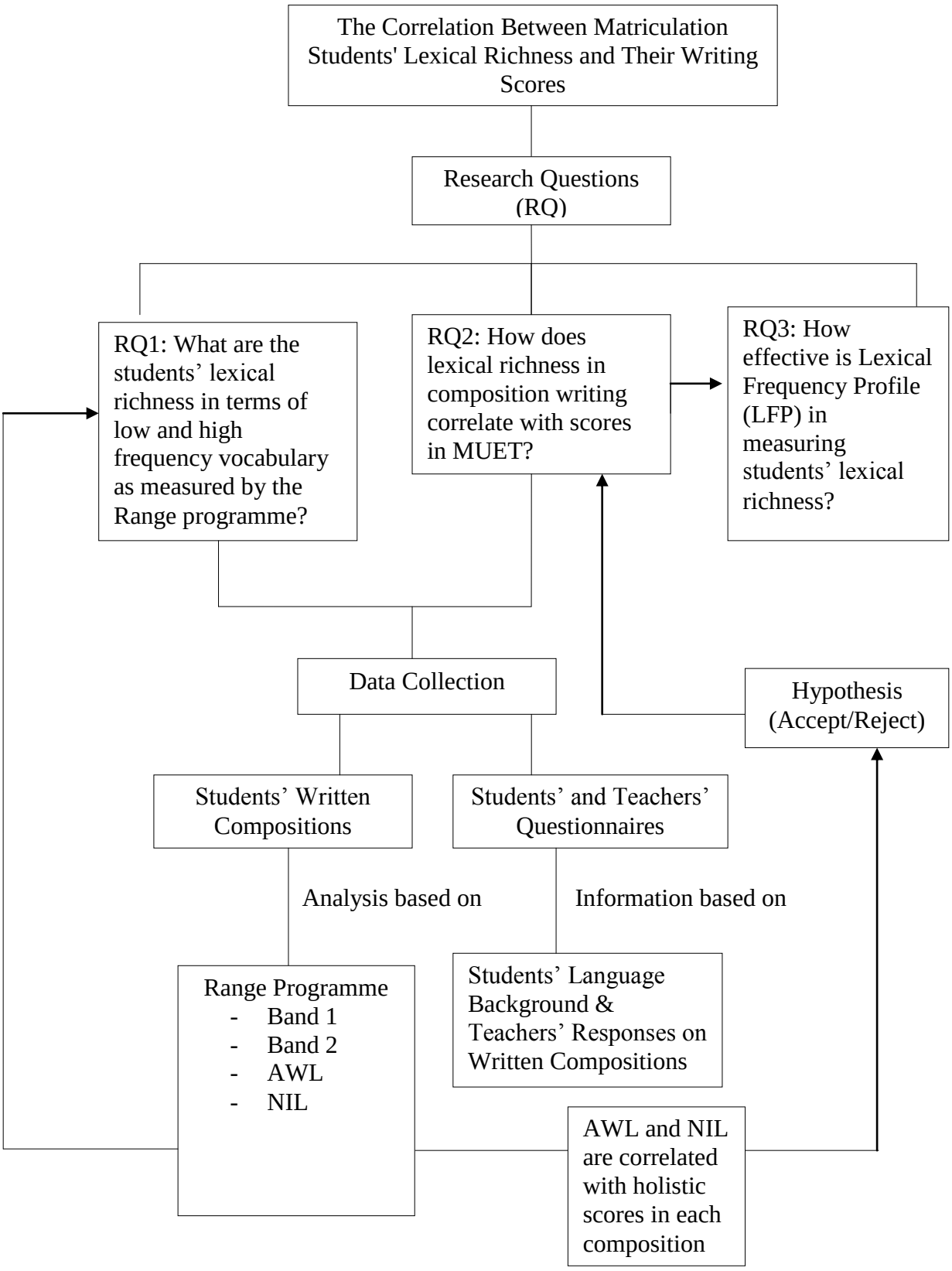
Apart from the conditions above, to ensure LFP analysis be executed well, there are six claims made by Laufer and Nation (1995) on the suitability of LFP for certain research designs that concentrate on group differences such as students of low English proficiency level and distinguishable vocabulary sizes. The consistency length of text for input is also taken into consideration before utilizing the LFP. LFP tool will churn out near-to-accurate results when learners' essays are divided into "rather broadly proficiency levels" (Smith, 2005).

Pokorny (2009) states that learners who use more frequently occurring words in writing signify their limitations in lexical knowledge compared to those who are able to use higher proportion of words that can be found in AWL and NIL lists which imply lexical richness. Laufer & Nation (1995) claim that "lexical richness is a reliable indicator of good quality writing". Li (1997), to validate claims made by Laufer and Nation, examined the lexical richness of a small group of EFL students in Hong Kong using LFP. The results show that LFP is able to distinguish good and weak compositions quite

accurately but it may be inaccurate in distinguishing compositions of average marks (Li, 1997; Smith, 2005).

2.8 Conceptual Framework

Having explained the various elements in this study, the conceptual framework below shows a schematic outline of the variables that underlies this study.



2.9 Summary

This chapter lays out the conceptual framework and review of relevant literature for the study. A detailed literature on the importance of writing in an academic context, the processes in writing and the effects of social, cultural and language backgrounds on writing proficiency are provided in this chapter. Looking at the importance of writing and the frequency of it being given to students in tertiary education as a form of assignment, the elements that make up a good writing are further explored. One of the key elements is lexis and how it affects the quality of writing and teachers' perception are discussed. Studies on lexical richness in written output and its effects on grades have been conducted by researchers using different electronic tools. After a detailed review of the available tools, Lexical Frequency Profile (LFP) is used in this study to find out the correlation between the two independent variables: lexis and holistic scores.

CHAPTER 3: METHODOLOGY

3.0 Introduction

This chapter will discuss the procedures of data collection followed by description and explanation regarding the participants, research instrument and data analysis method.

3.1 Research Design

This study employs a correlational research design through which two types of data are obtained, the lexical richness of the participants and the holistic scores given by the language instructors on the participants' written compositions. These variables will be labelled as quantitative variables and they will then be correlated to find out whether they have a positive or negative relationship using the Range programme. Data will be generated to answer the research questions in this study (Chapter 1). For easy reference, they are given below:

1) What are the students' lexical richness in terms of low and high frequency vocabulary as measured by the Range programme?

2) How does lexical richness in composition writing correlate with writing scores in MUET?

3) How effective is Lexical Frequency Profile (LFP) in measuring lexical richness in writing?

3.2 Data Collection Procedure

This study employs both quantitative and qualitative methods to the data collection and analysis. This subsection will explain the basis for the selection of the focal source of data for this research. The quantitative data are collected from two sources; questionnaires and a sample of compositions from the students. Data collection was first initiated by collecting students' compositions from Pusat Asasi Sains Universiti Malaya (PASUM). Then, the compositions were randomly selected according to the scores.

Questionnaires were distributed to students and experienced language instructors at PASUM. The respondents were required to answer all the questions and the researcher was present to increase the reliability of the answers provided by the respondents. The respondents had the opportunity to clarify any doubts with the researcher.

3.2.1 Pilot Study

A pilot study was conducted to validate the study. The aims of the pilot study were to:

a) ensure the obtainability of data and genuine sampling
b) provide the researcher with ideas and approaches to obtain richer findings in the main study
c) allow an in-depth check of the planned statistical and analytical measures before revamping the methods of data collection and analysis in the main study
d) lessen the number of problems which might arise.

The study was conducted with ten students from a Chinese independent secondary school in Kuala Lumpur. They were asked to write a composition on the "Dangers of

Smoking" in their English Language monthly test. Twenty compositions were collected and these were then typed into the Range programme for further analysis. The Range programme analysed the composition and the results generated by the programme was compiled in a table to show correlation between the students' lexical richness and the scores given by the language instructors. Findings of this pilot study guided the present research in terms of data obtainability, accessibility, convenience and economy of cost and resources into consideration of a study of this nature (Creswell, 2003).

3.2.2 Actual Study

The analysis and results from the pilot study helped to test and prove the reliability and validity of this study particularly for data collection. This study is a replication of a similar one done by Lemmouh (2008), which shows the relationship between students' lexical richness and their holistic scores.

The first objective of this research is to find out the students' lexical richness based on their written compositions. The Range programme was fully utilised to distinguish students' lexical richness in terms of low and high frequency vocabulary in their writing. The second objective of this research is to explore the causal relationship between lexical richness and the holistic scores awarded by the teacher-marker. Lastly, students' language backgrounds are taken into consideration to investigate its effects on determining the quality of writing.

In order to achieve the objectives mentioned above, a total of 100 MUET compositions were obtained from PASUM. The timed compositions were part of the students'

assessment conducted by PASUM for their examination. The papers were obtained from the English Language Coordinator at PASUM.

Firstly, a written letter of permission was provided to Director of PASUM and after the permission was granted, the 100 graded compositions were selected randomly by choosing a few samples from each of the scores range. There were 4 categories of scores range. On average, 20-30 samples for each category of scores were hand-picked by the researcher. For instance, 29 essays were obtained from the range of scores of 20-29, 31 essays from the range of scores of 30 - 39, 32 essays from the category of 40 - 49 and lastly, 8 essays were collected from the category of scores of 50 - 59.

To ensure accuracy of analysis from the Range Program, the Beyond 2000 (B2000) which means only the percentages in AWL/Band 3 and Not in the Lists (NIL), were used and they were served as a way to identify the participants' lexical profile. The rationale for applying the B2000 (the totalled up percentages of AWL/Band 3 and NIL in the 'Types' categories) is to examine these students' use of low frequency words or advanced vocabulary in their free writing.

There are other criteria in utilising the Range programme to measure students' use of low frequency vocabulary. Based on the rules set by Laufer & Nation (1995) to obtain a set of accurate and reliable data, the following steps need to be adhered to when entering the students' compositions into the Range programme:

a) the words used incorrectly are omitted as these words are not produced by the students' mental lexicon
b) the misspelled words are corrected

c) a bracket < > has to be typed in to take out a particular word such as proper nouns

d) a wrong derivative word is not to be considered an error

Apart from using the raw data, questionnaires were also designed for the students and instructors. Students were given a questionnaire focusing on their opinions on the importance of the four skills in language, language background, past results in examinations and other related questions. The students were given verbal instructions on how they can fill up the questionnaires. Once they were briefed, they began answering the questionnaire in the presence of the researcher, who was there to answer any questions or clear any doubts which may arise. This procedure of collecting information on participants' language background took a month, as it was felt that information given by the participants should be as accurate as possible.

Another questionnaire which consists of open-ended questions was given to five instructors of PASUM. They were asked for their opinion on how they mark and grade compositions. Again the presence of the researcher throughout this procedure was to help clarify any doubts if any should arise. On average the respondents took less than ten minutes each to answer the questionnaire.

3.3 Instruments

The instruments used in this study are the written compositions produced by PASUM students (n=100), questionnaires for students (n=50) and questionnaires for teacher-raters (n=5). The matriculation students are students who had studied in secondary schools and gone through the national syllabus. The teacher-raters who participated in the study were either experienced lecturers or experienced and freelance lecturers in PASUM. In this study, the final results are correlated based on the examiners' subjective judgment in marking compositions with the objective analysis of LFP. This study focuses on the correlation between the two variables to explore the effectiveness and reliability of LFP as an indicator of good quality writing based on the measurement of students' lexical richness. Apart from that, questionnaires were administered to students and teachers. The findings from these questionnaires will be used to support the findings from the written composition and also results generated from LFP.

3.3.1 Written Compositions

A total of 100 compositions which were marked and graded by the instructors were obtained from the Head of the English Department. These compositions were taken from the Preparatory English Course 1 examination which was a part of their assessment in preparation for MUET. The examination was held in the first semester of academic year at PASUM.

In the MUET, to test the students' abilities in writing, they were required to write two compositions in one and a half hours. The two tasks are transferring information from a non-linear source to a linear text and continuous writing. For the purpose of this study,

the continuous writing component was chosen because it puts a demand on students' free active vocabulary from their mental lexicon. The question given in the test, "Preserving the forests is the key to saving our planet. What is your opinion on this statement?" required students to give their opinion.

3.3.2 Questionnaires

Other than using written composition as the main source of data, questionnaires were also designed to further obtain information from the students and instructors. Through a random selection using the Microsoft Excel, 50 self-voluntary students were selected to answer the questionnaires. They were required to complete 20 close-ended questions which consist of three parts: personal information, language background and perceptions of English Language (See appendix E).

Although the researcher asked for the respondent's personal information, the information that they provided such as their names were kept confidential. Another questionnaire consists of 10 open-ended questionnaires were designed to acquire information from the instructors. They were asked for the criteria of marking a composition, the perceptions of the importance of vocabulary in writing and others. The rationale of administering this questionnaire is to support the raw data generated by Range programme.

MUET aims to test students on their ability and proficiency in listening, speaking, writing and reading. For the purpose of this study, only writing will be looked at. The students' writing proficiency will be examined and graded based on the five main

criteria: accuracy, appropriacy, coherence and cohesion, use of language functions and task fulfilment in genres such as report, article, letter and essay (Refer to appendix A).

3.4 Participants

Fifty students from the Foundation Studies in Science and Built Environment of PASUM participated in this research by providing the answers for the questionnaires. Besides, another 100 PASUM students written compositions were collected to be analysed. All the respondents have studied English as their second language for at least ten years and were exposed to English when they were studying in primary and secondary schools. Their highest academic qualification was *Sijil Pelajaran Malaysia* (SPM) examination, which they had taken in their fifth form, before entering PASUM to continue with their pre-university education.

Although the respondents are L2 learners, they do not come from various linguistic and cultural backgrounds. Instead they come from only the Malay language speaking background and they are mostly proficient in two languages: their mother tongue Malay and also English as their second language.

Apart from that, five teachers from PASUM were selected for this study. They are currently teaching matriculation students in the Foundation Studies in Science and Built Environment of PASUM. They were experienced NNSs who have been teaching MUET in PASUM for more than two years. The teachers were contacted and arrangement of time was made to meet them personally to obtain the relevant data.

3.5 Statistical Analyses

This section will discuss the rationale of using the statistical inferencing tools and how they are applied in this study. The first one, Range programme, is employed in this study to analyse the text and generate results in numbers and percentages of words and word families from each of the word lists as Lexical Frequency Profile (LFP) analysis. The programme categorised words in each composition into four levels. The four levels are Band 1, Band 2, Academic Word List (AWL) and Not in the List (NIL). This software was downloaded for free from Paul Nation's web site at http://www. vuw.ac.nz/lals/staff/paul_nation/index.html.

Next, two main statistical inferencing analyses, T-test analysis and scatter plot are employed in this study. The dependent-samples t-test aims to explore whether there is a statistically significant change in the mean scores between the holistic scores and profile B2000. The holistic scores are the scores given by examiners for students' writing while B2000 will be attained by adding up the AWL and NIL percentages based on the LFP analyses. The researcher also utilised the scatter plot, a diagram using Cartesian coordinate to show the values for two variables for a set of data which provides a good visual picture of the relationship and helps in interpreting the correlation coefficient between the two variables.

Having introduced the research design and procedures in this chapter, the next chapter will outline these descriptions in greater detail followed by analyses and explanations on the students' lexical knowledge based on their written compositions and information obtained through interviews and questionnaires.

CHAPTER 4: ANALYSIS OF DATA AND DISCUSSION OF FINDINGS

4.0 Introduction

This chapter will report the findings obtained from the data. Results analysed by the Range programme will be discussed in relation to the students' lexical richness and holistic scores. The results from the questionnaires will also be provided.

4.1 The Range Programme

The Range programme is used in this research to explore the correlation between two sets of independent data, which are lexical richness generated by the programme itself, and the holistic scores given by the instructors. This programme has been widely used by other researchers such as Harfitt (1999), Lemmouh (2008), Pokorny (2009) and Abbasian & Parizad (2011) to study students' lexical richness, and the findings from the studies have shown that it is a reliable programme for the intended purpose.

The Range programme was initially aimed at evaluating whether a particular text is suitable for language learners of a certain proficiency level. However, Laufer and Nation (1995) found that it was also beneficial in measuring students' lexical richness or productive vocabulary size. It is a programme which is relatively easy to use where raw text needs to be keyed in the system and a table of lexical frequency will be generated according to four bands. The bands consist of Band 1, Band 2, Band 3 or Academic Word List (AWL) and 'Not in the List' (NIL). This programme is capable of analysing a maximum of 32 texts at a time and this met the scope of this study.

The content of a composition is typed into the Range programme for further analysis and categorisation into the specific bands. The programme will then automatically group the lexis according to the frequency or difficulty of words used. Words that are categorised into Band 1 consist of the most frequent 1000 words of English. Band 2 includes the second 1000 most frequent words of English while Band 3 or AWL cover words that are not found in Band 1 and Band 2 but labelled as words that are of frequently used in the secondary and tertiary levels. Lastly, words that are categorised into the NIL will feature words that are not found in the other three bands. The use of these words is classified as low frequency words, which show that the particular student has a good level of receptive and productive vocabulary.

4.2 Total Types for Categories of Scores

The results generated from the Range programme are presented in Tables 3 - 6. There is a difference in sample size across the four categories of scores: 29 essays from the range of scores 20 - 29, 31 essays from the range of scores of 30 - 39, 32 essays from the category of 40 - 49 and lastly 8 essays were collected from the category of students who scored 50 - 59. The total marks for this written task is 60 marks, and each of the students is numerically identified and labelled in each category of scores.

4.2.1 The Lexical Richness of Students in PASUM

To answer the first research question in this study on '*What are the students' lexical richness in terms of low and high frequency vocabulary as measured by the Range programme?'*, the data were categorised according to the holistic scores of the matriculation students' essays. Adapting from Laufer and Nation study, these texts were

typed into the Range programme as input. Errors detected in students' writing were identified and amended based on certain conditions as discussed in Section 3.2.2.

Table 3: Lexical Profile of Students in the Range of Scores 20 - 29

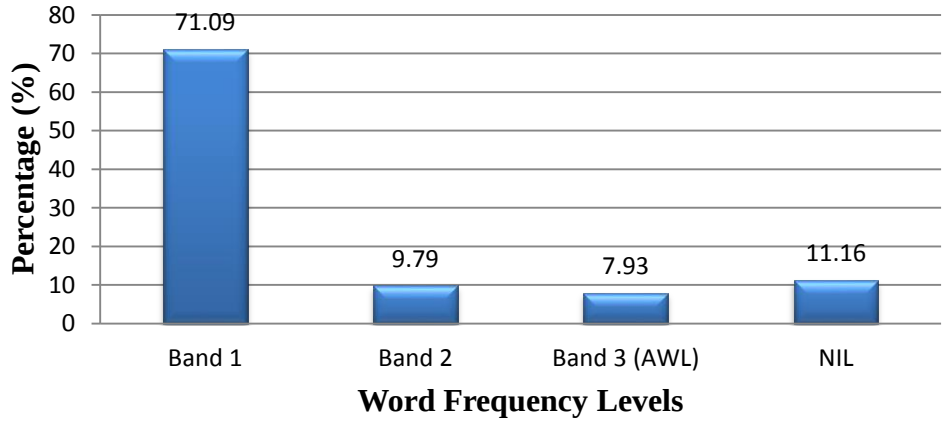
RANGE PROGRAMME	Band 1 (%)	Band 2 (%)	Band3/AWL (%)	NIL (%)	Total Types
Text 1	69.5	10.64	5.67	14.18	141
Text 2	66.82	10.14	10.14	12.90	217
Text 3	72.25	9.57	6.22	11.96	209
Text 4	75.17	12.75	4.70	7.38	149
Text 5	73.96	8.33	9.38	8.33	192
Text 6	59.8	10.29	12.25	17.65	204
Text 7	71.34	7.32	9.76	11.59	164
Text 8	68.02	11.63	8.14	12.21	172
Text 9	66.49	12.23	9.04	12.23	188
Text 10	64.97	8.12	11.68	15.23	197
Text 11	68.42	9.94	9.94	11.70	171
Text 12	65.90	10.14	11.98	11.98	217
Text 13	71.12	9.09	8.56	11.23	187
Text 14	70.28	11.32	6.60	11.79	212
Text 15	69.59	10.14	7.37	12.90	217
Text 16	69.94	6.94	12.14	10.98	173
Text 17	71.37	9.54	9.54	9.54	241
Text 18	68.52	11.11	5.56	14.81	162
Text 19	70.85	11.56	7.04	10.55	199
Text 20	76.19	8.84	7.48	7.48	147
Text 21	70.35	9.88	6.98	12.79	172
Text 22	83.01	7.84	5.23	3.92	153
Text 23	77.36	10.06	3.14	9.43	159
Text 24	71.11	10.00	8.89	10.00	180
Text 25	70.63	9.38	6.88	13.13	160
Text 26	74.48	7.59	6.21	11.72	145
Text 27	71.43	11.43	6.19	10.95	210
Text 28	77.22	10.56	5.56	6.67	180
Text 29	75.76	7.79	7.79	8.66	231

Based on Table 3, students who scored 20 - 29 used a minimum of 141 to a maximum of 241 total types or 184 total types on average. In the Range analysis, the words in a composition will be categorised into tokens, types and families (see appendix J). A type refers to any lexis which occur only once, regardless of how many more times it might

occur in a text while a token, on the other hand refers to lexis which occur in a text regardless of its first or *n*th occurrence (Goodfellow et al., 2002).

The data in Table 3 show that the students in PASUM utilised more high frequency words in writing. They tend to use more high frequency words as found in Band 1 compared to Band 2, Band 3 (AWL) and NIL. This might indicate that this group of students is less proficient writers. A more detailed analysis is done by counting the average percentage of Band 2, AWL and NIL, and this is shown in Figure 1.

Figure 1: The Average Word Frequency Levels of Texts in Category 20 -29



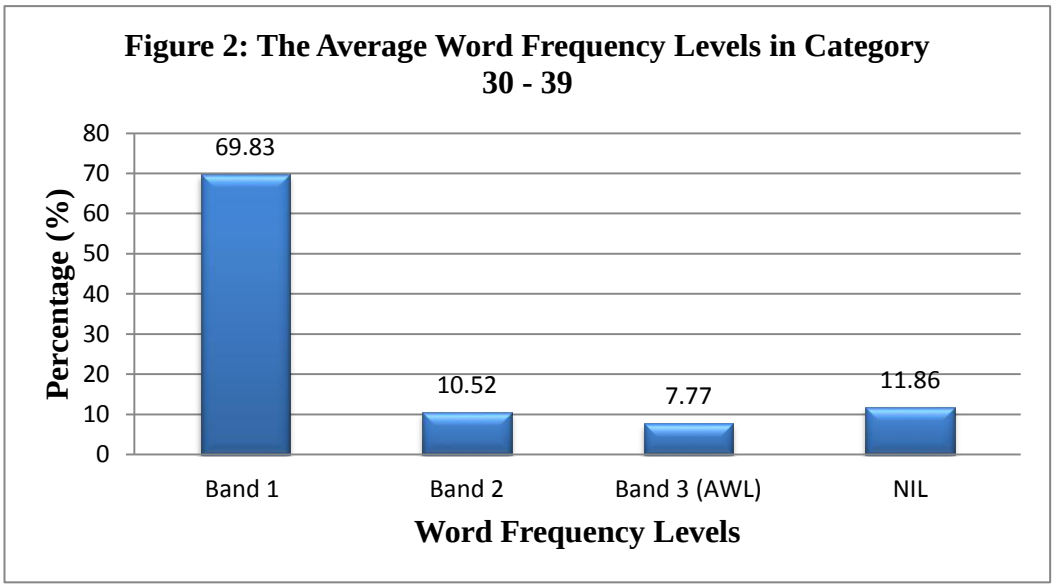
The data in Figure 1 shows the average word frequency levels in percentage of students who scored 20 - 29 marks in their writing. Students' active vocabulary encompasses words in Band 1 which average 71.09% in comparison to 9.79% in Band 2, 7.93% in Band 3 and 11.16% in NIL. The data reveal that students used more NIL words than words in the three other categories (Band 1, Band 2 and AWL). The next table (Table 4) below shows the lexical profile of students generated for the next 31 texts in category of 30 – 39 scores.

Table 4: Lexical Profile of Students in the Range of Scores 30 - 39

RANGE PROGRAMME	Band 1 (%)	Band 2 (%)	Band 3/AWL (%)	NIL (%)	Total Types
Text 1	68.35	9.49	8.86	13.29	158
Text 2	77.27	9.74	5.84	7.14	154
Text 3	79.47	11.05	4.21	5.26	190
Text 4	66.04	11.19	6.34	16.42	268
Text 5	68.16	11.73	6.70	13.41	179
Text 6	66.67	11.11	10.10	12.12	198
Text 7	73.81	10.48	4.76	10.95	210
Text 8	69.83	9.48	6.90	13.79	116
Text 9	75.90	11.28	5.13	7.69	195
Text 10	65.50	9.36	11.70	13.45	171
Text 11	71.43	7.98	7.98	12.61	238
Text 12	80.00	8.39	5.81	5.81	155
Text 13	71.58	12.02	7.65	8.74	183
Text 14	76.40	4.97	8.70	9.94	161
Text 15	66.52	9.82	7.59	16.07	224
Text 16	69.54	10.66	8.63	11.17	197
Text 17	64.04	8.87	10.34	16.75	203
Text 18	64.13	10.76	12.11	13.00	223
Text 19	63.48	9.57	11.74	15.22	230
Text 20	71.09	11.33	8.20	9.38	256
Text 21	69.47	11.45	6.49	12.60	262
Text 22	64.63	11.59	14.02	9.76	164
Text 23	69.90	10.19	6.31	13.59	206
Text 24	60.10	14.42	9.13	16.35	208
Text 25	74.15	9.76	6.34	9.76	205
Text 26	71.63	11.35	4.96	12.06	141
Text 27	68.25	12.17	9.52	10.05	189
Text 28	69.86	13.40	2.87	13.88	209
Text 29	75.31	11.11	3.70	9.88	162
Text 30	71.07	10.69	6.92	11.32	159
Text 31	61.39	10.76	11.39	16.46	158

From Table 4, it can be seen that the students used a minimum of 116 to a maximum of 268 total types or 192 total types on average. Students in this category are able to use more advanced words to express their ideas in writing compared to students in the category of 20 - 29 scores. This is evident in the number of total types counted by the Range programme, where the maximum used by a student was 241 total types in category 20 - 29 scores as compared to 268 total types in this category.

Figure 2 below shows the data of the average percentage of students who scored 30-39 marks in their writing.



In terms of the words usage, the students used an average of 69.83% in Band 1 but this figure is reduced quite drastically to 10.52% in Band 2. The students used a total of 7.77% words in Band 3 and 11.86% in NIL. This indicates that there is a drastic decrease of percentage from Band 1 to Band 2, Band 3 (AWL) and NIL; hence the pattern is similar to the data obtained from the 20 - 29 categories (Figure 1).

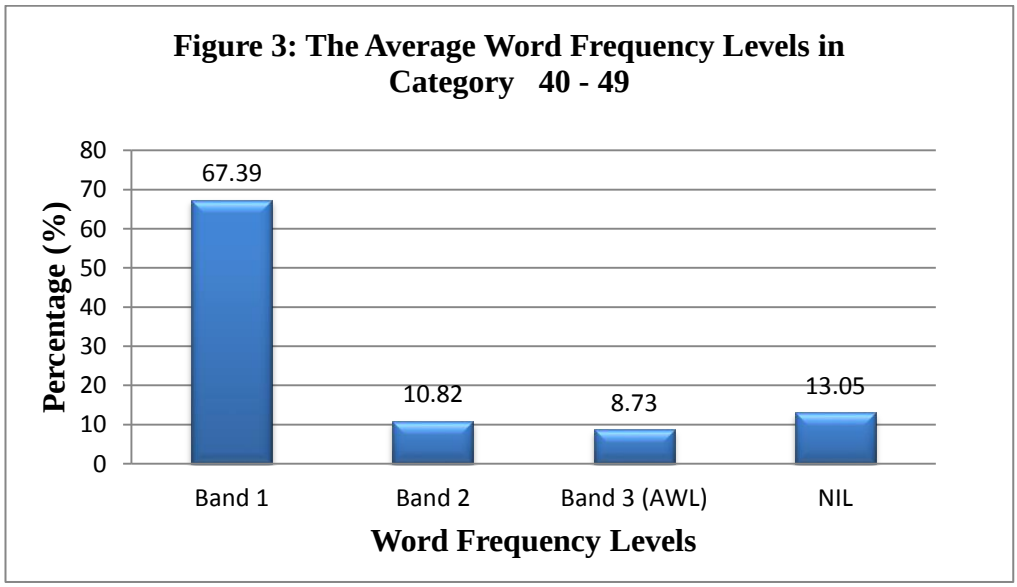
The data in Table 5 shows a total of 30 students' individual lexical profile in the category of 40 - 49. The Range programme analysed the lexis in each of the texts written by the matriculation students and the data reveals the results as shown in the following table:

Table 5: Lexical Profile of Students in the Range of Scores 40 - 49

RANGE PROGRAMME	Band 1 (%)	Band 2 (%)	Band 3/AWL (%)	NIL (%)	Total Types
Text 1	65.64	10.77	10.77	12.82	195
Text 2	72.18	12.41	7.14	8.27	266
Text 3	65.07	12.92	6.70	15.31	209
Text 4	71.21	9.73	8.56	10.51	257
Text 5	70.81	10.27	4.86	14.05	185
Text 6	75.43	12.00	5.71	6.86	175
Text 7	66.37	7.52	7.96	18.14	226
Text 8	68.67	7.23	12.05	12.05	166
Text 9	60.12	9.51	11.35	19.02	326
Text 10	66.88	11.88	9.38	11.88	160
Text 11	64.63	12.66	9.61	13.10	229
Text 12	64.73	10.85	10.85	13.57	258
Text 13	69.91	7.96	11.95	10.18	226
Text 14	72.50	10.50	7.00	10.00	200
Text 15	66.92	9.51	7.60	15.97	263
Text 16	64.53	10.26	10.26	14.96	234
Text 17	68.91	13.47	6.22	11.40	193
Text 18	69.49	9.60	9.04	11.86	177
Text 19	71.43	9.01	7.45	12.11	322
Text 20	63.82	12.06	10.05	14.07	199
Text 21	65.65	9.86	9.86	14.63	294
Text 22	64.47	10.66	11.68	13.20	197
Text 23	66.07	14.29	7.14	12.50	168
Text 24	70.00	11.00	7.00	12.00	200
Text 25	61.76	9.93	11.40	16.91	272
Text 26	61.54	10.41	12.67	15.38	221
Text 27	63.55	10.28	12.15	14.02	214
Text 28	74.31	11.07	3.95	10.67	253
Text 29	64.20	12.84	5.06	17.90	257
Text 30	67.89	11.93	5.96	14.22	218
Text 31	65.14	12.00	9.14	13.71	175
Text 32	72.78	12.03	8.86	6.33	158

From Table 5, it can be seen that the students used a minimum of 158 to a maximum 326 of total types or 216 total types in average. These students seem to be able to use a variety of words to express their ideas in writing the content of the essay, compared to students who are in the previous two categories.

Figure 3 below shows the average percentage of students' lexical profile in category 40 - 49. They used an average of 67.39% in Band 1, 10.82% in Band 2, 8.73% in Band 3 and 13.05% in NIL.



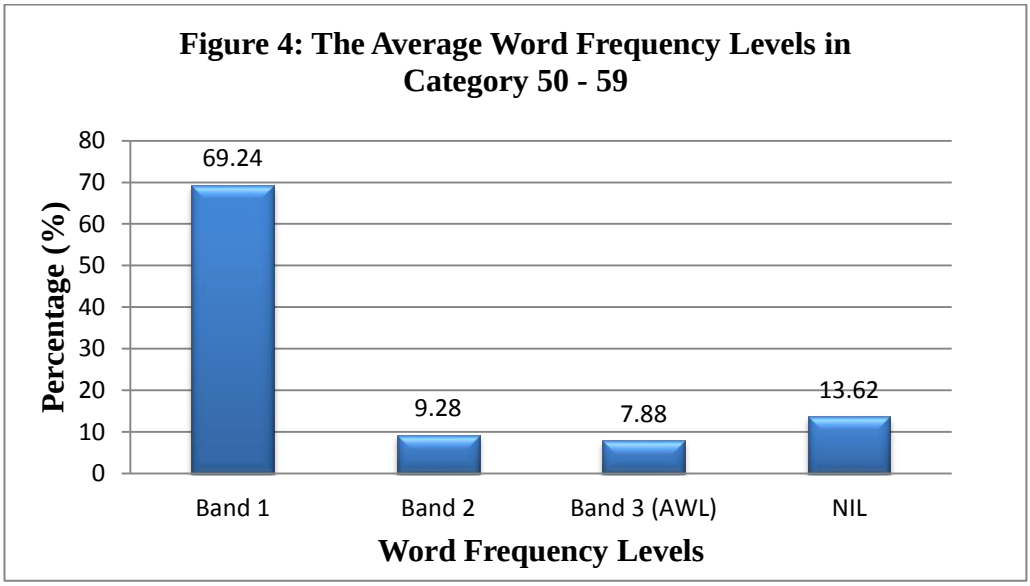
In compiling each student's lexical profile in the next category, Table 6 shows the category of students in the range of score of 50 - 59. The data reveal a total of 8 students' individual lexical profile as generated by the Range programme.

Table 6: Lexical Profile of Students in the Range of Scores 50 - 59

RANGE PROGRAMME	Band 1 (%)	Band 2 (%)	Band 3/AWL (%)	NIL (%)	Total Types
Text 1	70.54	8.30	9.13	12.30	241
Text 2	75.90	8.72	5.64	9.74	195
Text 3	75.22	9.57	5.65	9.57	230
Text 4	68.32	5.90	10.56	15.22	322
Text 5	65.00	10.45	10.45	14.09	220
Text 6	74.18	8.79	6.04	10.99	182
Text 7	65.60	9.17	6.88	18.35	218
Text 8	59.22	13.27	8.74	18.77	309

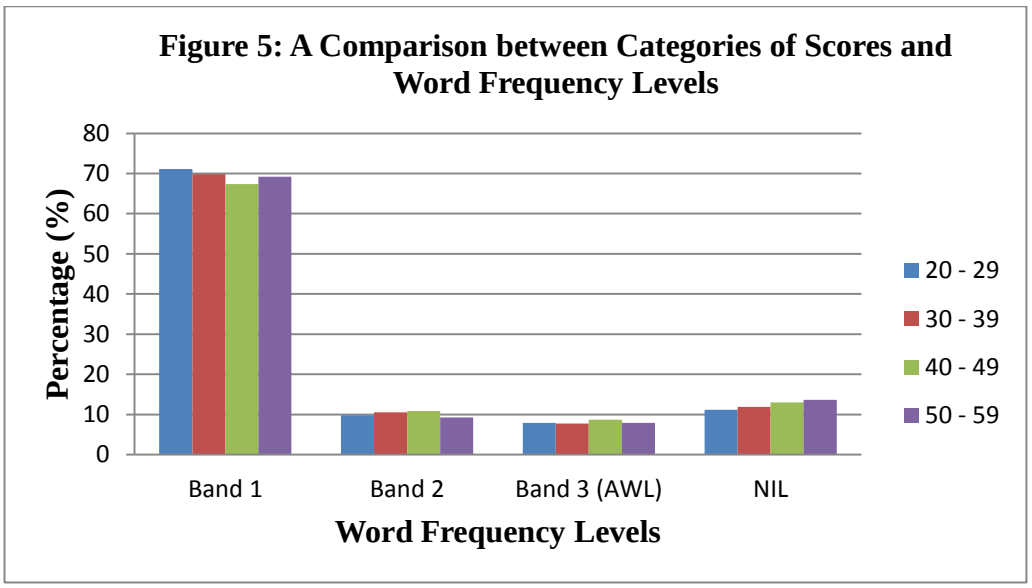
Based on Table 6, it is seen that the students utilised a minimum of 182 to a maximum of 322 total types or 239 total types in average. Students in this category appear to be

able to use more advanced vocabulary and are able to activate their mental lexicon to further express their ideas and refine their writing, which might be the possible explanation of their achievement of scores in their writing compared to those in other categories. To have a clear view of the students' lexical profile, Figure 4 below displays the average percentage of the four frequency word levels.



Based on the analysis done on the eight students' written compositions, they used a minimum of 182 to a maximum of 322 of total types or 239.62 on average, which is the highest total types among the four categories. In terms of the usage of words, students used an average of 69.24% in Band 1 and 9.28% in Band 2, 7.88% in Band 3 and 13.62% in NIL. This result shows similar characteristics as the ones in other categories, where the percentages of Band 1 and Band 2 are higher than Band 3 (AWL) and NIL.

In a further exploration of students' lexical richness, a comparison between students' usage of low and high frequency vocabulary was analysed and the results are presented in Figure 5.



The figure shows a comparison between the four categories of scores 20 - 29, 30 - 39, 40 - 49 and 50 - 59 with the four word frequency levels. For the first 1000 most frequent words in English, the data shows that students' in the category of 20 - 29 scores obtained the highest, i.e., 71.09%, followed by 69.83% for 30 - 39, 67.39% for 40 - 49, and lastly a slight increase is seen for the 50 - 59. Those in the category 20 - 29 recorded the highest number of words used in Band 1 compared to the other categories.

The data for the next 1000 most frequent words in English or Band 2 in Figure 5 shows a drastic drop of percentage compared to Band 1. The first category recorded 9.79%, 10.52% in the second category, 10.82% in the third category and lastly 9.28% in the last category. It is obvious that these results display a big difference of percentage between Band 1 and Band 2.

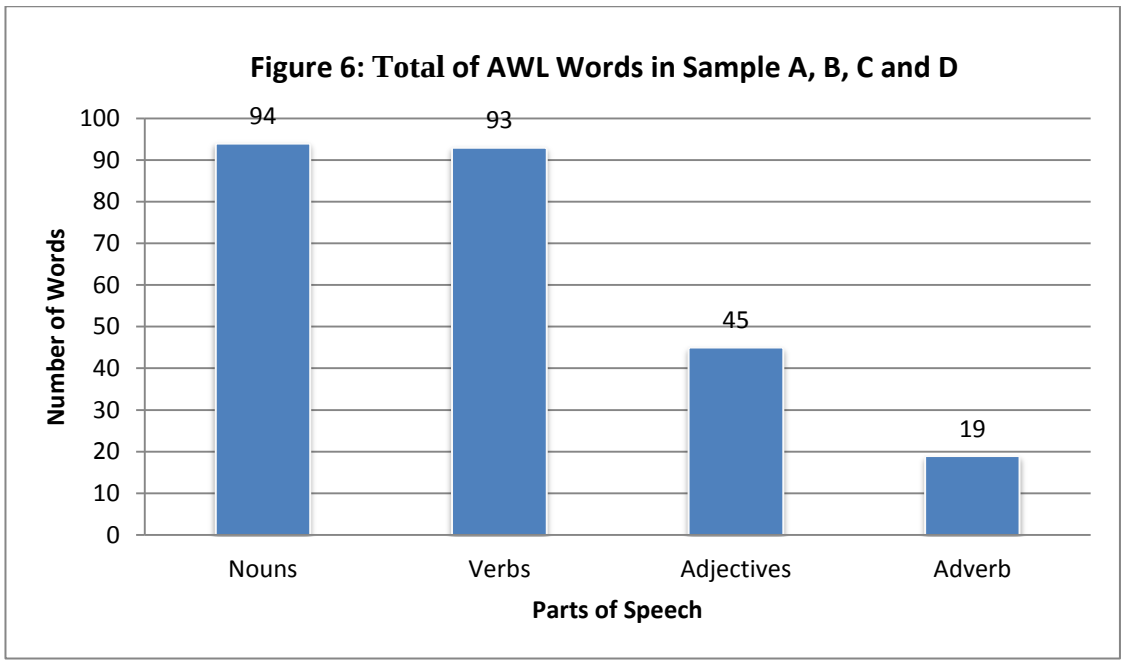
The results show that a majority of the matriculation students in the category of 20 - 29 have limited vocabulary. The bigger amount of high frequency words used in composing the essay has affected the writing scores. It might be due to the limited vocabulary in the matriculation students' mental lexicon which indicates that a majority

of them are exposed to high frequency words in their everyday life. These might have come from various sources, such as through conversation with lecturers and friends, instead of advanced vocabulary found in the academic reference books. The possibility of students who have only activated high frequency words compared to advanced words might be due to their low input of English words. Across all the four categories of scores in Band 1 and 2, it is clear that the Range programme generated a set of data that shows most of the students have limited vocabulary.

4.2.2 Categorisation of AWL and NIL Words into Parts of Speech

The data is further analysed to examine the words categorised in the Academic Word List (AWL) by the Range programme based on the students' compositions. The category of scores 20 - 29 will be labelled as Sample A, 30 - 39 as Sample B, 40 - 49 as Sample C and 50 - 59 as Sample D for ease of identification and explanation. All the words from each composition in the sample will be taken and each sample consists of words taken from the respective category of marks given by the teacher-rater. The words are taken from the end result of the Range programme which is distributed according to their frequency level.

Words from the Academic Word List (AWL) were extracted from the individual analysis of the Range programme before they were classified into four main parts of speech, which consist of nouns, verbs, adjectives and adverbs. The categorisation of words into these respective parts of speech is performed based on the syntactic context. For example, the word 'cycle' which was found repeatedly in the same parts of speech will not be taken into account. Figure 6 displays the total number of AWL words across the four samples.

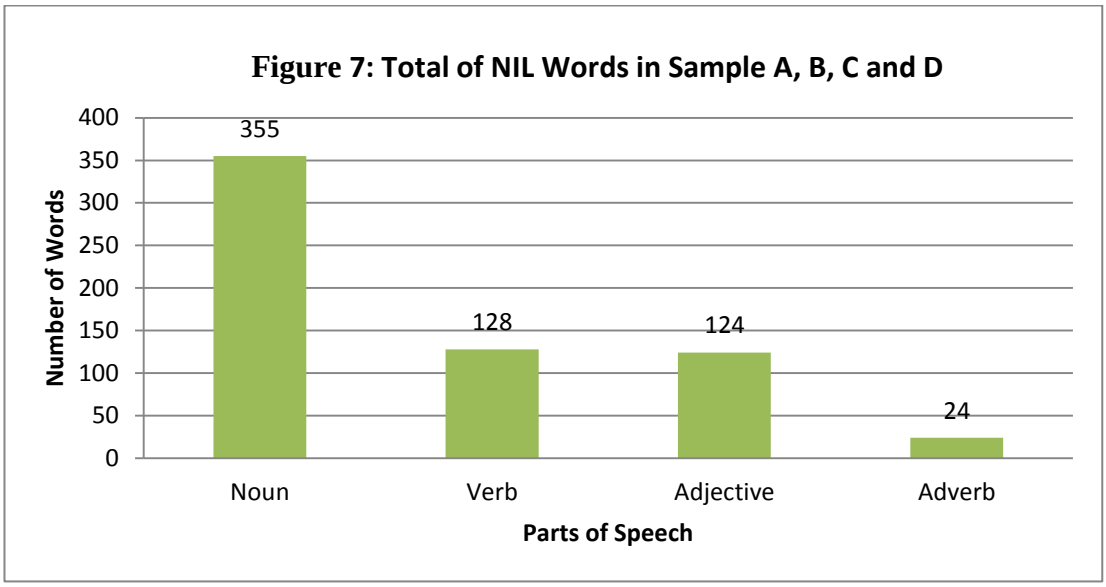


Based on the analysis of Academic Words List (AWL) in Figure 6,matriculation students used more nouns than verbs, adjectives and adverbs. Nouns found in the AWL or Band 3 is recorded as the most, followed by the other three parts of speech. There are 94 nouns out of a total of 252 AWL words, which show a small difference between verbs (93) and a big difference from adjectives (46). The use of adverb forms in the compositions across the four samples is found to be the least, whereby only 19 adverbs were used among the 100 compositions.

Due to the fact that noun forms are recorded as the highest in AWL words, it might be worth conducting a detailed study on the 94 nouns as it might provide some insights into the nature of the AWL or low frequency words. The noun forms in written compositions are regarded as high frequency words and its high usage among matriculation students might reveal that they probably learned and utilised the nouns quite often at some point in their life, be it for social or academic purposes. Referring to appendix B, nouns recorded in the AWL word list are *area, benefit, conclusion, cycle,*

job, items, role, sum, and others.However, these noun forms need to be affirmed for its frequency in their textbooks of different subjects.

Looking not only at the Academic Word List (AWL), the low frequency words also include words found in 'Not in The List (NIL). Figure 7 below shows the total number of words in NIL across the four samples.



According to the categorisation of NIL words into parts of speech in Figure 7, it is apparently a norm for students to utilise more nouns in their writing compared to the other three parts of speech. The words categorised in AWL and NIL show that students were inclined to use nouns in their compositions. Based on the data in Figure 7, nouns in NIL formed the biggest cluster compared to verbs, adjectives and adverbs. A total of 355 nouns out of 631 NIL words were found in the 100 compositions. The number of nouns outnumbered the total of verbs (128), adjectives (124) and adverbs (24).

It may be inferred that the students have a high tendency to use a variety of noun forms, such as the neutral, singular, plural and proper nouns, in their writings. Nouns are

naming and identifying words which demand little or no complex processing on their mental lexicon. They might be more familiar in using nouns to name and identify things which are a common practice in most academic subject classes. In a more traditional teaching context in Malaysian classrooms, nouns are the first part of speech that the students are exposed to.

Although the students used a variety of noun forms, neutral noun forms, and either singular or plural noun forms, are taken into consideration. Repetitive words found in the same sample will only be selected and recorded once regardless of its form as singular or plural nouns. For example, between the singular noun form ‘consumer’ and plural noun form ‘consumers’, only one of the words will be recorded. The probability of the greater use of noun forms in writing might be due to the prescribed English Language teaching syllabus in the country. Language instructors who adhere to the English curriculum specifications set by the country’s Ministry of Education will be more likely to teach nouns in the first English lesson because it is specified as the first component under grammar. Thus, the chances of students being exposed to noun forms are considered relatively high, and they become familiar with the use of this part of speech in writing and speaking.

Pedagogically, the matriculation students seem to be more competent in using a variety of noun forms in naming words rather than describing words. It is in the researcher’s expectation that the basic component in the students’ writings will be mainly from noun forms. Apart from the fact of nouns being the first item to be taught in the English curriculum specifications, the high frequency usageof noun forms in essay might be due to the role of nouns. The role of noun forms in a text is to bring out the primary meaning of a sentence by indicating an object and a subject in a sentence. The inclination to use

nouns more than the other parts of speech also implies that the students' mental productive lexicons are capable in retrieving nouns to effectively name objects and subjects in a sentence.

According to the detailed analysis of word by word done on NIL, it is found that the words categorised in this level are sophisticated ones, such as *methane, chlorofluorocarbon, nitrogen, carbon dioxide, carbon monoxide,phytoplankton, chlorofluorocarbon,* and *tuberculosis.* These words are labelled as low frequency words which are not found in any other levels of frequency using the Range software. This implies that these words may have stretched beyond the 3,100 words of 570 word families in scholastic texts. Many of the words categorised in the NIL word list are words related to pollution, environment and diseases: for example *mammals, reptiles, algae, cancer, cataract, marine, ozone, cheetah, ecosystem,* and *organisms.* This is obvious as the composition topic is on environmental issues and thus; to a certain extent, does influence their lexical choices.

Apart from that, NIL words found in the 100 compositions such as *carbon dioxide, carbon monoxide* and *traffic jam* are analysed separately (as two words) and categorised into the Not in The List (NIL) by the Range programme. This is one of the disadvantages stated by several researchers such as Smith (2005), Calzolari et al. (2002) and Wiktorrson (2001). Calzolari et al (2002) termed it as 'multi-word expressions' (MWEs)', where it is considered as distinct, but is related to fixed or semi-fixed phrases, compound words, support verb forms, idioms, phrasal verb collocations and others. Generally, all of these phenomena may be briefly termed as "*a sequence of words that acts as a single unit at some level of linguistic analysis*" according to Calzolari et al (2002).

Besides, attention is not only focused on the use of noun forms in the students' writings but also on the verb forms. The verb forms recorded the second highest after noun forms in the students' written output. Verbs are used to show the action carried out by a subject to make sentences in the writing more vivid. However, if compared to other forms, students are found to be more competent in using the noun forms in their writing.

Adjectives and adverbs are generally used less in all the samples. In Figures 6 and 7, adjectives and adverbs play a minor role in students' compositions. The data in Figure 6 show that students used some adjectives (45) and adverbs (19) which are categorised in the AWL word list while some used have slightly more adjectives (124) and adverbs (24) in the NIL word lists. Based on this analysis, it appears that the students have difficulty to effectively use the two parts of speech in a sentence compared to the use of noun and verb forms. The lack of adjective and adverb forms in a text may denote that students are fairly weak in describing and articulating their ideas on the topic given.

This research found that the 100 sample of compositions written by PASUM students show a tremendous use of high frequency words (words categorised in the Band 1 as generated by the Range programme) and a relatively low percentage in the use of Band 3 (AWL) and NIL words. Besides, the students found to be more competent in the use of noun forms compared to verb, adjective and adverb forms as shown in Figure 6 and Figure 7.

4.3 The Relationship between Lexical Richness and Holistic Scores

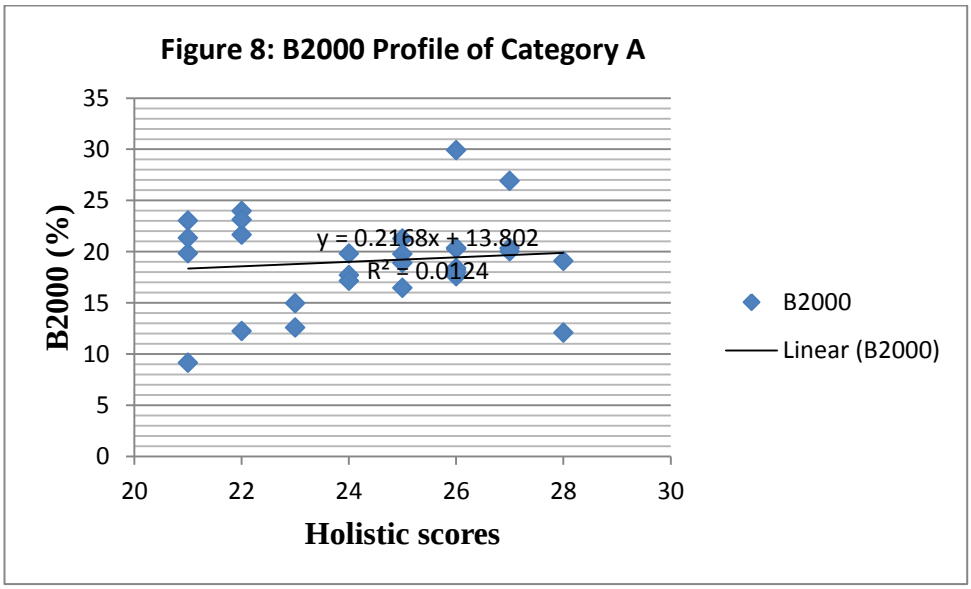
The words found in the matriculation students' writings are categorised into AWL and NIL before sorting the words to their parts of speech in Figure 6. In an attempt to

answer the second research question *"How does lexical richness in composition writing correlate with writing scores in MUET?"*, the impact of low frequency words on the holistic ratings of scores in compositions were analysed by utilising a condensed profile named Beyond 2000.

This profile measures the number of words listed in AWL and NIL: words that are not listed in Band 1 and Band 2. To answer this research question, AWL and NIL of each student in the category is totalled and compared against the holistic scores attained in the essays. The rationale for adding AWL and NIL percentages is that the higher number of B2000 words used in a text, the richer the lexical knowledge of a student.

In an attempt to examine the correlation between students' lexical richness and holistic scores in MUET writing, correlation coefficient and a scatter plot are employed to validate the relationship between the two variables. The correlation coefficient of 0.1112 shows there is a significant difference between lexical richness and holistic scores. To measure how the two variables are related and move together, a scatter plot of B2000 profile in Category A students (20 - 29 scores) is displayed in Figure 8.

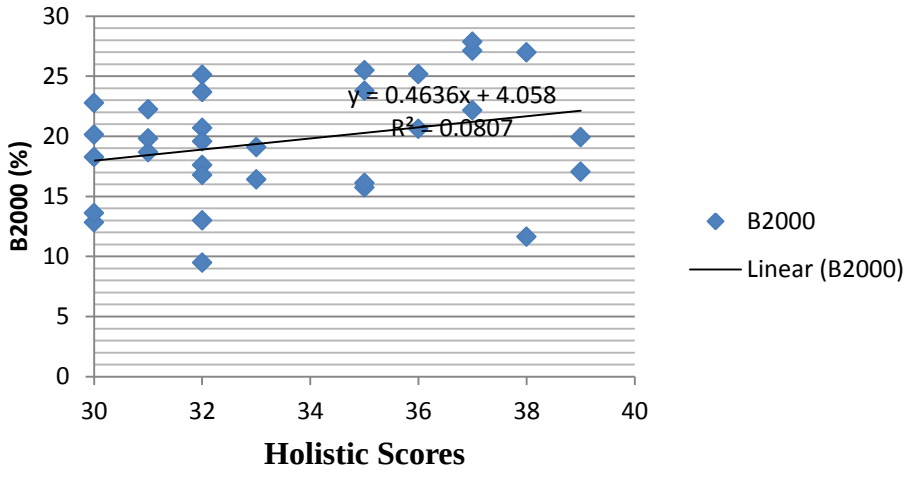
Figure 8: B2000 Profile of Category A



The scatter plot shows the relationship between B2000 and holistic scores of Category A. It reveals that a positive correlation exists between the two variables where the patterns of dots slope from lower left to upper right. The equation *y=0.2168x + 13.802* denotes that the variation of *y* depends on the variation of *x*.Besides, a run of T-test to determine whether there is a statistically significant change in the mean scores between the holistic scores and profile B2000 found that the P-value is 6.87E-31. The use of both statistical analyses in this study can be summed up that the null hypothesis (H0) is rejected. It also indicates that there is a relationship between the use of advanced words (B2000) and the holistic scores given by teacher-raters.

In Category B, the same method of using correlation coefficient and scatter plot is used to show the relation between the variables. The correlation coefficient of 0.2840 shows a statistically significant difference between advanced vocabulary and holistic scores given by the teacher-raters. To measure how the two variables are linked and move together, a scatter plot of B2000 profile in Category B students (30 - 39 scores) is displayed in Figure 9.

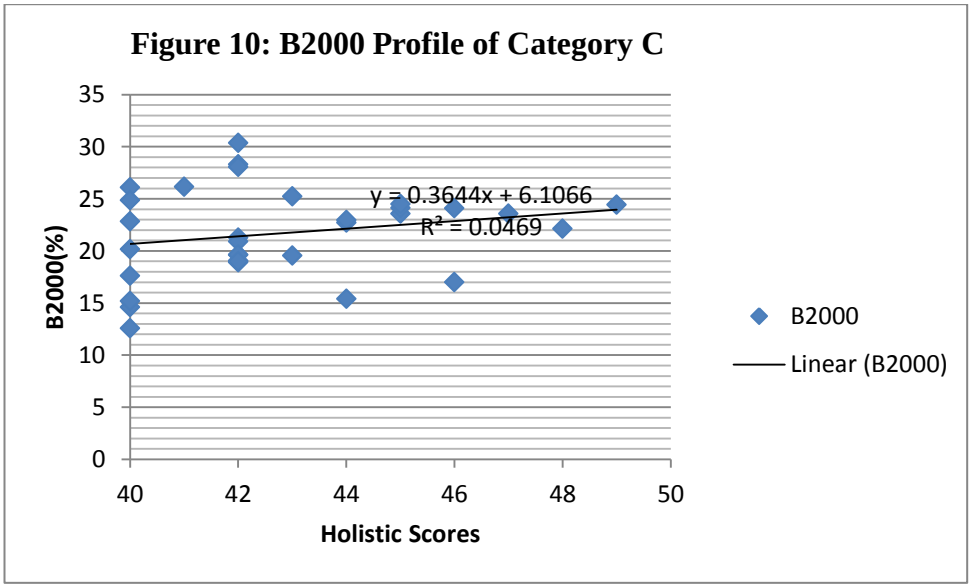
Figure 9: B2000 Profile of Category B



The scatter plot shows the relationship between B2000 and holistic scores of Category B. It shows a positive correlation between the two variables. The equation *y=0.4636x + 4.0548* signifies that variation of *y* depends on the variation of *x*.Besides, a run of T-test to determine whether there is a statistically significant change in the mean scores between the holistic scores and profile B2000 found that the P-value is 6.87E-31. Both the statistical analyses prove that the null hypothesis (H0) is rejected. It denotes that there is a relationship between the use of advanced words (B2000) and the holistic scores given by teacher-raters.

To examine the B2000 profile in Category C, the correlation coefficient is calculated to show how the variables are correlated with each other. The correlation coefficient of 0.2166showing the use of advanced words in a text is significantly correlated to the holistic scores graded by teacher-raters. In line with the purpose of finding out how the variables are related and the effect of each variable on one and another, a scatter plot of B2000 profile in Category C students (40 - 49 scores) is presented in Figure 10.

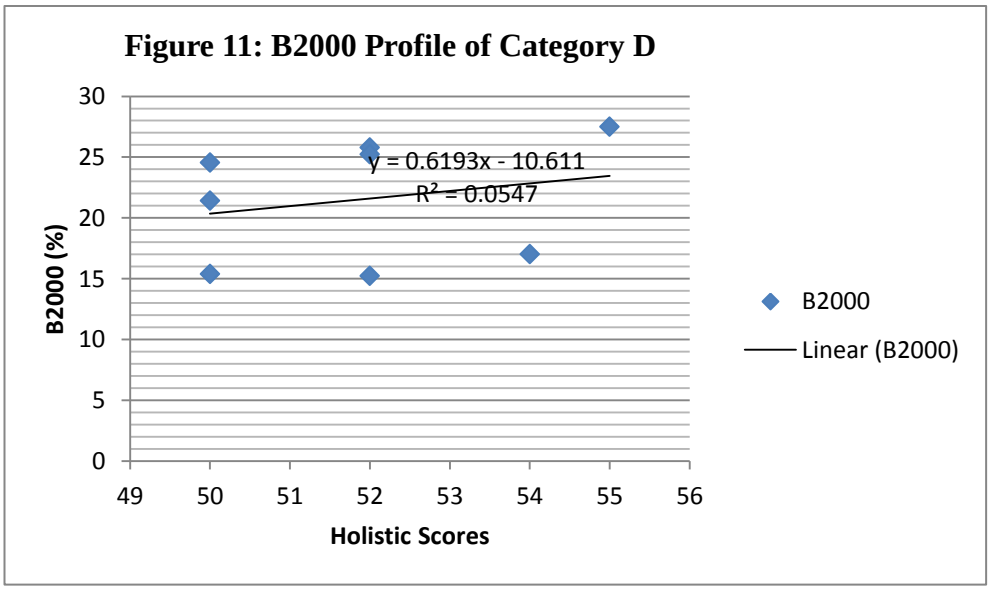
Figure 10: B2000 Profile of Category C



The scatter plot exemplifies the relationship between B2000 and holistic scores of Category C. It indicates positive correlation between the two variables. The equation *y=0.3644x + 6.1066* signifies that the variation of *y* depends on the variation of *x*.Besides, a run of T-test found that the P-value between the mean of holistic scores and profile B2000 is 6.87E-31. Using the scatter plot and T-test analyses, it can be summed up that the null hypothesis (H0) in this study is rejected and it implies that there is a relationship between the two variables.

In Category D, the B2000 profile is analysed to find the correlation of the variables. The correlation coefficient is 0.2338 and shows that it is statistically significant, where the presence of low frequency words in a text is positively related to the holistic scores. To measure how the lexical richness and scores are related to each other, a scatter plot of B2000 profile in Category D students (50 - 59 scores) is shown in Figure 11.

Figure 11: B2000 Profile of Category D



The scatter plot above displays the relationship between B2000 and holistic scores of Category D. It indicates positive correlation between the two variables. The equation *y=0.6193x + 10.611* indicates that the variation of *y* depends on the variation of *x*.

Besides, the P-value between the mean of holistic scores and profile B2000 is 6.87E-31. Based on the scatter plot and T-test analyses, it can be deduced that the null hypothesis (H0) is rejected and it implies that students' lexical richness has certain effects on the holistic scores.

To sum up, to answer the second research question, the finding shows a correlation between the two independent variables: the use of advanced words and holistic scores. Unlike Lemmouh's study (2008) which revealed that there is no relationship between lexical richness and the scores given by instructors, this study has found that there is a positive correlation which signifies the importance of advanced vocabulary in determining the scores of written output.

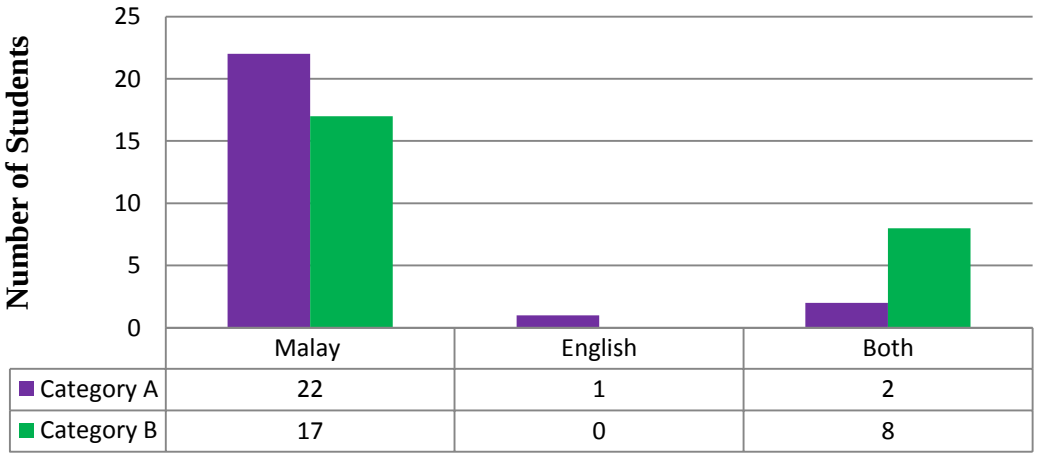
It is evident that the equation in each category is different and it actually refers to how elasticity of the data changes according to the equation. The most extreme and most elastic of the equation are shown in Figure 11: the B2000 Profile of Category D students. The equation $y=0.6193x + 10.611$ in Category D students exemplifies that it has the most elasticity when the use of advanced vocabulary increases and thus this highly affects the scores. However, it is obvious that the sample from Category D is the smallest. There are only eight pieces of writing, however it is statistically found to be valid.

Although lexical richness makes up a major part of a written text, it can be predicted that a strong foundation in English language plays a significant role in determining the students' scores. The factors that governed the students' lexical richness might be related to their language background. The researcher investigated further by analysing the same distributed questionnaire on the language background.

In the questionnaire, the students were asked to provide information such as years of studying English language, number of languages they speak, their SPM and recent MUET grades, choice of language use with family, friends, and lecturers, their exposure to English reading materials, and their opinions on the most difficult language skill. To find out the relationship between the students' language background and their scores in their MUET written examination, the students-respondents in this research will be categorised into two distinct categories: the first category of students (Category A) who scored less than 30 marks, and the second category of students (Category B) who achieved more than 30 marks for Question 2 in the MUET Paper 4.

Figures 12, 13 and 14 will illustrate the language that the two categories of students use to communicate with their family members, friends and lecturers.

Figure 12: The Languages Used by Matriculation Students with Family Members



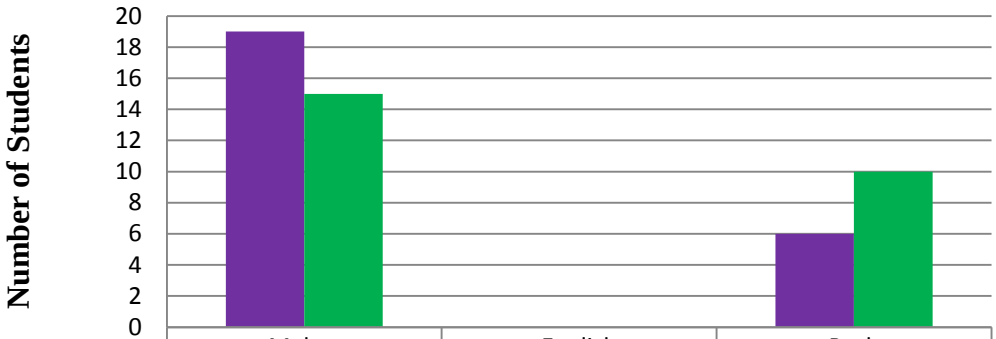
	Malay	English	Both
Category A	22	1	2
Category B	17	0	8

Figure 12 shows the students' responses to the first question *"Which language do you use to communicate with your family members?"*.The respondents stated that they use Malay, English and also both languages. In Category A, 22 out of 25 students speak purely Malay with their family members, and according to them, the reason is Malay is

their mother tongue. They are exposed to this language since birth and they are obviously more comfortable using it with their family members as it creates close bonding. Only 1 out of 25 students selected English with family members, while only two students used both languages at home. On the other hand, 17 out of 25 students in Category B chose Malay as their primary language to use with their family members. None of the students in this category used solely English at home. However, the data show that there are 8 students who are proficient in using both languages to communicate with their kin.

Figure 13 below shows the language choice of matriculation students when they are communicating with their friends.

Figure 13: The Languages Used By Matriculation Students with Friends



	Malay	English	Both
Category A	19	0	6
Category B	15	0	10

It is clearly shown that matriculation students prefer to use their mother tongue when they are talking to their friends. In Category A, 19 out of 25 students used Malay with their friends and 6 of them chose to use both English and Malay to communicate with their friends. Furthermore, only 15 out of 25 students in Category B chose Malay as a medium of communication with their friends while 10 of them enjoy communicating

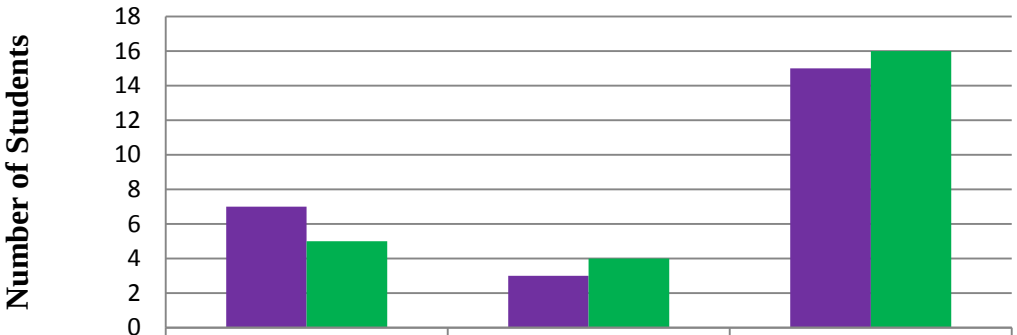
with their friends using both Malay and English language. None in Category A and Category B used English as their language choice when communicating with their friends.

The data in Figure 13 is similar to Figure 12: matriculation students are found to prefer to use their mother tongue in informal situations. The probability of using Malay over English might not only be limited to the ease of speaking the language fluently and confidently, but also due to their language exposure. The students are more familiar in using Malay to chat with their family members and friends because their mental productive lexicon is bound to yield vocabulary for sharing information with their closed ones about daily lives, problems or just for small talk.

In contrast, none of the students selected English as their choice of language to communicate with their family and friends. There might be many other reasons of their language choice, and it is highly probable that because of the limited vocabulary they have. Hence they are not able to effectively share information in informal situations due to their limited exposure to the language in social contexts. Their vocabulary only revolves around academic settings as most of the time they will learn new academic words instead of vocabulary for social purposes.

Apart from examining the language preference of students with family members and friends, Figure 14 shows the choice of language used by matriculation students with lecturers.

Figure 14: The Languages Used by Matriculation Students with Lecturers



	Malay	English	Both
Category A	7	3	15
Category B	5	4	16

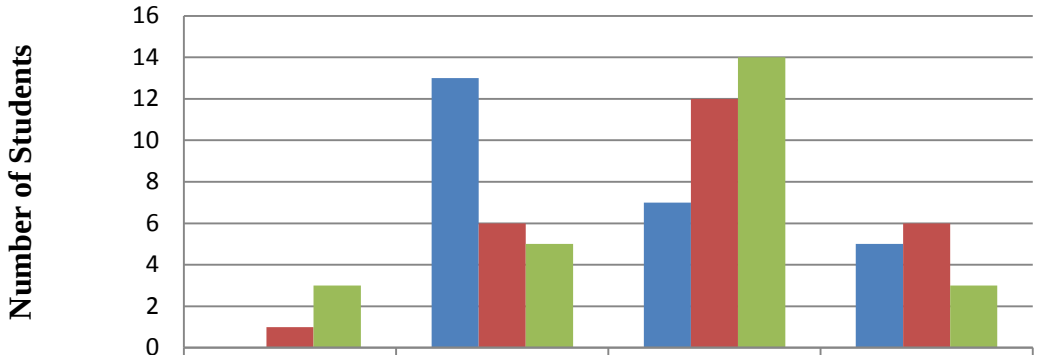
In a formal situation such as talking to lecturers, matriculation students tend to be bilingual, instead of using a single language. In Category A, there are 7 out of 25 students who still opt to use their mother tongue to speak with their lecturers while only 3 students used English effectively when asking their lecturers about academic-related questions. The majority of the students in Category A prefer to use both languages with their lecturers. In Category B, there are only 5 students who used Malay and 4 of them used primarily English with their lecturers. There are a total of 16 students in this category who used both Malay and English when communicating either face to face or indirectly by using emails with their lecturers.

Based on the data presented in the Figures 12 - 14, it may be inferred that Category A students are mostly Malay speaking students while Category B students are more bilingual speakers. Category B students usually use both Malay and English when they are communicating either with their family members, friends or lecturers.

Apart from the choice of spoken language used by matriculation students, the amount of exposure to English reading materials such as the newspapers, magazines and books, is

also taken into account. In the students' questionnaire, they were asked to select the amount of time spent on each reading material. The first question *"How often do you read English newspapers?"* required students to select one out of the four options given: *Everyday, 2-3 times a week, 2-3 times a month* and *Never*. Figure 15 displays the data on the frequency of Category A students reading English newspapers, magazines and books.

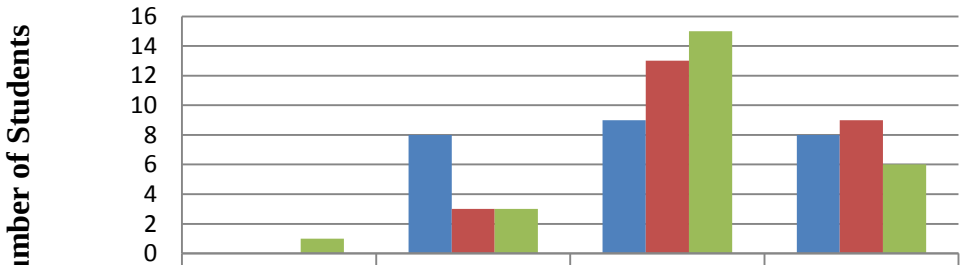
Figure 15: The Frequency of Category A Students Reading English Newspapers, Magazines and Books



	Everyday	2-3 times a week	2-3 times a month	Never
Newspapers	0	13	7	5
Magazines	1	6	12	6
Books	3	5	14	3

The data in Figure 15 shows that students in Category A never read English newspapers on a daily basis. However, 13 out of 25 will do it 2 - 3 times a week, while 7 of them will read English newspapers 2-3 times a month. There are 5 out of 25 students who never read the English newspapers. The students may be labelled as less frequent magazine readers. Although only 1 of them read English magazines for leisure, 6 out of 25 students will pick up a magazine to read in 2-3 times a week, while 12 out of 25 students read 2 - 3 times a month. There are 6 of them who never read any English magazine.

Figure 16: The Frequency of Category B Students Reading English Newspapers, Magazines and Books



	Everyday	2-3 times a week	2-3 times a month	Never
Newspapers	0	8	9	8
Magazines	0	3	13	9
Books	1	3	15	6

Figure 16 clearly illustrates that students in Category B seldom read English reading materials on a regular basis. The students only read 2 - 3 times a month. Only 1 student reads a book on a daily basis. Students in this category have limited exposure to English reading materials because the data in the figure above displays that students do not have a daily reading habit. There are only 8 students who read the newspaper 2 - 3 times a week, whilst the other 3 read magazines and books 2 - 3 times a week.

Analysing the data in the students' questionnaire for reading frequency in a month, 15 of the students still pick up a book to read, followed by 13 students who read magazines, followed by 9 students who read the newspapers. Nevertheless, the number of students who have never read a book, newspaper or magazine is relatively high. There are 8 students who never read the newspapers followed by 9 students who never read the magazines and 6 students who never read a book.

Although reading materials such as the newspapers and magazines are easily available within the vicinity of the university and even convenient stores outside the premise of the university, students still have limited exposure to these reading materials. In an

attempt to see the difference between Category A and B students on the frequency of reading English newspapers, magazines and books, it is interesting to note that students in Category A have more exposure to English reading materials despite the fact that they have lower scores in writing compared to Category B students.

Category B students have low receptive vocabulary due to limited exposure in leisure reading but they are capable of attaining higher scores in their writing compared to Category A students. This might be due to the nature of the question given in the MUET written paper. The question *"Preserving the forests is the key to saving our planet. What is your opinion on this statement?"* required students to write a composition which contains a high number of scientific words.

The participants in this research are currently pursuing their foundation studies in science and built environment and thus they are frequently exposed to scientific words in their area of discipline. Therefore, leisure reading among matriculation students does not play a major role in their writing and tend to have less impact on their scores given by the language instructors.

4.4 Teachers' Responses on Students Written Assessment

Besides looking at the language background of the students, the researcher distributed the questionnaire to five experienced teachers in order to elicit more information on their marking criteria in writing. It is felt to be necessary to understand the teachers' perceptions towards students' writing because their perceptions will indirectly affect the grades. The teacher-respondents are currently teaching MUET at the institution and they were given a 14-item questionnaire to elicit information on the weight they put on

lexical features. Two respondents viewed vocabulary of equal importance to other features such as grammar, while the other three respondents looked at vocabulary as the second most crucial item when allocating marks in a written composition. Even though they viewed vocabulary as an important aspect in writing, the feedback on the criteria for a good piece of written work they gave varied.

The first question *"Do you agree that advanced vocabulary (low frequency words) is important for students to score better in writing essays"* was intended to elicit information on the significance of advanced vocabulary in writing. However, none of the respondents stated the direct importance of advanced vocabulary in writing. Some of the responses were "the most important aspects of a good essay are to have a clearly defined thesis statement, minor grammatical error and in overall, organised" and "not important as it will only create confused writing pieces".

The respondents added that even though low frequency words are not placed as the most significant aspect to better scores, it will be an added advantage if students are able to use advanced and varied vocabulary accurately. Based on this information, it can be inferred that teachers might not be so particular with the use of low frequency words in essays to be perceived as good, but what is more crucial is that the students are able to make use of low frequency words fluently and accurately. The use of low frequency words inaccurately may actually distort an essay.

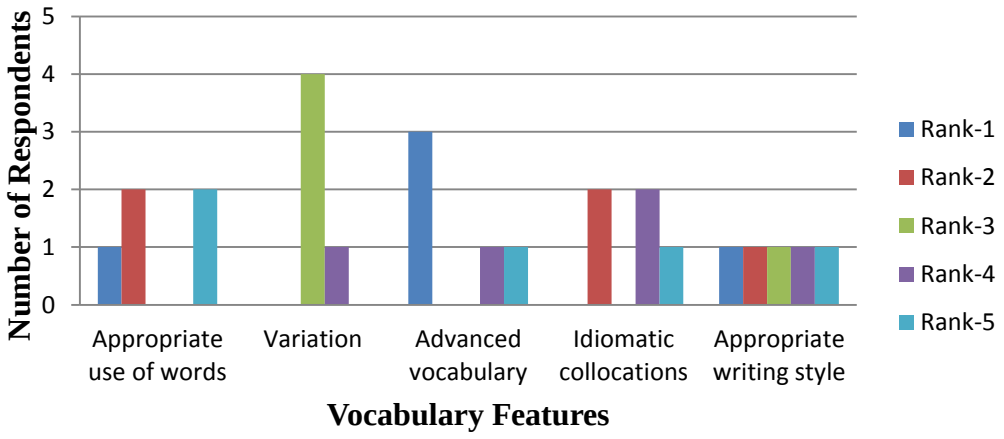
A good written composition will always be linked to the amount of vocabulary that the students were exposed to. The next question *"Do you teach vocabulary specifically in class? If yes, please state the allocation of time for students to learn vocabulary"* aimed to look at the students' frequency in learning vocabulary in class. The respondents

stated that they did not assign a specific period of time for students to learn vocabulary in class. Instead, the students were usually exposed to the low frequency words through reading comprehension practices in class. The teachers incorporated the learning of vocabulary in reading comprehension by asking students to identify the difficult words in the passage.

Next, they are required to look for the meaning and make sentences with the particular word to familiarise them with the usage of word in the correct context. When the respondents were further asked for the reason of why vocabulary is not taught as an individual lesson in class, they responded that the time allocated for MUET lesson was not sufficient. The teachers had to teach the four skills in MUET: listening, speaking, reading and writing, therefore it is not possible to teach vocabulary separately in a class.

Finally, the last question was *"The following are the different vocabulary features of a well written work. Please rank the vocabulary features according to its importance from 1 being the least important to 5 being the most important"*. The respondents were asked to rank the five given vocabulary features: appropriate use of words, variation, advanced vocabulary, idiomatic collocations and appropriate writing style. These five vocabulary features were adapted from Lemmouh's study (2008). Figure 17 illustrates the respondents' answers to this question.

Figure 17: Importance of Vocabulary Features of a Written Work



The vocabulary features that the teachers considered as important in a well written work are *advanced vocabulary* which was selected as the main element in an essay, while the *appropriate use of words* fell as the last choice.

Appropriate use of words is primarily concerned with the depth of lexical knowledge which includes the meaning of the word and the register constraints. Students who are able to use appropriate words must not only understand the word meaning but also the use of words in the right context. This vocabulary feature is not clearly defined in the LFP analysis as the programme only examines the presence of advanced vocabulary regardless of its appropriateness in an essay. Therefore, LFP programme which analysed a word as correct in an essay might be assessed by the teacher as incongruously used, and thus ensuing in a poor assessment of the quality of the essay. Nevertheless, the teacher-respondents categorised this feature as the least significant in comparison with the other four vocabulary elements.

The next vocabulary feature, *variation*, was ranked in the third place by the respondents. *Variation*, deals with the variety of words in a written work and the ability to use a

variety of words instead of using the same word repeatedly. Having variation contributes to a positive impression and indirectly influences the examiners' judgment on the students' written output. Based on a thorough analysis of the 100 samples, it was found that the matriculation students used certain words repeatedly. In view of the essay topic which was on giving an opinion on forest preservation, it basically deals with the environment. Thus, certain words such as environment, forest, planet, global warming, deforestation, carbon dioxide and oxygen were frequently repeated in their essays, and this is inevitable.

Meanwhile, *advanced vocabulary* was rated as the most important vocabulary feature in classifying a good essay. The presence of advanced vocabulary or low frequency words in an essay might positively affect the examiners' perceptions and directly influence the grades. This is reflected in the function of LFP programme to categorise words used in an essay to low and high frequency according to the four frequency lists: Band 1, Band 2, AWL and NIL. Advanced vocabulary comprises of words in AWL and NIL or words beyond the 2000 most frequent words in English (Beyond 2000 vocabulary) and they are considered as low frequency words. Credits might be given to mark up the students' scores if they are able to use a variety of advanced vocabulary in writing an essay.

Idiomatic collocations are deemed as unnecessary to assess an essay of good quality. This vocabulary feature was rated as both Rank 2 and Rank 4 by the teacher-respondent in this study. The last feature *Appropriate writing style* was voted only once by all the respondents. Nevertheless, students must adhere to the correct writing style according to the given question, taking into account that the respondents considered this feature as important, yet not necessary to contribute to a good essay.

In addition to the importance of vocabulary features in an essay, the degree of prominence between vocabulary and grammar are also to be taken into consideration to understand the teachers' opinion on this matter. The featured question "*Do you think a student can have a good essay with poor grammar?*"

Based on the responses to this question, four respondents answered *yes* while the remaining one answered *no.* Majority of the respondents agreed that students can have a good essay with poor grammar. They commented that poor grammar does not impede the quality of the essay if the content is clear, whereas the one respondent who disagreed with this statement stated that poor grammar in an essay will affect the quality and thus results in low scores.

In the next related question "*Do you agree that advanced vocabulary (low frequency words) is important for students to score better in writing essay"*, a majority of the respondents disagreed that the high degree of advanced vocabulary or low frequency words in essay resulted in better scores because the five teacher-respondents regarded grammar and advanced vocabulary the same. Some of them commented that plain English words with clear thesis statement are deemed important in an essay whereas the presence of advanced vocabulary only creates further confusion into a piece of writing if the student uses it inaccurately. Hence, the responses given by the respondents for both question construed that poor grammar and low degree of advanced vocabulary tend to have no effect on the scores of an essay.

However, the LFP analysis of the 100 compositions indicated different findings. Students who used advanced vocabulary have a tendency to score better than the students who used less advanced vocabulary. The examiners adhered to specific written

criteria when marking students' essays but the judgment to distinguish a good and bad essay is rather subjective. Examiners might not list advanced vocabulary in their marking scheme yet the presence of advanced vocabulary in general might affect the examiners' perceptions in grading the written compositions.

This study is carried out with pre-university students as subjects of the study and their writing output as a variable data to investigate the relation of lexical richness and holistic scores in the Malaysian context. To my knowledge, very limited studies have been conducted in this field, and researchers like Mokhtar (2010), Nadarajan (2011) and Teoh (2009) have used the Range programme to measure students' compositions which yielded different findings in their studies.

Based on the findings in this study, it has shown a positive relationship between the use of low frequency words in written text and holistic scores given by the teachers. It also validated the effectiveness of Range programme in measuring students' lexical richness based on the use of low and high frequency words in composition. It can be inferred that low frequency words affect the composition scores and thus it shows a need for students to use a variety of advanced words.

4.5 Evaluation of the Lexical Frequency Profile (LFP)

This section answers the third research question *"How effective is Lexical Frequency Profile (LFP) in measuring lexical richness in writing?"*. Laufer and Nation (1995) introduced the LFP which utilises a computer programme called the Range programme. This programme matches the four categories of words lists against a text that is typed in to see what words are present or absent in the lists. This innovation has sparked interest

among researchers to investigate the students' lexical richness regardless of writing or speaking skills. The Range programme has the ability to analyse students' texts and determine their proficiency based on the use of advanced vocabulary. A student is considered a highly proficient student if the Range programme analysis displays a high percentage of low frequency words in an essay.

In this study, the Range programme has proven to be a reliable tool in identifying good quality writing by analysing the words use in a text. Based on the analyses of T-test and scatter plot, it can be deduced that the null hypothesis (H0) in this study is rejected. Thus, it implies that there is positive correlation between the students' lexical richness and the holistic scores.

Despite the advantages, the programme has its disadvantage in categorising words in a text according to the four categories of words; Band 1, Band 2, Academic Word List (AWL) and Not in the List (NIL). Wiktorrson (2001) challenged one of Laufer and Nation (1995) claims of LFP or the Range programme as a reliable and valid measure of lexical use in writing.

Wiktorrson who classified multiword expressions (MWE) as 'prefabricated phrases' has conducted a study to compare the written output between NSs and NNSs and found out that both groups used a similar high amount of prefabricated phrases; not taking repetitions and assuming prefabricated phrases only consist of two words. Some examples of the prefabricated phrases are *hard facts*, *baby boom, look back, wake up, the same, the next,* and *most of*.

Smith (2005) states that LFP has no capacity to recognize and count MWEs which in turn weakens the earlier claim made by Laufer and Nation on LFP as a reliable and valid lexical richness measurement tool.

To further resist the claim, Smith points out that on the account of LFP measures phrasal verbs and prepositional verbs separately as individual word: it will be analysed by LFP as highly frequent prepositions such as in, on, at, by and off and highly frequent verbs such as *bring, carry, take* and *go.* Nevertheless, if the particular phrasal verbs are analysed together such as *take in, carry on* and *take off* will be counted as low frequency words which therefore makes counting or not counting MWEs could produce an inconsistent set of data and cause invalid results in a research.

Some of the words categorised in the NIL level are of high frequency words. The NIL words list includes words such as kids, television, fox, zoo, lions, bee, crab, and honey. The Range programme does not group these words in any of the three word frequency levels. These words are easily found in the early stage of education which denotes that these high frequency nouns should not be labelled as words in the NIL. Perhaps, Laufer and Nation could refine the list of Band 1, Band 2 and AWL lists in order to give words like these a suitable category instead of placing them on the NIL lists and referred to as low frequency words.

In this study, the analysis of each composition has shown that the claims made by Wiktorson (2001) are true, because the words which are categorised in NIL are not necessarily less frequent words as defined by Laufer and Nation (1995). The lexis in a text is compared against the lists of words that are already installed in the Range programme in ASCII format. The list of words is not updated and the vocabulary which

Laufer and Nation (1995) categorised as advanced words in AWL and NIL in year 1995 may not be relevant in the current era. The lists of words should be revised and certain words such as *television, cancer, hydropower* and others should be removed from the current NIL list. (Refer to the NIL word lists in the appendices for the full view of the words detected in students' essays). In the current era, the society is familiarised with these words as these words have frequently appeared in the mass media compared to the situation of more than a decade ago when the words are considered "alien" to them.

Although LFP or the Range programme has its disadvantages, it has proven to be a consistent tool in assessing lexical richness in written compositions. It can be improved by updating the four categories of words according to the frequency of words in our daily lives to effectively generate a list of words that is current and relevant.

CHAPTER 5: CONCLUSION

5.0 Introduction

This chapter summarises and concludes the findings in this study. Insights gained from this study will also be discussed and some suggestions for future research on LFP in the Malaysian context are offered.

5.1 Summary of Findings

The answers to the three research questions formulated in Chapter 1 have been answered in Chapter 4 and a summary of the findings is provided in this section. The first research question "*What are the students' lexical richness in terms of low and high frequency vocabulary as measured by the Range programme?"* revealed that low proficient students in Category A (20 - 29) and B (30 - 39), incorporated less advanced vocabulary in their writing while high proficient students in Category C (40 - 49) and D (50 - 59) utilised more low frequency words in their essays. The findings also identified the similarities between the less proficient and high proficient students in their choice of words when composing essays. The analysis which was carried out using the Range programme found that the four categories of students who used high frequency words can be found in Band 1 followed by Band 2, NIL and lastly Band 3. This is interesting to note as the data shows that even the proficient students used a bigger number of high frequency words in their essays.

The results that pertain to the second research question "*How does lexical richness in composition writing correlate with writing scores in MUET"*, show that the

matriculation students' use of low and high frequency of words, to some extent, affected their writing scores. Low frequency words are defined as words that stretched beyond the 2000 words, or in the B2000 category are words found in AWL and NIL of the Range programme. This categorisation in the compositions obtained according to the level of frequency in the programme, has confirmed that the choice of words used by students in their writing has made a small yet noticeable influence on their scores. As mentioned in Chapter 4, the relationship between students' lexical richness and holistic scores in their writing illustrated that the holistic scores increased with the use of low frequency words. From Figure 5 (section 4.2.1), Category A students used less advanced words in their writing. The total average of AWL and NIL increased steadily with the holistic scores. Therefore, it may be assumed that students who used more advanced words results in a positive impression, yielding better scores.

Apart from the evidence of low frequency words in the written output, the differences of the average word types utilised in the pool of essays were further analysed and discussed in Chapter 4. The word types are found to be in the ascending level, from Category A to Category D. Although the examiners used a set of marking criteria system to evaluate the students' writing performance, it seems that the task is not as easy or straightforward. To assess a piece of writing is subjective. Some instructors perceived a good piece of written output as one with low frequency words.

From the instructors' feedback, it is also noted that variation was selected as the third most important feature when assessing an essay. Although LFP did not measure variation in a text, it can still be analysed through the number of word types indicated in the LFP analysis. The findings reveal that Category A has the least variation where the compositions contained a relatively high number of repeated words, while the highly

proficient students in Category D used a wider variety of words. Those in Category D obtained better scores compared to those in other categories. Thus, holistic scores might be influenced by variation and the frequency of words in a composition.

For the third research question *"How effective is Lexical Frequency Profile (LFP) in measuring lexical richness in writing?"*, the researcher found that the findings of this study prove the Range programme to be a reliable tool to measure students' lexical richness in written compositions used in this study.

The LFP programme has been proven to be a reliable tool as several other studies which adopted it as a tool to measure students' lexical richness in writing also reveals this reliability. Abbasian and Parizad (2011) validate the computer programme as a measure of lexical richness by looking at the written discourse of 50 Iranian EFL students who were majoring in English Translation. The students were motivated to produce two written compositions on general topics to generate LFP analysis. With the aim to seek reliability of the programme, both compositions were correlated to validate the LFP. They found that the LFP analysis of two compositions of different topics correlated significantly. Due to the significant correlation coefficient, LFP analysis is conservatively safe and to certain extent is a very reliable and valid measurement instrument.

Although the Range programme is proven to be an effective tool in measuring lexical richness in writing, certain words are considered to be off-list in which the words are not found in Band 1, Band 2 and AWL (Refer to Section 4.6 for NIL words). These words might be included in Band 1 (1000 most frequent words) and Band 2 (next 1000 words most common words) as high frequency words. The Range programme was

developed in year 1995 and words are constantly changing over time, thus there is a rise of need to revise and update the words to reflect current usage.

The researcher also investigated the effects of the participants' language backgrounds on the quality of their writing. The finding from the questionnaires showed that these students from PASUM are learning English as their second language. They have acquired a formal education in school and most of the time they are exposed to the English language in school but they rarely speak it at home with their family members or relatives.

On the other hand, the five instructors who provided feedback on the learners writing stated that a good composition is evaluated based on several criteria. With regard to vocabulary features, they were asked to rank the five given features: appropriate use of words, variation, advanced vocabulary, idiomatic collocations and appropriate writing style. The findings demonstrated that advanced vocabulary is the most important feature in a good essay, and this was reflected in the students' holistic scores in this study.

5.2 Limitations of the Study

This study is relatively small as it only investigates lexical richness of 100 graded essays written by pre-university students at one Malaysian matriculation centre, and thus the findings in this study cannot be generalized. The samples of essays (n = 100) and the number of questionnaires (n = 50) obtained provided additional support to the analysis. Based on the limited sampling and voluntary participation of the respondents, which were relatively small in number, the findings can only be applicable to this study,

and the results cannot be generalised to represent the student or teacher population in Malaysia.

5.3 Pedagogical Implications

From a pedagogical point of view, the LFP can be used to measure lexical richness in written texts and it also serves as an important tool to classifying students writing based on their level of proficiency. LFP has proven to be a reliable diagnostic tool to identify students who are less proficient in writing or to identify students who have higher chances in failing due to poor writing, based on the LFP analysis on their written text.

Through early identification of students who are less proficient in writing, measures to improve these students writing proficiency can be done before they are asked to withdraw from a certain course. Students with difficulties in writing can be taught by exposing them to a larger repertoire and richer vocabulary knowledge. This could be done by motivating them to read more. Their productive vocabulary could increase if they are exposed to more receptive vocabulary. This suggestion concurs with what Laufer and Paribakht (1998) said that the exposure to low frequency words will activate their receptive vocabulary.

Language teachers may also implement useful vocabulary practices in the classroom and expose students to several techniques to learn vocabulary, such as using contextual clues, making connections of the word to mental images, using mnemonics, playing word games or to a certain extent, look up for the word meaning in a good dictionary. Further, students could be rewarded based on their lexical richness in their writing and

this may result in raising their awareness on the importance of advanced vocabulary in obtaining high scores in their writing.

5.4 Suggestions for Future Research

Based on the limitations of this study, it may be inferred that future researchers who are interested to pursue a study of this nature, may look into examining the lexical richness of each student using relevant productive and receptive vocabulary tests, exploring the students' proficiency in writing, and taking into account the individual differences of each student, such as language background, age and gender. The investigation into the role of different genres on the productivity of advanced vocabulary could also be carried out as it appears that when students are asked to write narrative or descriptive texts, their lexical richness will differ, due to the need to use different word choices to fulfil the requirement of these genres. Future researchers might also find it interesting and worthwhile to increase the number of written text samples to be analysed using the Range programme and administer a larger number of questionnaires on respondents. More researches could also be carried out on varying groups of learners, or at different levels, as in primary and secondary school learners in Malaysia. A comparative study could also be done on the different ethnic groups writing ability and their holistic scores using LFP. This may contribute to the field of language learning and culture influences.

5.5 Conclusion

In this study, three research questions have been designed to find out the students' lexical richness in terms of low and high frequency vocabulary as measured by the Range programme, the correlation between lexical richness in composition writing and

writing scores in MUET and the effectiveness of Lexical Frequency Profile (LFP) in measuring lexical richness in writing. Previous studies such as Lemmouh (2008) found no significant relationship between lexical richness and holistic scores while Teoh (2009) conducted a similar study and found that there was a low positive correlation between holistic vocabulary scores and the LFP.

Unlike previous studies which found no relationship or a weak correlationbetween the two independent variables, this dissertation reveals slightly different findings. The results in this study have proved that there is indeed a positive relationship between lexical richness and holistic scores. Based on the data from the teacher-respondents questionnaires, it was found that low frequency words or advanced vocabulary is not the main criterion in judging a good piece of text. However, there are some apparent contradictions when the LFP analysis of the 100 samples of compositions revealed that advanced vocabulary has its effects on scores. It is not easy to offer an objective assessment on the quality of writing due its subjective nature.

> "One of the effective ways for vocabulary assessment is through evaluation of the language learners' free writings, but free writing evaluation has always been a thorny task for language educators for its heavy reliance on the subjective judgment of the human raters." (Abbasian & Parizad, 2011, p.57).

Through the use of LFP which is a reliable tool, it is hoped that the analysis may be used by teachers to identify well written essays and thus they may be able to provide a more objective view when awarding the grades.

REFERENCES

Abbasian, G. R. & Parizad, M. S. (2011).Validation of lexical frequency profiles as a measure of lexical richness in written discourse. *Journal of Technology & Education, 5(3).* Retrievedfrom http://jte.srttu.edu/browse.php?a_id=287&slc_lang=fa&sid=1&ftxt=1

Llach, A.M. P. & Gallego, T.M. (2009). Examining the relationship between receptive vocabulary size and written skills of primary school learners. *Journal of the Spanish Association of Anglo-American Studies, 31, 129- 147.*Retrieved from www.atlantisjournal.org/ARCHIVE/31.1/2009Agustin_Terrazas.pdf 129-47.

Alamargot, D. & Chanquoy, L. (2001).Through the models of writing. Retrieved from http://books.google.co.uk/books?id=jtUshuUzsHIC&printsec=frontcover&dq=Denis+Alamargot+and+Lucile+Chanquoy&source=bl&ots=kVYUCy5pn9&sig=0vzcMZ-T-0dV9KbXUyOo0T9gMdI&hl=en&ei=fmyTZ_2FIf4rQetvbHIDQ&sa=X&oi=book_result&ct=result&resnum=1&ved=0CBcQ6AEwAA#v=onepage&q&f=false

Alderson, J.C. (2000). *Assessing reading*. Cambridge: Cambridge University Press.

Andrews, S. J. (2009). *Educational background as predictor of lexical richness among Libyan and Saudi Arabian ESL students*. Master's Thesis, University of Pittsburgh.

Astika, G.G. (1993). Analytical assessments of foreign students' writing.*RELC Journal 24: 61-71.*

Bacha, N. N. (2002). Developing learners' academic writing skills in higher education: A study for educational reform. *Language and Education,16(3), 161-177.* Retrieved from www.tandfonline.com/doi/abs/10.1080/09500780208666826

Bauer, L. & Nation, P. (1993).Word families.*International Journal of Lexicography, 6(4).*Retrieved from www.victoria.ac.nz/lals/...nation/1993-Bauer-Word-families.pdf

Beckett , G. H. & Gonzalez, V. (2004). Content-based ESL writing curriculum: A language socialization model. *Journal of Research and Practice,2(1), 161-175.* Retrieved from njrp.tamu.edu/2004/PDFs/Beckett%20.pdf

Berman, R. & Cheng, L. (2010). English academic language skills: Perceived difficulties by undergraduate and graduate students, and their academic achievement. Retrieved from blade2.vre.upei.ca/ojs/index .php/cjal/article/download/169/139

Bialystok, E. (2001). Bilingualism in Development. Retrieved from http://assets.cambridge.org/97805216/32317/frontmatter/9780521632317_frontmatter.pdf

Bogaards, P. & Laufer, B. (2004).*Vocabulary in a second language: Selection, acquisition and testing.* Amsterdam: Benjamins

Bonanno, H. & Jones, J. (2007). The MASUS procedure: Measuring the academic skills of university students: A diagnostic assessment. Retrieved from sydney.edu.au/stuserv/documents/learning_centre/MASUS.pdf

Burrough-Boenisch, J. (2003). Shapers of published NNS research articles.*Journal of Second Language Writing, 12, 223–243.*

Calzolari, N., Fillmore, C., Grishman, R., Ide, N., Lenci, A., MacLeod, C.,& Zampolli, A. (2002). Towards best practice for multiword expressions in computational lexicons. In proceedings of the 3rd international conference on language resources and evaluation, Las Palmas, Canary Islands, Spain. 29-31 May, 2002 (pp. 1934-1940). Paris: European Language Resources Association.

Chapman, K. (2007). Varsity students do badly in MUET. The Star. Retrieved from http://thestar.com.my/news/story.asp?file=/2007/1/23/nation/16647460&sec=nation

Chase, B.J. (2011). An analysis of the argumentative writing skills of academically underprepared college students. Retrieved from http://academiccommons.columbia.edu/download/fedora_content/download/ac:131454/CONTENT/Chase_columbia_0054D_10083.pdf

Chen, X. (2007).Lexical cohesion in Chinese college EFL writing.*CELEA Journal, 30(5).* Retrieved from www.celea.org.cn/teic/75/75-46.pdf

Choy, S. C. & Troudi, S. (2006). An investigation into the changes in perceptions of and attitudes towards learning English in a Malaysian college. *International Journal of Teaching and Learning in Higher Education, 18(2), 120-130.* Retrieved from http://www.isetl.org/ijtlhe/pdf/IJTLHE99.pdf

Crossley, S. A. C. & McNamara, D. S. (2009). Computational assessment of lexical differences in L1 and L2 writing. *Journal of Second Language Writing,18, 119–135*. Retrieved from www.sciencedirect.com/science/article/pii/

S1060374309000046

Crossley, S. A., Salsbury, T., McNamara, D. S. & Jarvis, S. (2010). Predicting lexical proficiency in language learner texts using computational indices. *Language Testing.*Retrieved from http://129.219.222.66/pdf/Predicting_ Lexical_Proficiency_In _Language_Learners.pdf

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

Creswell, John W. (2003). *Research design: Qualitative, quantitative, and mixed methods approaches.*Thousand Oaks, CA: Sage Publications.

Darus, S. & Khor , H. C. (2009). Common errors in written English essays of form one Chinese students: A case study. *European Journal of Social Sciences*. Retrieved from www.eurojournals.com/ejss_10_2_07.pdf

David, M. K. (2008). Language policies – Impact on language maintenance and teaching focus on Malaysia, Singapore and the Philippines. Retrieved from http://www.scribd.com/doc/17081287/ Language-Policies- Impact-on-Language

Douglas, S. R. (2010). Non-native English speaking students at university: Lexical richness and academic success. Retrieved from dspace.ucalgary.ca/bitstream/1880/48195/1/2010_Douglas.pdf

Elman, L. J. (2004). An alternative view of the mental lexicon.*Trends in Cognitive Sciences, 8(7), 301-306.* Retrieved from http://crl.ucsd.edu/~elman/ Papers/elman_tics_opinion_2004.pdf

Durukan, E. (2011). Effects of cooperative integrated reading and composition (CIRC) technique on reading-writing skills. *Educational Research and Reviews,6(1), 102-109.* Retrieved from www.eric.ed.gov/ERICWebPortal/ detail?accno=EJ923632

Engber, C. (1995). The relationship of lexical proficiency to the quality of ESL compositions. *Journal of L2 Writing 4(2), 138-155.*

Erling, E. J. ,& Richardson, J. T. (2010). Measuring the academic skills of university students: Evaluation of a diagnostic procedure. *Assessing Writing 15,177–193.* Retrieved from www.sciencedirect.com/science/article/pii/S1075293510000383

Farahzad, F. & Emam, A. (2010). Reading-writing connections in EAP courses: Implications and applications.*Journal of Language Teaching and Research, 1(5), 596-604.* Retrieved from ojs.academypublisher.com/index.php/jltr/ article/view/0105596604

Feng, J. (2008). Enhancing college students active use of vocabulary- An empirical study on vocabulary in acquisition. *CELEA Journal, 31(1).* Retrieved from www.celea.org.cn/teic/77/08070211.pdf

Fitzgerald, J. & Shanahan, T. (2000).Reading and writing relations and their development. *Educational Psychologist, 35(1), 39-50.*

Fuentes, A. C. (2009). Case study corpus for academic English writing by NNS authors. Retrieved from www.um.es/lacell/aelinco/contenido/pdf/75.pdf

Giridharan, B. & Conlan, C. (2003). L2 Vocabulary acquisition: Investigating the key to lexical comprehension. Retrieved from citeseerx.ist.psu.edu/viewdoc/ download?doi=10.1.1.118...pdf

Goodfellow, R., Lamy, M. N. & Jones,G. (2002). Assessing learners' writing using lexical frequency. *ReCALL, 14(1),133–145.* Retrieved from dl.acm.org/citation.cfm?id=984943

Graham, S.&Perin, D. (2007).*Writing next: Effective strategies to improve writing of adolescents in middle and high schools.* Washington, D.C: Alliance for Excellent Education.

Grimes, B. F. (Ed.) (2000). Ethnologue*: Languages of the world,* (14th ed.). Dallas, TX: SIL International. ISBN 1-55671-106-9. Retrieved fromhttp://www.ethnologue.com/

Hamzah, M. S. G. & Abdullah, S. K. (2009). Analysis on metacognitive strategies in reading and writing among Malaysian ESL learners in four education institutions. *European Journal of Social Sciences, 11(4).* Retrieved from www.eurojournals.com/ejss_11_4_13.pdf

Harfitt, G. J. (1999).A comparison of lexical richness in samples of written and spoken English from a group of secondary six students in Hong Kong. Retrieved from http://hub.hku.hk/bitstream/10722/27623/1/ FullText.pdf?accept=1

Hazenberg, S. & Hulstijn, J.H. (1996). Defining a minimal receptive second-language vocabulary for non-native university students: An empirical investigation. *Applied Linguistics, 17, 145-163.* Retrieved from http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.129.7732&rep=rep1&type=pdf

Hinkel, E. (2004). *Teaching academic ESL writing*. New York: Lawrence Erlbaum Associations.

Hrabincova, E. (2002). *Individual strategies in learning English vocabulary.* University of Pradubice.

Hyland, K. & Tse, P. (2007). Is there an 'academic vocabulary'? TESOL Quarterly, 41(2).

Ikah (2006). Unpublished Journal: The correlation between students achievement in vocabulary and reading ability. Jakarta: State Islamic University Syarif Hidayatullah.

Jafarpur, A. (1996). Native speaker performance validity: In vain or for gain?. *System, 24, 83-95.*

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

Johansson, V. (2008). Lexical diversity and lexical density in speech and writing: A Developmental Perspective. Lund University, Dept. of Linguistics and Phonetics Working Papers,53, 61-79. Retrieved from nile.lub.lu. se/ojs/index.php/LWPL/article/view/2273/1848

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

Kitao, K. (April, 2006). Using the internet to teach English. 5(8). Retrieved from http://iteslj.org/Articles/Kitao-WhyTeach.html

Komolafe, A. T. & Yara, P. O. (2010). Sentence combining strategy and primary school pupils achievement in written English in Ibadan, Nigeria. *European Journal of Scientific Research,40(4), 531-539.* Retrieved from http://www.eurojournals.com/ejsr.htm

Kuang, C.Y. (2011). The impact of EFL students' vocabulary breadth of knowledge on literal reading comprehension. *Asian EFL Journal, 51.* Retrieved from www.asian-efl-journal.com/PTA/April-2011-kyc.pdf

Laufer, B. (1997). *The lexical plight in second language reading: Words you don't know, words you think you know, and words you can't guess*. In J. Coady& T. Huckins (Eds.), Second language vocabulary acquisition: A rationale for pedagogy (pp. 20-34). New York: Cambridge University Press

Laufer, B. (1998). The development of passive and active vocabulary in a second language: Same or different? *Applied Linguistics, 19(2), 255-27.*

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

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

Laufer, B. & Paribakht, T. S. (1998). Relationship between passive and active vocabularies: Effects of language learning context. *Language Learning 48: 365-391*

Leki, I. (2009). Before the conversation: A sketch of some possible backgrounds, experiences, and attitudes among ESL students visiting a writing center. In S. Bruce & B. Rafoth (Eds.), ESL writers: A guide for writing center tutors (2nd

ed., pp. 1–17). Portsmouth, NH: Boynton/Cook Publishers.

Lemmouh, Z. (2008). The relationship between grades and the lexical richness of student essays. *Nordic Journal of English Studies, 7(3), 163 - 180.* Retrieved from http://130.241.35.204/ojs/index.php/njes/article/viewFile/157/155

Lewis, M. (1993).The Lexical Approach.*Journal of Cambridge Studies, 128.*

Li, C. (1997). Is lexical richness an essential criterion in judging a piece of writing? Retrieved from hub.hku.hk/handle/10722/26757

Lillis, T. & Scott, M. (2007). Defining academic literacies research: Issues of epistemology, ideology and strategy. *Journal of Applied Linguistics, 4, 5–32.* doi:10.1558/japl.v4i1.5

Llach, M. P. A. (2010). An overview of variables affecting lexical transfer in writing: A review study. *International Journal of Linguistics,2(1).*Retrieved from www.macrothink.org/journal/index.php/ijl/article/view/445

Llach, M. P. A. & Gallego, M. T. (2009). Examining the relationship between receptive vocabulary size and written skills of primary school learners. *Journal of the Spanish Association of Anglo-American Studies,31(1), 129-147.* Retrieved from www.atlantisjournal.org/ARCHIVE/ 31.1/2009Agustin_Terrazas.pdf

Maros, M., Tan, K. H. & Salehuddin. K. (2007). Interference in learning English: Grammatical errors in English essay writing among rural Malay secondary school students in Malaysia. *e-Bangi Journal 2 (2):1-15*

McNeill, A. (1994). *Vocabulary testing: defining word knowledge*. In Boyle, J & Falvev, P. English Language Testing in H.K., Chinese University Press.

Meara, P. (1996). *The dimensions of lexical competence.* Gilliam Brown, Kirsten Malmkjaer and John Williams (Eds.) Performance and Competence in Second Language Acquisition. Cambridge: Cambridge UP. 35-53.

Meara, P. (1996). The vocabulary knowledge framework. Retrieved from http://www.swan.ac.uk/clas/vlibrary/pm96d

Meara, P. & Alcoy, O. (2008). Words as species: an alternative approach to estimating productive vocabulary size. *22*(1). Retrieved from nflrc.hawaii.edu/rfl/ April2010/articles/meara.pdf

Meara, P. & Bell, H. (2001). P_Lex: A simple and effective way of describing the lexical characteristics of short L2 texts. *Prospect, 16 (3), 5–19.*

Meara, P. M. & Miralpeix, I. (2007). *D_Tools ver2.0 (computer program).* Swansea: Lognostics.

Mehring, J. G. (2005). Developing vocabulary in second language acquisition: From theories to the classroom. Retrieved from http://www.hpu.edu/CHSS/ LangLing/TESOL/ProfessionalDevelopment/200680TWPfall06/03Mehring.pdf

Mehrpour, S., Razmjoo, S. A. & Kian, P. (2011). The relationship between depth and breadth of vocabulary knowledge and reading comprehension among Iranian EFL learners.*Journal of English Language Teaching and Learning, 222.* Retrieved from www.ensani.ir/storage/Files/20120427094740-8079-5.pdf

Mokhtar, A. A. (2010). Achieving native-like English lexical knowledge: The non-native story. *Journal of Language Teaching and Research,1(4), 343-352.* Retrieved from http://ojs.academypublisher.com/index.php/jltr/article/download/0104343352/1962

Muncie, J. (2002). Process writing and vocabulary development: Comparing lexical frequency profiles across drafts.*System 30: 225-235.*

Nadarajan, S. (2008). Assessing in-depth vocabulary ability of adult ESL learners. *The International Journal of Language Society and Culture, (26).*Retrieved from http://www.educ.utas.edu.au/users/tle/JOURNAL/issues/2008/26-9.pdf

Nadarajan, S. (2011). The challenges of getting L2 learners to use academic words in their writings. *Electronic Journal of Foreign Language Teaching, 8(2), 184–200.* Retrieved from http://e-flt.nus.edu.sg/v8n22011/nadarajan.pdf.

Nakamaru, S. (2010). Lexical issues in writing center tutorials with international and US-educated multilingual writers.*Journal of Second Language Writing, 19(2), 95-113.* Retrieved from www.sciencedirect.com/science/article/pii/S1060374310000123

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

Nation, I.S.P.& Newton, J. (1997).*Teaching vocabulary*. In J. Coady& T. Huckin (eds.) Second Language Vocabulary Acquisition. Cambridge Applied Linguistics.

Nation, P. (2001). Learning vocabulary in another language, Cambridge: Cambridge University Press.

Nor Azni Abdullah (1993). Beyond the obvious: developing critical thinking skills in academic writing. In the *2nd National ESP Seminar Proceedings* titled ESP on: Applications and Implications for Human Resource Development. UTM, Johor.

Olive, T. (2003). Working memory in writing: Empirical evidence from the dual-task technique. *European Psychologist.*Retrieved from http://cogprints.org/3246/1/Olive_Euro_Psy.pdf

Ooi, V. B. Y. (2001). Upholding standards or passively observing language? Corpus evidence and the concentric circles model. Retrieved from courses.nus.edu.sg/course/.../upholdingstandards-vincentooi.pdf

Pandian, A. (2000, July). A study on readership behaviour among multi-ethnic, multi-lingual Malaysian students. Paper presented at the Seventh International Literacy and Education Research Network (LERN) Conference on Learning, Melbourne.

Pang, E. S., Muaka, A., Bernhardt, E. B. & Kamil, M. L. (2003). *Family Science Review*. Retrieved from www.ibe.unesco.org/fileadmin /user_upload/archive/.../prac12e.pdf

Paton, M. (2007).Why international students are at greater risk of failure: An inconvenient truth. *International Journal of Diversity in Organisations, Communities and Nations, 6 (6), 101–111.*

Pokorny, B. (2009). Language Frequency Profiling of Written Texts by Students of German as a Foreign Language. Retrieved from http://uwspace.uwaterloo.ca/ bitstream/10012/4680/1/Pokorny_Bjanka.pdf

Pustejovsky, J. (1996). *The Generative Lexicon*.MIT Press, Cambridge, MA.

Qian, D.D. (1999). Assessing the roles of depth and breadth of vocabulary knowledge in reading comprehension.*Canadian Modern Language Review, 56, 282-308.*

Read, J. (2000). *Assessing vocabulary.* Cambridge: Cambridge University Press.

Rethinasamy, S. & Kee M. C. (2011). The Malaysian University English Test (MUET) and its use for placement purposes: A predictive validity study. *Electronic Journal of Foreign Language Teaching, 8(2), 234–245.* Retrieved from http://eflt.nus.edu.sg/ v8n22011/rethinasamy.pdf.

Richards, J. C. & Rogers, T. S. (2001). *Approaches and methods in language teaching (2nd ed.)*. Cambridge, UK: Cambridge University Press.

Ruetten, M. (1994). Evaluating ESL students' performance on proficiency exams. *Journal of Second Language Writing*, 3, 85–96.

Sánchez, A. & Rosa M. (2007). Research on second language vocabulary acquisition and learning: An introduction. *International Journal of English Studies, 7(2), vii-xvi.* Retrieved from http://digitum.um.es/xmlui/ bitstream/10201/2582/1/2595023.pdf

Schmitt, N. (2000). *Vocabulary in language teaching.* Retrieved from catdir.loc.gov/catdir/samples/cam032/99057110.pdf

Segler, T. M. (2001). Second language vocabulary acquisition and learning strategies in ICALL environments. Retrieved from www.tandfonline.com/ doi/abs/10.1076/call.15.4.409.8272

Severino, C. (2009). Avoiding appropriation. In S. Bruce & B. Rafoth (Eds.), ESL writers: A guide for writing center tutors (2nd ed., pp. 51–65). Portsmouth, NH: Boynton/Cook Publishers.

Shafie, L. A. & Nayan, S. (2010). Employability awareness among Malaysian undergraduates. *Language Testing, 5(8).* Retrieved from http://ccsenet.org/ journal/index.php/ijbm/article/download/6917/5487

Shamsudin, S. & Kaur, M. (2011). Extracting noun forms: A lesson learnt. *International Journal of Language Studies (IJLS), 5(4), 19-32.* Retrieved From www.ijls.net/volumes/volume5issue4/kaur1.pdf

Schmitt, M. & McCarthy, M. (2003). *Vocabulary description, acquisition and Pedagogy.*Cambridge University Press, Cambridge.

Smith, B. (2005). The relationship between negotiated interaction, learner uptake, and lexical acquisition in task-based computer-mediated communication. *TESOL Quarterly, 39(1), 33-58.*

Snow, C. E. & Biancarosa, G. (2003). *Adolescent literacy and the achievement gap: What do we know and where do we go from here?* New York: Carnegie Corporation of New York.

Sripada, P. N. (2008). Mental lexicon. *Journal of the Indian Academy of Applied Psychology, 34(1), 181-186.*

Van Gijsel, S., Speelman, D. & Geeraerts, D. (2005). A variationist, corpus linguistic analysis of lexical richness. Retrieved from http://www.bhamlive2.

bham.ac.uk

Taiwo, R. (2004). Helping ESL learners to minimize collocational errors. *The Internet TESL Journal, 10 (4).* Retrieved from http://iteslj.org/Techniques /Taiwo-Collocation.html

Tamariz, M. (2004). Exploring the adaptive structure of the mental lexicon. Retrieved from www.ling.ed.ac.uk/~monica/tamariz_thesis.pdf

Teoh, L.N. (2009). *Lexis in compositions of ESL students in a selected university college in Nilai, Malaysia*. PhD thesis, Universiti Putra Malaysia.

Tercanlioglu, L. (2004). Pre-service EFL teachers beliefs about foreign language learning and how they relate to gender. *Electronic Journal of Research in Educational Psychology 5-3.1: 145-62.*

Todd, R. W., Khongput, S. & Darasawang, P. (2007). Coherence, cohesion and comments on students' academic essays. *Assessing Writing, 10-25.* Retrieved from www.eric.ed.gov/ERICWebPortal/record Detail?accno= EJ796721

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

Mohamad, M. (1991). Malaysia. The way forward. Kuala Lumpur: Centre for Economic Research & Service, Malaysian Business Council. Paper presented at the inaugural meeting of the Malaysian Business Council, Kuala Lumpur, Malaysia.

Wan, I. I., Shafinah, M. D. & Azhari, M. H. (2006). Utilizing needs analysis study to determine the current English language needs in the industrial sector. *Wahana Akademik Journal, 5, 75-80.*

Wang, L. (2005). Lexical richness in L2 academic writing. *CELEA Journal, 28*(5). Retrieved from www.elt-china.org/teic/63/63-18.pdf

Waring, R. (1999). Tasks for assessing second language receptive and productive vocabulary. Retrieved from http://www1.harenet.ne.jp/~waring/papers/phd/abstract.html.

Waring, R. & Nation, I. S. P. (1997).Vocabulary size, text coverage, and word lists. In *Vocabulary: Description, Acquisition and Pedagogy*.

Weigle, S. C. (2004). Integrating reading and writing in a competency test for non-native speakers of English. *Assessing Writing 9, 27-55.* Retrieved from www.clal.org.cn/personal/testing/MA/writing/weigle2004.pdf

White, C. P., Reichelt, S. & Woods, B. (2011). Low stakes writing as an instructional strategy to engage students. *Family Science Review,16(1).* Retrieved fromwww.familyscienceassociation.org/.../FSR_16.../White%20final.pdf

Wiktorsson, M. (2001). Register differences between prefabs in native and EFL English. In S. Manninen& C. Paradis (Eds.). The Department of English in Lund: Working Papers in Linguistics, 1, 85-94. Lund: Lund University. Retrieved from http://www.englund.lu.se/research/workingpapers/pdf/Maria.pdf.

Wu, Y. (2009). The application of CLT in college English vocabulary teaching. *Journal of Cambridge Studies*, 4(3),128-131.Retrieved from

http:// journal.acs-cam.org.uk/data/archive/2009/200903-article13.pdf

Xue, G. & Nation, P. (1984). A university word list. *Language learning and communication, 3, 215-229.*

Ying , Z. (2009). An empirical study of L2 learners' use of lexical chunks and language production. Retrieved from hkr.diva-portal.org/smash/get/diva2:229050/ FULLTEXT01

Yong, S. E., Tan, B. H., & Yong, M. F. (2012). Speech disfluencies and mispronunciations in English oral communication among Malaysian undergraduates. *International Journal of Applied Linguistics & English Literature,1(7), 19-32.*ISSN 2200-3592.

Zareva, A. (2005). Models of lexical knowledge assessment of second language learners of English at higher levels of language proficiency. *System 33, 547-562.*

Zeng, D. (2005). The process-oriented approach to ESL/EFL writing instruction and research. *CELEA Journal.28(5).*

Zhou, S. (2010). Comparing receptive and productive academic vocabulary knowledge of Chinese EFL learners. *Asian Social Science, 6(10).* Retrieved from ccsenet.org/journal/index.php/ass/article/view/6313

Zimmerman, K. J. (2004). *Role of vocabulary size in assessing second language proficiency.*Unpublished M.A. thesis. Brigham YoungUniversity (BYU), America

APPENDICES

APPENDIX A: SPECIFICATIONS OF WRITING COMPONENT IN MUET

COMPONENT	TEST SPECIFICATIONS
Writing	Candidates are assessed on their ability to write various types of text covering a range of rhetorical styles. Assessment will cover the following: **(i) accuracy** • using correct spelling and mechanics • using correct grammar • using correct sentence structures **(ii) appropriacy** • using varied vocabulary and expressions • using clear varied sentences • using language appropriate for the intended purpose and audience • observing conventions appropriate to a specific situation or text type **(iii) coherence and cohesion** • developing and organising ideas • using appropriate markers and linking devices • using anaphora appropriately together with other cohesive devices **(iv) use of language functions** • defining, describing, explaining • comparing and contrasting • classifying • giving reasons • giving opinions • expressing relationships • making suggestions and recommendations • expressing agreement and disagreement • persuading • interpreting information from non-linear texts • drawing conclusions • stating and justifying points of view • presenting an argument **(v) task fulfilment** • presenting relevant ideas • providing adequate content • showing a mature treatment of topic **Possible genres:** Report, article, letter, essay

(Taken from the Malaysians Examinations Council, 2006)

APPENDIX B: MUET BAND DESCRIPTION

AGGREGATED SCORE	BAND	USER	COMMUNICATIVE ABILITY	COMPREHENSION	TASK PERFORMANCE
260 – 300	6	Highly proficient user	Very fluent; highly appropriate use of language; hardly any grammatical error	Very good understanding of language and context	Very high ability to function in the language
220 – 259	5	Proficient user	Fluent; appropriate use of language; few grammatical errors	Good understanding of language and context	High ability to function in the language
180 – 219	4	Satisfactory user	Generally fluent; generally appropriate use of language; some grammatical errors	Satisfactory understanding of language and context	Satisfactory ability to function in the language
140 – 179	3	Modest user	Fairly fluent; fairly appropriate use of language; many grammatical errors	Fair understanding of language and context	Fair ability to function in the language
100 – 139	2	Limited user	Not fluent; inappropriate use of language; very frequent grammatical errors	Limited understanding of language and context	Limited ability to function in the language
Below 100	1	Very limited user	Hardly able to use the language	Very limited understanding of language and context	Very limited ability to function in the language

APPENDIX C: ACADEMIC WORDS LIST (AWL) FOUND IN DATA

NOUN	VERB	ADJECTIVE	ADVERB
EROSION	MAINTAIN	DISTINCT	HENCE
OBJECTIVE	CONCLUDE	GLOBAL	FURTHERMORE
PROCESS	AFFECT	ENVIRONMENTAL	CONSEQUENTLY
PROFESSIONALS	CONTRIBUTE	AVAILABLE	NORMALLY
AREA	CONSUME	CRUCIAL	DEFINITELY
BENEFIT	EXPLOITING	ENORMOUS	ENORMOUSLY
GENERATION	OCCUR	MEDICAL	ABNORMALLY
PERCENTAGE	RELEASE	RESTRICTED	AUTOMATICALLY
ROLE	IMPOSE	ABNORMAL	EVENTUALLY
CONCLUSION	STABILIZE	BENEFICIAL	FINALLY
LAYER	UNDERGO	ILLEGAL	DESPITE
COOPERATION	RESEARCHES	INSUFFICIENT	EVENTUALLY
UNIQUENESS	GENERATE	MAJOR	CONSTANTLY
DIVERSITY	OCCUPIED	SIGNIFICANT	SIGNIFICANTLY
LICENSES	MINIMISE	ECONOMIC	OBVIOUSLY
COMMITMENT	FUNCTIONING	AFFECTED	ULTIMATELY
COMMUNITY	MAINTAIN	NEGATIVE	APPROXIMATELY
ENERGY	AFFECTING	TRADITIONAL	VOLUNTARILY
RESOURCES	CONCLUDE	NORMAL	SUBSEQUENTLY
SOURCE	OCCUR	RESIDENTIAL	IMPLICITLY
CONSUMPTION	PARTICIPATE	ENORMOUS	VIRTUALLY
CYCLE	RELEASE	DYNAMIC	FURTHERMORE
SOURCE	SURVIVE	CHALLENGING	
AUTHORITIES	REMOVE	STABLE	
AWARENESS	ERODES		
ENFORCEMENT	ACCUMULATE		
TASK	APPRECIATE		

DECADES	CATEGORISED	STABLE	
ELEMENT	ENSURE	COMPLEX	
RATIO	SURVIVE	LEGAL	
SUM	AWARE	PASSIVE	
CHEMICAL	COLLAPSE	MEDIUM	
RESIDENTS	DISPOSE	MATURED	
SOURCE	ENFORCE	INITIAL	
STRUCTURE	REVERSED	PRIMARY	
CONTRIBUTOR	EXTRACT	SOLE	
INDIVIDUAL	FOCUSED	INEVITABLE	
OCCURRENCE	INSTANCE	OBVIOUS	
PHENOMENON	CREATE	UNIQUE	
CREATION	ADAPT	UNAWARE	
PERIOD	LOCATED	DRAMATIC	
AUTHORITY	REMOVING	PREVIOUS	
FACTORS	INVOLVE	ULTIMATE	
CONSUMER	REJECTED	MINIMUM	
ISSUE	PROHIBIT	DOMESTIC	
REGION	REGULATING	IGNORANT	
TRANSPORTS	ANALYZE	POSITIVE	
INTERACTION	AID	CONSTANT	
MINORITY	CLARIFY	PERCENT	
REQUIREMENT	CONSTRUCT	PHYSICAL	
IMPLICATIONS	CONVERT	MAXIMUM	
INCOME	DENIED	PROPORTIONAL	
STABILITY	REQUIRED	SUSTAINABLE	
SURVIVAL	ACCUMULATE		
GOALS	ENHANCED		
INITIATIVE	DECLINING		
MEDIA	MIGRATE		
METHOD	ACHIEVE		
QUOTE	STABILISE		
TREND	ABANDONED		
VISION	EXPLOIT		
ALTERNATIVE	PROMOTE		
OUTPUT	LINKED		

SUMMATION	COMMIT		
ACCUMULATION	JUSTIFIED		
ENVIRONMENTALISTS	IMPLEMENTED		
SECTOR	REINFORCE		
STRATEGY	ALTERED		
CONTRIBUTION	ADJUST		
JOB	ISSUED		
TEAM	IGNORE		
CONSEQUENCES	RESTORE		
RESOLUTION	SYMBOLIZED		
TECHNOLOGY	ESTABLISHING		
ACCOMMODATION	FEATURES		
INJURIES	GRANTED		
ASPECT	RETAIN		
CONCENTRATION	ACKNOWLEDGE		
CONSTRUCTIONS	ERODED		
ECONOMY	IMPLY		
GLOBALIZATION	EMPHASIZE		
INCIDENT	ASSURE		
STRESS	UTILIZED		
INTENSITY	MONITOR		
REACTION	RELAX		
SCOPE	APPRECIATE		
TASK	ESTABLISHED		
ADJUSTMENT	EXPOSED		
ITEMS	EXCEED		
VEHICLES	EXPOSED		
APPRECIATION	STRESSED		
ASSISTANCE	CYCLES		
AWARENESS	DISTORTED		
COMPONENT	INTERACT		
IDEOLOGY	CONSIST		
IMPACTS	SUSTAIN		
PORTION	COMMUNICATE		
SCHEDULES	DIMINISHING		
RESTORATION	COOPERATE		

CONVERSION THEORY IGNORANCE MECHANISMS TENSION	RECOVER ENFORCING IMPLEMENTING ELIMINATE REGULATE		
TOTAL: 94	**93**	**45**	**19**

APPENDIX D: NOT IN THE LIST (NIL) FOUND IN DATA

NOUN	VERB	ADJECTIVES	ADVERB
CAMPAIGN	ABSORB	ADVERSE	EFFICIENTLY
CARBON	ERASED	ECO- FRIENDLY	FOREMOST
DESTINATION	POLLUTED	DEVASTATING	SOMEDAY
DIOXIDE	CURB	DRASTIC	STRAIGHTLY
DROPLETS	GRIP	OPTIMUM	FUTHERMORE
ECOSYSTEM	LOGGING	COASTAL	CONTRASTLY
EXTINCTION	RECYCLING	ECOLOGICAL	DRASTICALLY
GREENHOUSE	OCCURING	UNCOUNTED	UNIVERSALLY
HABITAT	VOWED	ENDANGERED	BIOLOGICALLY
NUTSHELL	BANNED	HARSH	HOPEFULLY
OXYGEN	FINED	HAZARDOUS	FOREVER
PHOTOSYNTHESIS	RECYCLE	SHADDY	TREMENDOUSLY
PLANET	CATER	VITAL	MASSIVELY
POLLUTION	SPREADED	BIOLOGICAL	CONTINUALLY
SPECIES	CONVERGING	WARMER	CONS
CLIMATE	INSTIL	UNSCRUPULOUS	PROS
GLASIERS	PRACTICING	GASEOUS	UTTERLY
INSULATOR	IMPRESSED	RENEWABLE	UNINTENTIONALLY
OCCURENCE	EXECUTE	REPLENISHING	CRITICALLY
PENALTIES	REAPPLY	BELOVED	INVERSELY
POLE	INHERIT	INDISPENSABLE	NEVERMORE
RESPIRATION	INDULGE	CROOK	UNDENIABLY
RUMOUR	DEPLETED	IMBALANCED	SIMULTANEOUSLY
TIMBER	REUSE	NAKED	VICE -VERSA
TORNADO	WITHSTAND	AMAZING	
TYPHOONS	GRAB	ULTRAVIOLET	
ATMOSPHERE	PRECIPITATE	TOXIC	
DEHYDRATION	ENTITLED	TREMENDOUS	
DISASTER	CAPTURE	URBAN	
FAUNA	HIKING	ABIOTIC	
FLORA	CRAVING	BIOTIC	
INTAKE	COMBAT	HUGE	
INTERNET	BLOOM	UNEXPECTED	

KIDS	MODERNIZE	AQUATIC	
LANDSLIDE	OCCURED	FATAL	
POLLUTANT	OVERFLOW	AGGRESSIVE	
SURFING	URBANIZE	CONDUCIVE	
VEGETABLES	REPLENISHED	TROPICAL	
OZONE	DEPOSITED	SURPRISED	
TELEVISION	ABUSED	TERRESTRIAL	
DEFORESTATION	FED	DEPLORABLE	
PROVERB	LOOSING	CLIMATIC	
ABSORBER	ELABORATE	INERT	
BEACH	CONSERVE	WANTON	
FERNS	SYNTHESIZE	MASSIVE	
MANGROVE	STARVING	MEAGRE	
SWAMP	PENETRATE	DISASTROUS	
ALARM	SUBMERGED	EXISTENT	
CANCER	ERADICATE	CHRONIC	
ECOLOGY	NURTURE	UNTREATED	
FUELS	REVOKE	PANICKED	
HARMONY	SPEW	MOTIVATIONAL	
LAUNCHED	STRENGHTEN	OPTIMALLY	
ORGANISM	ABOLISHING	PROACTIVE	
RAINFALL	DESTRUCT	AWFUL	
SEDIMENTATION	COOING	RURAL	
CATCHMENT	MUTATED	RAINY	
DROUGHT	FULFILL	UNAFFORDABLE	
HYDROPOWER	OBLIGED	SPACIOUS	
METHANE	OXIDIZE	SUPERB	
NITROGEN	SADDENING	CALMER	
USAGE	JAILED	COMMUTE	
CHORES	ARMING	COPIOUS	
CLIMAX	DIFFUSE RECRUITING	HEART-WRENCHING	
DAM	PROTEST	HERBIVOROUS	
GENERATOR	FEED LOGGED23	LEGUMINOUS	
HOUSEBUILDING	ENERGIZED	LUSH	
PETROLEUM	PARROTING	STURDY	
PHOSPHORUS	DONATE	BIOCHEMICAL	
REFORESTATION	EXTINCT	WELL-DEVELOPED	

WEB	BOOMED	WELL-KEPT	
HAZARD	BREEDING	WEARY	
INTERRUPTION	VANISHED	DETRIMENTAL	
STARVATION	ALLEVIATE	GIANT	
VIRUS	DEPLETING	ABSURD	
ACID	MONOPOLIZED	SOLAR REVOLVING	
COUNSELLING	REPLANT	IRRITATING	
ELECTRONIC	EMBRACE	MIGHTY	
IMPRESSION	CHOPPED	UNCIVILISED	
JAM	DISRUPT	VAST	
TRAFFIC	EXHALE	OUTRAGEOUS	
ERA	LEACHING	NON-RENEWABLE	
ICEBERG	INHALE	DIRTIER	
TRAGEDIES	ERODED	VULNERABLE	
TSUNAMI	BULLY	DIRE	
DRAINS	REBUILDING SPANNING	GEOGRAPHICAL	
LOGGERS	REVOKING	GREEDIEST	
OFFSPRINGS	JEOPARDISE	ACIDIC	
CULPRITS	REVIVE	NON-LIVING	
DEPLETION	DETERIORATES	NON-STOP	
DOINGS	DIGEST	INTERDEPENDENT	
SYLLABUS	DISAPPEAR	CUTE	
TEXTBOOK	STUCKED	HAZY	
OFFENDERS	GASP	BALD	
CHLOROPHYLL	RESPIRE	PRICELESS	
JAIL	CHOPPING	RAMPANT	
PENALTY	OBSTRUCT	SELF-CENTRED	
CIRCULATION	TOOTS	WHOLE-	
CIVILIZATION	TREKKING	HEARTEDLY	
COMBUSTION	INFORCED	BIODEGRADABLE	
DRUGS	INVADING	SHADY	
MEDICATION	BUSTLES	ALARMING	
PLANTATION	RETREAT	KEEN	
MONSOON	PERISH	UTMOST	
ASSET	VANISHES	OVERWHELMING	
DIABETES	REINVENTING	SELF-CENTEREDNESS	
HERITAGE	INCULCATE	TOLERANCE	

JUNGLE	REAMING	WELL-BALANCED	
PLASTIC	CLOGGED	UNTOUCHED	
SQUAD	COPE	DELETERIOUS	
CONDITIONER	CRAWLING	INHUMANE	
THICKENING	EXCRETE	IRONIC	
BLOG	EMBED	MONETARY	
URBANISATION	REBALANCE	RESPIRATORY	
CATARACT	UNPLANNED	BREATHTAKING	
DINASOUR	CHIPPED	DEPENDABLE	
EXTICNTION	CRUMBLE	MAGNETIC	
HERBIVORES	DISTRACTED	ARROGANT	
HERBS	ENDANGERING	EXOTIC	
LIFESTYLE	SHRINK	NUTRIENTS	
MAMMALS	STOMP	ORGANIC	
REPTILES	DECOMPOSE	SPIRITUALLY	
SPAN	GLITTERING	SURPRISINGLY	
SYMPTOMS	HUSTLE	RADIOACTIVE	
ORPHANS	STROKING		
BARRIER			
LOVERS			
TRAGEDY			
MONOXIDE			
NUTRITION			
PROTECTOR			
TECHNOLOGIES			
GEOLOGIST			
TIGER			
AEROSOL			
ALGAE			
CHLOROFLUOROCARBON			
FISHERIES			
HYBRID			
HYDROFLUOROCARBON			
MARINE			
LEFTOVER			
REFRIGERATOR			
SEDIMENT			

CHEETAH			
HERBAL			
LEOPARD			
INCREMENT			
ASTHMA			
CRAB			
PROTEIN			
LITHOSPHERE			
BREED			
BEE			
HIVE			
HONEY			
DEERS			
DISAPPEARANCE			
GONERS			
INHABITANTS			
LIFESPAN			
LIONS			
MANKINDS			
POLAR			
WILDLIFE			
ZOO			
EROSIONS			
GOODS			
GUARDIAN			
TRAUMA			
ANCESTORS			
CARROTS			
RAINFOREST			
SEASHORES			
SEASIDES			
ALTITUDE			
CUBES			
DECOMPOSERS			
RUBRICS			
VICINITY			
COMPOSITIONS			

ORCHARD			
RECOMBINANT			
CANOPY			
CELL			
METABOLISM			
RADIATION			
TRIPHOSPHATE			
DEFICIENCY			
DISTRACTION			
HOLDER			
MISSION			
PIONEER			
REFLECTOR			
NITROGENOUS			
VIRGIN			
HAZE			
PERMISSIONS			
HEIGHT			
FOX			
HEADLINES			
INDUSTRIALIZATION			
ABUNDANCE			
GLACIERS			
HAVOC			
LOWLAND			
ERUPTION			
MICROORGANISMS			
ABSORPTION			
BOAR			
ACTVITIES			
VERGE			
BIOSPHERE			
HYDROSPHERE			
SPONGE			
VELOCITY			
APARTMENT			
CASINO			

CONNECTORS			
RANGERS			
CASUALTIES			
HIGHLANDS			
LONGEVITY			
PACE			
TROPICS			
WRATH			
ANTS			
BUCKS			
CONCRETE			
FOLKS			
LIZARD			
ADVICES			
FRESHY			
GALAXY			
GREENERY			
ORGANIZERS			
PANDA			
PEERS			
ROUTINE			
THANKFULNESS			
OBLIGATION			
BRONCHITIS			
CHARITY			
CLEARANCE			
FINES			
GIRAFFES			
HEADACHES			
MATING			
PHYTOPLANKTON			
PRECAUTION			
AMBIENCE			
BACTERIA			
BREEZE			
BURDEN			
FUNGI			

HOMELAND			
KOALAS			
LANDSCAPES			
MARVEL			
MYRIAD			
NODULES			
SQUIRRELS			
STALKS			
HEALTH-RELATED			
CAREER			
HOUSEHOLD			
LITTER			
BIOLOGIST			
ECO-TOURISM			
CORAL			
REEFS			
ROTAN			
TOURISM			
CONDITIONING			
INDUSTRIALISATION			
ALUMINIUM			
CONTINUATION			
DUMPSITES			
HOTSPOT			
MAGNITUDE			
TRASH			
ACTIVISTS			
LAWAS			
FLASHFLOOD			
HANDY			
HERO			
IMPRISONMENT			
SEDIMENTS			
TRANQUILITY			
FUMES			
HUMIDITY			
TUBERCULOSIS			

EQUILIBRIUM			
FILTER			
IRRESPONSIBILITIES			
RESERVOIR			
SAMARITAN			
SKYROCKET			
SMUG			
SOOTS			
ZEITGEIST			
MIGRAIN			
OLDIES			
RIVERBANK			
MODERNIZATION			
TISSUES			
BUNGALOWS			
MALLS			
SKYSCRAPERS			
CAB			
HOBBY			
HORNBILLS			
RAINWATER			
TRANSPIRATION			
POLYMERS			
CATASTROPHES			
HUMANITY			
RECREATIONAL			
SALAMANDERS			
SERENITY			
GUTS			
LOOP			
POLYSTYRENES			
SUPERMARKET			
PENETRATION			
CORROSION			
EMISSION			
HUMANKIND			
SULPHUR			

FOOTPRINT SURROUNDINGS TWIN FURNITURES ABORIGINALS COMPLIMENT EARTHLING ECO-CENTRISM ESSAY SANCTUARY STARCH CASUALITIES FATALITY NIGHTMARE OBSTACLES RAGE VOLCANO BOMB CIRCULATORY DISRUPTION GINGKO MINIATURE PARASITE PULCHRITUDE TOXINS WORKLOAD			
TOTAL: 355	**128**	**124**	**24**

APPENDIX E: STUDENTS QUESTIONNAIRE

Questionnaire Consent Form

The attached questionnaire was developed as part of my postgraduate dissertation. It intends to identify the challenges of matriculation students in writing English compositions. Your contribution will help in identifying the problems in writing among students in *PusatAsasiSainsUniversiti Malaya* (PASUM). You have been selected as the best source of information to contribute to this study by responding to the attached questionnaire.

The questionnaire is divided into three parts and consists of 5 pages. It is designed to be answered within 10 minutes. Please answer all the given questions accurately. Instructions are provided for each question. Please note that the information that you provide will be used for research and publication purposes only. Your name will be kept confidential in this research. Please inform that by completing this questionnaire you agree that the researcher is allowed and permitted to use the information that you provide for research and publication purposes only.

Your participation in this research study is completely voluntary. The researcher will answer any further questions about the research, now or during the course of the project, and can be reached by telephone at: 0146226490 or email at: charinelhs@hotmail.com. I am grateful and appreciate for your time in completing this questionnaire.

Thank you.

LEE HUI SHEIN

Part A: Personal Information

Name: _____

Matriculation number:_____

Gender (circle one): M F

Part B: Language Background

1. Years of studying English language:_____
2. Number of languages you speak: _____
3. Highest academic qualification:_____
4. What grade did you achieve for English language in the SPM and MUET examinations?

   _____

5. Which language do you use to communicate with your family members?

   ❑Malay ❑English

   ❑Mandarin ❑Others, please state _____

6. Which language do you use to communicate with your friends?

   ❑Malay ❑English

   ❑Mandarin ❑Others, please state _____

7. Which language do you use to communicate with lecturers?

❑Malay ❑English

❑Mandarin ❑Others, please state _____

8. How often do you speak English?

❑Everyday ❑2-3 times a month

❑2-3 times a week ❑Never

9. How often do you write in English (excluding writing academically)?

❑Everyday ❑2-3 times a month

❑2-3 times a week ❑Never

10. How often do you read English newspapers?

❑Everyday ❑2-3 times a month

❑2-3 times a week ❑Never

11. How often do you read English magazines?

❑Everyday ❑2-3 times a month

❑2-3 times a week ❑Never

12. How often do you read English books?(excluding school reference book/textbook)

❑Everyday ❑2-3 times a month

❑2-3 times a week ❑Never

13. Why do you learn English?

❑It is required in the education
❑I love English
❑To communicate better
❑To further my studies
❑To get a better job
Others, please state _____

14. Which level of the *writing* English language skill enables you to perform well in your academic effectively? How do you rate your English writing ability?

❑Very good
❑Good
❑Satisfactory
❑Very poor
❑Poor

15. Rank the following English language skills in terms of importance in your studies. Please rank using numbers from

[most important = 4; important = 3; fair = 2; not important =1]

Listening _____

Speaking _____

Reading _____

Writing _____

16. Rank the following English language skills in terms of difficulty you have in your studies. Please rank using numbers from

[most difficult = 4; difficult = 3; moderate = 2; easy =1]

Listening _____

Speaking _____

Reading _____

Writing _____

Part C: Perceptions on English Language

17. How important is it to have a high level of English proficiency when performing the following activities? Please circle the number that corresponds to your answer.

[most important = 4; important = 3; fair = 2; not important =1]

		Listening	Speaking	Reading	Writing
a)	Phone Conversations	1 2 3 4	1 2 3 4		
b)	Asking questions in class	1 2 3 4	1 2 3 4		
c)	Assignments			1 2 3 4	1 2 3 4
d)	Forms/Applications			1 2 3 4	1 2 3 4
e)	Reports			1 2 3 4	1 2 3 4
f)	Emails and Faxes			1 2 3 4	1 2 3 4
g)	Accessing Internet	1 2 3 4		1 2 3 4	
h)	Class presentations	1 2 3 4	1 2 3 4	1 2 3 4	1 2 3 4
i)	Others, please state_____ _____	1 2 3 4	1 2 3 4	1 2 3 4	1 2 3 4

18. How would you rate your knowledge of English before you began college? Please tick one answer.

❑ Very good
❑ Good
❑ Satisfactory
❑ Very poor
❑ Poor

19. How did the Malaysian University English Test (MUET) lessons that you studied help you in the following tasks? Please look at the scale below and circle the appropriate number accordingly.

[a lot = 4; somewhat = 3; a little = 2; did not help at all =1]

a) Converse in English
1 2 3 4

b) Write essays in English
1 2 3 4

c) Read books, articles, and magazines in English Language
1 2 3 4

d) Understand classroom instructions, lectures, and homework
1 2 3 4

e) Listen to radios and conversation in English language
1 2 3 4

20. How would you rate your knowledge of English after you attended MUET lessons? Please tick one answer.

❑ Very good
❑ Good
❑ Satisfactory
❑ Very poor
❑ Poor

Thank you for your participation

APPENDIX F: TEACHERS QUESTIONNAIRE

Questionnaire Consent Form

The attached questionnaire was developed as part of my postgraduate dissertation. It intends to identify the challenges of matriculation students in writing English compositions. Your contribution will help in identifying the problems in writing among students in *PusatAsasiSainsUniversiti Malaya* (PASUM). You have been selected as the best source of information to contribute to this study by responding to the attached questionnaire.

The questionnaire is divided into two parts and consists of three pages. It is designed to be answered within 15 minutes. Please answer all the given questions accurately. Instructions are provided for each question. Please note that the information that you provide will be used for research and publication purposes only. Your name will be kept confidential in this research. Please inform that by completing this questionnaire you agree that the researcher is allowed and permitted to use the information that you provide for research and publication purposes only.

Your participation in this research study is completely voluntary. The researcher will answer any further questions about the research, now or during the course of the project, and can be reached by telephone at: 0146226490 or email at: charinelhs@hotmail.com. I am grateful and appreciate for your time in completing this questionnaire.

Thank you

LEE HUI SHEIN

Instructions: Please read and answer the questions in the order they appear in the questionnaire.

A) Background Information

1. Name:_____

2. Native speaker/ Non-native speaker of English Language

3. Male/Female

4. Years of teaching experience:_____

5. Hours of teaching MUET per week:_____

6. Highest academic qualification:_____

B) Please answer the following questions:

1. What are the difficulties you have in teaching MUET?

_____

2. How frequent do students write compositions?

_____

3. What kind of composition topics do you usually ask the students to write?

_____

4. Do you adhere to any specific written criteria when marking essays? If yes, please provide a short description below.

_____

5. If no to question 4, how do you decide what grade to give when marking an essay?

_____

6. What are the problems you faced when marking students' essays? Please comment.

_____

7. Within language how much weight do you put on lexical features/vocabulary? (Please give a rough estimate as a percentage, e.g. 50% grammar/50% vocabulary)

_____

8. Do you think a student can have a good essay with poor grammar? Please explain.

Yes No

_____

9. Do you agree that advanced vocabulary (low frequency words) is important for students to score better in writing essay? Please comment.

Yes No

_____

10. Do you teach vocabulary specifically in class? If yes, please state the allocation of time for students to learn vocabulary. If no, please proceed to question 12.

_____

11. What are the challenges in teaching vocabulary to students?

_____

12. How do you think you can improve students' vocabulary?

_____

13. The following are the different vocabulary features of a good written work. (Please rank the vocabulary features according to its importance from 1 being the least important to 5 being the most important)

☐Appropriate use of words

☐ Variation

☐ Advanced vocabulary

☐ Idiomatic collocations

☐Appropriate writing style

14. Based on your teaching experience, what is the most difficult language skill experienced by the students? Please explain.

_____

Thank you for your time!

APPENDIX G: A SAMPLE OF STUDENT WRITTEN ESSAY (CATEGORY A)

FB0111(4)/06

plant → O_2 →

present landslide, no flood occured

Question 2

*You are advised to spend about **50 minutes** on this task*

Preserving the forests is the key to saving our planet. What is your opinion on this statement?

You should write at least 350 words. [60 marks]

~~Nowadays, our planet are harmed with pollutions~~

Nowadays, pollutions is the most popular issue in newspapers and telenscas too. It is harmed our planet uncontrully. Air, pollution, water pollution, global warming, green house effect are occured around us. Should we just keep quite and not do anything ~~in order to save our world~~? As a responsible person, we have to think creatively and in work progressively to save our planet. Why we need to do so? Because planet is just like our home ~~place for~~. So, preserving the forests is the key to saving our planet.

~~As we know, Malaysian is a developed country. Now, we are work hard to develop our country to fulfill our mission to get title as Wawasan 2020. So,~~ ~~just like Japan and America~~

Why I said so? It is because, as we know forest have a variety of plants. Plants are very important in our life. It can do photosynthesis process which release oxygen gas to atmosphere. As we know, human need oxygen in the presence of chlorophyll.

~~are very im~~ to live in the earth. Without plant, the percentage of carbon dioxide in the atmosphere will ~~is~~ increase~~d~~. For example, in rural areas, ~~there are lo~~ it have ~~they are~~ alot of forest that we can found compared in metropolitan cities like Kuala Lumpur and Johor Bahru. We can feel that air in ~~differentiate how freshness of the air~~ in rural area are more fresh compared to in metropolitan city. ~~Be~~ Because that, we ~~can~~ will feel relax and rest & living ~~peace~~ in rural area than in city.

Besides, ~~forest is im~~ we need to preserve forest because to ensure our planet free from global warming. This is because, if our ~~forest a~~ plants are being cut, so the sun ray will directly approaching to the earth. This contributes to the increase of earth's temperature. So, human will feel hot. ~~Be~~ Moreover, this situation will ~~gi~~ give negatife effect to human like cataracts, skin cancer and droughness ~~others~~. If the earth is too hot, it is not favouirable for human to do their daily activities smoothly.

In addition, ~~the~~ forest are need to preserve because to save our planet from flashflood ~~flashflood~~. ~~and if let's~~ It is because if our forest are not preserve, so during the raining day, landslides will be occured. Why this problem happened? It is because, there is no more plants to absorb the water into the soil, so this leads to landslide. The impact is, flash flood

will be occured. But if there is plant it will occured reverse. ~~If in coming day, the roots of plants will absorb the water by the root. The~~ Flash flood ~~can cause~~ gives the opportunities for the spreading of disease as our world is not clean anymore. ~~Besides, a lot of houses or vehicles will~~

As a conclusion, I agree with the statement that preserving forest is the key to saving our planet. Planet is the gift from God that we need to appreciate it. ~~If we are~~ Planets is place for us to live, place for us to make money and place for us to enjoy our life in this world. What will happen if our planet is not save anymore?. Disease will spread among the human, animals and aquatic plants will die and ~~life also cannot~~ no development can occur in that country. So, as a responsible human, we need to take our steps from now and work together with the government on how to save our world from pollution and ~~many~~ others for our happiness.

TF 4 0 26/60

APPENDIX H: A SAMPLE OF STUDENT WRITTEN ESSAY (CATEGORY B)

FB0111(4)/06

Question 2

*You are advised to spend about **50 minutes** on this task*

Preserving the forests is the key to saving our planet. What is your opinion on this statement?

You should write at least 350 words. [60 marks]

~~Nowadays, depletion of forest has become a big issue in our country. As we know, Malaysia has large area of tropical forest. However, deforestation activity has make the large area of forest reduce days by days. International media.~~

Nowadays, our planet is getting sick. There are so many activities done by human in order to improve the quality of life. They cut trees to build factories, develop residential areas and set up a shopping malls. It is done under the name of development. All these activities has make our planet getting weak. We, as the guardian of the Earth must do something in order to save our world. I strongly believe that preserving the forest is the key to save our planet because forest ~~trees are~~ is needed to balance earth temperature.

provide habitat to animals and plants and prevent environmental disaster.

First and foremost, forest is needed to balance earth temperature. It is means that forest can reduce the hot earth's temperature that increase gradually increase in the period of two decades ago. Science studies pro show that plants produce oxygen gas during the process of photosynthesis. It also know that bigger trees are the largest oxygen supply large When we replant more trees, the proportion of oxygen gas that increase in earth atmosphere will reduce balance the high composition of carbon dioxide in atmosphere which contribute to the increases of ~~the hot temperature~~ ~~Therefore, forest should preserved in order to help our plant~~ earth's temperature. ~~to reduce decrease it hotness~~ Hence, we need to replant more trees in order to help our planet on reducing it's temperature.

Secondly, to by preserve the forest, we can provide the habitat for plvariety of plants and animals. This is because, floral and fauna has lost their homes ~~as forest~~ due to deforestration activities. Moreover, the illegal deforestration activities has contribute to the extinctances of wild life. By & preserving the forest, we can save these ~~wild life~~ earth's property and let our future generation get to know all the & unique wild life. Therefore, we should replant the forest to avoid our planet lost its property of animals and plants.

Last but not least, preserving the forest can save our ~~plant~~ planet as it can prevent environmental disaster. This is due to the fact that less number of trees can be the reason of flash flood and soil erosion that occur each time ~~when~~ after rain. ~~Soil flash flood and soil erosion will cause a lot of problems such as eroded the sea bottom death and reduce.~~ Each time these disaster happen, human face a lot of problem such as ~~to~~ lost of electricity supply and lost of food production, ~~rate~~. Hence, preserving the forest can help our planet to overcome this matter. ~~as the root of the plant can prevent flash flood and~~

In short, the reasons I have mentioned why preserving the forest is the key ~~to~~ to save our plant ~~is~~ are, forest ~~are is needed to~~ can balance earth temperature, provide animal and plants their habitat and ~~prevent~~ environmental disaster. I strongly ~~felt~~ that human should play the role ~~is to ensure that we can~~ the earth is the best place to live on. Our government should highlight this issue ~~on~~ media mass ~~by d~~ so that people will aware ~~to~~ play the role in saving our planet lovely.

TP 31/60

APPENDIX I: A SAMPLE OF STUDENT WRITTEN ESSAY (CATEGORY C)

FB0111(4)/06

Question 2

*You are advised to spend about **50 minutes** on this task*

Preserving the forests is the key to saving our planet. What is your opinion on this statement?

You should write at least 350 words. [60 marks]

Much has been heard recently about the terrible effects that global warming brings to our earth. Some of the effects of global warming are ~~the~~ thinning in the atmosphere layer, green house effect, drastic increase in temperature and melting of ice ~~at~~ at the poles. We as legal guardians of Mother Nature, should find ways in protecting our earth. /57

~~One of the many ways to save our planet is by preserving the forests.~~

One of the many ways to save our planet is by preserving the forests. Preserving the forests is the main key ~~to~~ in saving our planet. Clearing and removing ~~tres~~ ~~tress~~ trees in the forests can cause a lot of harm to ~~et~~ a lot of parties. For example,

animals that live in the forests. If we destroy the forest, the animals would not have anywhere ~~a~~ to stay. ~~The forest is the animal's habitat.~~ The forest is a home to the animals, just like a house is a home to us humans. How would you like ~~if~~ if somebody destroy your house for no good ~~reason~~ reason? ~~Be It is important for us to preserve the forests, so that the~~ In preventing this from happening, preservation of the forests should takes places.

Preserving the forests has many benefits. ~~For~~ For instance, increase in the quality of our breathing air. When forests are preserve, the quality of air that we breathe in would be better. This is because when forests are preserve, more plants would give off ~~clean~~ oxygen that can clean the dirty air around us.

Through preserving the forests, we can also keep our ecosystem balanced. This is because when we preserve the forests, the animals' habitat would not be destroyed. Hence, extinction of animals would not occur. At the same time, our ecosystem is kept balanced.

Other than that, by preserving the forests, we can

prevent
~~save~~ the human era from ending. As we know it, humans need oxygen in order to survive or in other words, to live. without plants or trees, the world would be lacking of oxygen for the human to breathe in. when this happens, that would be the end of the human era as we know it.

In order to live longer, ~~to~~ we as guardian ~~of~~ of the earth should definitely preserve the forests. There are many advantages would come by doing so. saving the planet is our obligation as citizen of the earth. Do not let our grandchildren or great grandchildren miss out on the beauty and the wonders that the Mother Nature gets to offer. It would be a sad thing, if they only manage to see grey skies and bare hills.

10

/60

APPENDIX J: A SAMPLE OF STUDENT WRITTEN ESSAY (CATEGORY D)

FB0111(4)/06

Question 2

*You are advised to spend about **50 minutes** on this task*

Preserving the forests is the key to saving our planet. What is your opinion on this statement?

You should write at least 350 words. [60 marks]

"Preserve our forest and save the earth". We often hear this quote everywhere. But why are we asked to save the earth? What is happening to our earth ? These are some of the questions that run through our mind. Nowadays, there are many disasters occuring in our planet such as global warming, climate change, green house effects and the rising of sea level. Have we ever thought why are all these happening ? Is it because of our doings or is it just occuring by itself?

In my opinion, it is due to the forest that is not taken care of enough by man. For example, unplanned forest cutting is practised as well as open burning activity.

These activities causes the forest to be extinct. As we all know, the forest plays an important role in all our lives. It maintains the continuity of live on earth.

Among some of the importance of the forest to our planet are it is the habitat for flora and fauna. We can find many rare and valuable flora and fauna species in the forest. The forest gives these species shelter and prevents them from extinction. However, when the forest is not taken care of and destroyed, the flora and fauna are also destroyed along with it. Therefore, we loose the valuable life on earth.

Besides that, the forest helps balance the content of oxygen in the air which will indirectly cool down the earth. The plants in the forest releases oxygen when they undergo the photosynthesis process and absorbs carbon dioxide. When the trees in the forest are cut down, the carbon dioxide content in the air increases and this reduces the content of oxygen on earth. Indirectly, the

temperature of the earth rises and this cause climate change, global warming and also the melting of ice.

If the forest is preserved, all this disasters can be prevented from occuring. We can preserve the forest by implementing and enforcing rules for those who practise unplanned forest cutting and also open burning. Besides that, we can form a group that takes special care of the forest in case anything happens.

In conclusion, if all the drastic, concrete and effective measures are taken to preserve the forest, we will be able to save the planet. We should not be selfish and think of our own benefit only but also the benefit of others as well. In short, I do agree that preserving the forests is the key to saving our planet.

TF = 25

LTO = 29

54

APPENDIX K: A SAMPLE OF LEXICAL FREQUENY PROFILE (LFP) ANALYSIS

Processing file: C:\Users\charinelhs\Desktop\Research Data\30-39 (31)\FAH111045 (31).txt

Number of lines: 1

Number of words: 402

Reading: C:\Users\CHARIN~1\AppData\Local\Temp\Rar$EX06.415\BASEWRD1.txt

Reading: C:\Users\CHARIN~1\AppData\Local\Temp\Rar$EX06.415\BASEWRD2.txt

Reading: C:\Users\CHARIN~1\AppData\Local\Temp\Rar$EX06.415\BASEWRD3.txt

WORD LIST	TOKENS/%	TYPES/%	FAMILIES
one	314/78.11	137/69.54	118
two	36/ 8.96	21/10.66	18
three	22/ 5.47	17/ 8.63	15
not in the lists	30/ 7.46	22/11.17	?????
Total	402	197	151

Number of BASEWRD1.txt types: 4119 Number of BASEWRD1.txt families: 998

Number of BASEWRD2.txt types: 3708 Number of BASEWRD2.txt families: 988

Number of BASEWRD3.txt types: 3107 Number of BASEWRD3.txt families: 570

Table of Ranges: Types

197 Words appear in 1 input files

Table of Ranges: Families

151 Words appear in 1 input files

Types Found In Base List One

TYPE	RANGE	FREQ	F1
A	1	1	1
ABLE	1	1	1
AGAIN	1	1	1
ALL	1	1	1
ALSO	1	5	5
AMONG	1	1	1
AMOUNT	1	1	1
AND	1	16	16
ANIMALS	1	2	2
ARE	1	4	4
AS	1	2	2
AT	1	3	3
BE	1	3	3
BEING	1	1	1
BELIEVE	1	1	1
BESIDES	1	1	1
BIGGER	1	1	1
BOTH	1	1	1
BRING	1	1	1
BUILDINGS	1	1	1
BURNING	1	2	2
BY	1	4	4
CALLED	1	1	1
CAN	1	1	1
CASE	1	1	1
CAUSE	1	4	4
CAUSES	1	1	1
DANGEROUS	1	1	1
DESERTS	1	1	1
DO	1	1	1
DONE	1	1	1

DRIED	1	1	1
DUE	1	2	2
EARTH	1	6	6
ENOUGH	1	1	1
EVERY	1	1	1
EVERYONE	1	1	1
FOOD	1	1	1
FOR	1	4	4
FOREST	1	2	2
FORESTS	1	10	10
FROM	1	3	3
FULL	1	2	2
GAS	1	2	2
GETTING	1	1	1
HAD	1	2	2
HAPPENED	1	1	1
HAVE	1	1	1
HEAT	1	1	1
HELP	1	2	2
HIGHER	1	1	1
HOLD	1	1	1
HOW	1	1	1
HUMAN	1	5	5
I	1	1	1
IMPORTANT	1	2	2
IMPORTANTLY	1	1	1
IN	1	6	6
INCREASING	1	2	2
IS	1	10	10
IT	1	1	1
ITSELF	1	1	1
KIND	1	1	1
KNOW	1	1	1
LAND	1	3	3
LATELY	1	1	1

LEAD	1	1	1
LEADERS	1	1	1
LESSER	1	2	2
LEVEL	1	1	1
LOWER	1	1	1
MANY	1	3	3
MAY	1	1	1
MORE	1	1	1
MOST	1	1	1
MUCH	1	1	1
MY	1	1	1
NATURE	1	1	1
NO	1	1	1
NORTH	1	1	1
NOWADAYS	1	1	1
OF	1	16	16
ON	1	3	3
ONE	1	1	1
ONLY	1	1	1
OPEN	1	2	2
ORDER	1	1	1
ORGANISATIONS	1	1	1
OTHER	1	1	1
OUR	1	10	10
OURSELVES	1	1	1
OUT	1	1	1
OWN	1	1	1
PERSON	1	1	1
PLACES	1	1	1
PLANTS	1	6	6
POPULATION	1	1	1
PRESENCE	1	1	1
PROVIDE	1	1	1
RISING	1	1	1
SAVE	1	2	2

SAVING	1	3	3
SEA	1	1	1
SHOULD	1	2	2
SHOWS	1	1	1
SIDE	1	1	1
SINGLE	1	1	1
SO	1	1	1
SOUTH	1	1	1
STAYING	1	1	1
STRONGLY	1	1	1
SUCH	1	1	1
SUNLIGHT	1	1	1
SUPPLIED	1	1	1
TAKEN	1	1	1
THAT	1	4	4
THE	1	21	21
THEM	1	1	1
THESE	1	1	1
THEY	1	2	2
THIRD	1	1	1
THIS	1	5	5
TO	1	13	13
TOP	1	1	1
TOWN	1	1	1
TREES	1	1	1
TWICE	1	1	1
US	1	5	5
USED	1	1	1
VIEW	1	1	1
WATER	1	2	2
WE	1	5	5
WHICH	1	4	4
WHY	1	1	1
WILL	1	4	4
WITH	1	2	2

WITHOUT	1	1	1

Types Found In Base List Two

TYPE	RANGE	FREQ	F1
ACCIDENTS	1	1	1
AVOIDED	1	1	1
BARELY	1	1	1
BREATHE	1	1	1
BREATHING	1	1	1
COMPOSED	1	1	1
CREATURES	1	2	2
ESPECIALLY	1	2	2
GREEDY	1	2	2
ICE	1	1	1
KEY	1	2	2
MELTED	1	1	1
NECKS	1	1	1
PRESERVE	1	3	3
PRESERVED	1	1	1
PRESERVING	1	6	6
RAINING	1	1	1
SOIL	1	4	4
TALL	1	1	1
TEMPERATURE	1	2	2
WARMING	1	1	1

Types Found In Base List Three

TYPE	RANGE	FREQ	F1
AREA	1	1	1
AREAS	1	1	1
BENEFITS	1	2	2
COMMUNITIES	1	1	1

CONTRIBUTES	1	1	1
CONVERT	1	1	1
COOPERATION	1	1	1
ERODED	1	1	1
EROSION	1	3	3
EXPLOIT	1	1	1
GLOBAL	1	1	1
HENCE	1	1	1
ILLEGAL	1	1	1
MAINTAINING	1	3	3
PHENOMENON	1	1	1
SOURCE	1	1	1
UNDERGO	1	1	1

LIST OF FAMILY GROUPS

BASE ONE FAMILIES	RANGE	TYFREQ	FAFREQ	F1
A	1	1	1	1
ABLE	1	1	1	1
AGAIN	1	1	1	1
ALL	1	1	1	1
ALSO	1	5	5	5
AMONG	1	1	1	1
AMOUNT	1	1	1	1
AND	1	16	16	16
ANIMAL	1	0	2	2
AS	1	2	2	2
AT	1	3	3	3
BE	1	3	18	18
BELIEVE	1	1	1	1
BESIDE	1	0	1	1
BIG	1	0	1	1
BOTH	1	1	1	1
BRING	1	1	1	1

BUILD	1	0	1	1
BURN	1	0	2	2
BY	1	4	4	4
CALL	1	0	1	1
CAN	1	1	1	1
CASE	1	1	1	1
CAUSE	1	4	5	5
DANGER	1	0	1	1
DESERT	1	0	1	1
DO	1	1	2	2
DRY	1	0	1	1
DUE	1	2	2	2
EARTH	1	6	6	6
ENOUGH	1	1	1	1
EVERY	1	1	2	2
FOOD	1	1	1	1
FOR	1	4	4	4
FOREST	1	2	12	12
FROM	1	3	3	3
FULL	1	2	2	2
GAS	1	2	2	2
GET	1	0	1	1
HAPPEN	1	0	1	1
HAVE	1	1	3	3
HEAT	1	1	1	1
HELP	1	2	2	2
HIGH	1	0	1	1
HOLD	1	1	1	1
HOW	1	1	1	1
HUMAN	1	5	5	5
I	1	1	2	2
IMPORTANT	1	2	3	3
IN	1	6	6	6
INCREASE	1	0	2	2
IT	1	1	2	2

KIND	1	1	1	1
KNOW	1	1	1	1
LAND	1	3	3	3
LATE	1	0	1	1
LEAD	1	1	2	2
LESS	1	0	2	2
LEVEL	1	1	1	1
LOW	1	0	1	1
MANY	1	3	3	3
MAY	1	1	1	1
MORE	1	1	1	1
MOST	1	1	1	1
MUCH	1	1	1	1
NATURE	1	1	1	1
NO	1	1	1	1
NORTH	1	1	1	1
NOW	1	0	1	1
OF	1	16	16	16
ON	1	3	3	3
ONE	1	1	1	1
ONLY	1	1	1	1
OPEN	1	2	2	2
ORDER	1	1	1	1
ORGANIZE	1	0	1	1
OTHER	1	1	1	1
OUT	1	1	1	1
OWN	1	1	1	1
PERSON	1	1	1	1
PLACE	1	0	1	1
PLANT	1	0	6	6
POPULATION	1	1	1	1
PRESENT	1	0	1	1
PROVIDE	1	1	1	1
RISE	1	0	1	1
SAVE	1	2	5	5

SEA	1	1	1	1
SHOULD	1	2	2	2
SHOW	1	0	1	1
SIDE	1	1	1	1
SINGLE	1	1	1	1
SO	1	1	1	1
SOUTH	1	1	1	1
STAY	1	0	1	1
STRONG	1	0	1	1
SUCH	1	1	1	1
SUN	1	0	1	1
SUPPLY	1	0	1	1
TAKE	1	0	1	1
THE	1	21	21	21
THEY	1	2	3	3
THIS	1	5	10	10
THREE	1	0	1	1
TO	1	13	13	13
TOP	1	1	1	1
TOWN	1	1	1	1
TREE	1	0	1	1
TWO	1	0	1	1
USE	1	0	1	1
VIEW	1	1	1	1
WATER	1	2	2	2
WE	1	5	21	21
WHICH	1	4	4	4
WHY	1	1	1	1
WILL	1	4	4	4
WITH	1	2	2	2
WITHOUT	1	1	1	1

BASE TWO FAMILIES	RANGE	TYFREQ	FAFREQ	F1
ACCIDENT	1	0	1	1

AVOID	1	0	1	1
BARELY	1	1	1	1
BREATHE	1	1	2	2
COMPOSE	1	0	1	1
CREATURE	1	0	2	2
ESPECIAL	1	0	2	2
GREED	1	0	2	2
ICE	1	1	1	1
KEY	1	2	2	2
MELT	1	0	1	1
NECK	1	0	1	1
PRESERVE	1	3	10	10
RAIN	1	0	1	1
SOIL	1	4	4	4
TALL	1	1	1	1
TEMPERATURE	1	2	2	2
WARM	1	0	1	1

BASE THREE FAMILIES	RANGE	TYFREQ	FAFREQ	F1
AREA	1	1	2	2
BENEFIT	1	0	2	2
COMMUNITY	1	0	1	1
CONTRIBUTE	1	0	1	1
CONVERT	1	1	1	1
COOPERATE	1	0	1	1
ERODE	1	0	4	4
EXPLOIT	1	1	1	1
GLOBE	1	0	1	1
HENCE	1	1	1	1
LEGAL	1	0	1	1
MAINTAIN	1	0	3	3
PHENOMENON	1	1	1	1
SOURCE	1	1	1	1
UNDERGO	1	1	1	1

Types Not Found In Any List

TYPE	RANGE	FREQ	F1
ALTITUDE	1	1	1
CARBON	1	1	1
COASTAL	1	1	1
CUBES	1	1	1
DECOMPOSERS	1	1	1
DEFORESTATION	1	3	3
DETRITIVORES	1	1	1
DIOXIDE	1	1	1
DRAIN	1	1	1
ECOSYSTEM	1	2	2
HABITAT	1	1	1
HERBAL	1	1	1
NUTSHELL	1	1	1
OXYGEN	1	2	2
PHOTOSYNTHESIS	1	1	1
PLANET	1	5	5
POLES	1	1	1
PRECAUTION	1	1	1
SUBMERGED	1	1	1
TIMBER	1	1	1
URBANISATION	1	1	1
WILDLIFE	1	1	1

time taken was : 8 Seconds

Number of cache nodes read:	41603
Number of cache nodes written:	2749
Number of disk nodes read:	0
Number of disk nodes written:	0
Number of nodes per second,	0
Total Number of words,	402

Number of words per second,	50
Number of unique words in tree,	197
Number of unique words per second,	24
Memory used,	0.96 MB
Size of node,	316 bytes
Memory used by nodes,	0.06 MB

...Finished