



University of Bahrain

College of Arts

Department of English Language and Literature

**Measuring the Lexical Complexity of University Students in EAP and
ESP Contexts: A Comparative Study**

A Thesis Submitted in Partial Fulfillment of the Requirements for the
Master's Degree in Applied English Language Studies

Prepared by:

Zainab Nader Mahmood

20135181

Supervised by:

Dr. Diana AbdulKarim Aljahromi

Assistant Professor of Linguistics

University of Bahrain

Kingdom of Bahrain

April 2023



University of Bahrain

College of Arts

Department of English Language and Literature

M.A Program in English Language Applied Studies

Defense Committee's Approval

Student **Zainab Nader Abdulretha**, ID number **20135181** defended his/her thesis by which he/she fulfilled the requirements for the degree of M.A in English Language Applied Studies, entitled **(Measuring the Lexical Complexity of University Students in EAP and ESP Contexts: A Comparative Study)** on this day of 26th June 2023

The committee recommends awarding of M.A Degree with the grades of: A.

(Defense Committee)

Name

Signature

Dr. Diana Al Jahrami

Supervisor

on behalf of Dr Diana

Dr. Madani Osman

Internal Examiner

Prof. Abdul Rahman Saleh

External Examiner

Declaration

I declare that this work, which I submitted to the University of Bahrain, contains my own investigation. This research has not been previously accepted for any degree and is not currently considered for any other degree at any other university.

Name: Zainab Nader Abdulretha

ID Number: 20135181

Date: May 2023

The thesis defence committee considered the thesis title:

Measuring the Lexical Complexity of University Students in EAP and ESP Contexts: A Comparative Study

Submitted to the Department of English Language and Literature for the degree of Master in Applied English Language Studies.

Name

Dr. Diana Al Jahrami

Dr. Madani Osman

Prof. Abdul Rahman Saleh

Signature

Supervisor

Internal Examiner

External Examiner

on behalf of Dr Diana
Signature
Signature
Signature
Signature

Acknowledgments

First and foremost, my deepest and sincerest gratitude and thanks go to God for giving me the endurance, strength, motivation, and confidence to move forward with this process. Without the support of God, none of this thesis would have been possible, and without the very command of the Divine, this entire process would not have commenced nor have been completed.

Many have stood by my side throughout this process, and I would like to take the time to thank this group of wonderful people for all their efforts in pushing me towards the completion of this thesis. I would like to extend my gratitude to Dr Diana Al Jahromi, my supervisor and my mentor who guided me with utmost patience, support, and resolve. I would also like to thank the educators and members of the academic committee who have been a continuous source of support during this journey; Dr Mohammed Redha Qader, Dr Madani Osman, Dr Nuri Ageli, and Dr Martin Parker.

I would also like to convey my gratitude towards my parents for supporting and caring for me over the course of my life and always pushing me towards achieving more, becoming more, and believing that I will always be more. I would also like to thank my grandparents for their prayers and gentle gestures of kindness in supporting me and sheltering me.

Next, I would like to thank my dearest aunt, who has not only been my greatest source of benevolent faith and guidance but has also been my peer and classmate throughout this entire journey. I would like to thank her for her trust,

for the company she gave me during hardship, and for every challenge we overcame together during this process.

A special word of gratitude also goes to my dear partner, a man who picked me up and pushed me towards doing my best, even at times when I barely had the strength to do so.

My final word of gratitude goes to my closest childhood friends who have not only been a source of support and encouragement, but also role models, whose own efforts towards their education motivated me to accomplish my own goals.

Abstract

The research for this thesis explored and compared the degree of lexical complexity demonstrated in the written discourse of EAP and ESP learners studying at university in the Kingdom of Bahrain. A total of nine thousand five hundred words were analyzed from learners across each context and varying methods of quantitative analysis were conducted to measure the lexical density, diversity, and sophistication of the discourse produced. The results were then further analyzed using the SPSS software, which demonstrated that the results were indeed significant, with EAP learners attaining slightly higher lexical density in comparison to the ESP learners. When lexical sophistication was considered, the TAALES 2.0 program was utilized to compare the written discourse of the learners in contrast to the British National Corpus, the Brown Verbal Frequency, as well as N-gram indices. Further qualitative analysis was also conducted to explore the perspectives of learners from both EAP and ESP contexts towards lexical complexity.

Keywords: EAP, ESP, lexical density, lexical diversity, lexical sophistication

Table of Content

Topic	Page No.
Defence Committee's Approval Page	A
Declaration	B
Acknowledgments	C
Abstract	E
Table of Contents	G
List of Tables	I
List of Figures	J
List of Abbreviations	K

Table of Content

Topic	Page No.
Defence Committee's Approval Page	A
Declaration	B
Acknowledgments	C
Abstract	E
Table of Contents	F
List of Tables	H
List of Figures	I
List of Abbreviations	J

Table of Contents

CHAPTER ONE		
Introduction		
1.1	Background Information	2
1.1.1	English for Academic Purposes Versus English for Specific Purposes	3
1.1.2	Lexical Complexity	4
1.2	Research Problem	7
1.3	Research Questions	8
1.4	Purpose of the Research	9
1.5	Methodology	9
1.6	Significance of the Research	10
CHAPTER TWO		

LITERATURE REVIEW		
2.1	Literature Review	12
2.1.1	Theoretical Background	12
2.1.1.1	English for Academic Purposes Versus English for Specific Purposes	12
2.1.1.2	Measures of Lexical Complexity	17
2.2.1.2.1	Lexical Density	17
2.2.1.2.2	Lexical Diversity	19
2.2.1.2.3	Lexical Sophistication	21
2.1.2	Previous Studies	23
CHAPTER THREE METHODOLOGY AND PROCEDURES		
3.1	Methodology	31
3.1.1	Participants	31
3.1.2	Quantitative Data	31
CHAPTER FOUR FINDINGS AND DISCUSSION		
4.1	Lexical Diversity (TTR) Against Lexical Density	35
4.1.1	Lexical Diversity (TTR) Against lexical Density of EAP Learners	35
4.2	Type Token Ratio of ESP Learners Across Divisions	50
4.2.1	Lexical Diversity Analysis: Type Token Ratio Against Lexical Density	51
4.3	Comparative Analysis of EAP and ESP Samples	67
4.3.1	Statistical Analysis of EAP and ESP Samples	67
4.4	Lexical Sophistication	70
4.5	Interview Analysis	78
4.5.1	EAP	78
4.5.2	ESP	88
CHAPTER FIVE CONCLUSION		
5.1	Introduction	95
5.2	Limitations and Delimitations	98
REFERENCES		
APPENDIX		

List of Tables

4.1	Statistical Comparison of EAP and ESP Lexical Density	67
<u>4.2</u>	<u>Leven's Test for Equality of Variances on Lexical Density</u>	68
4.3	Statistical Comparison of EAP and ESP Type Token Ratio	69
4.4	Analysis of TAALES Word Range Amongst EAP and ESP Contexts for Lexical Sophistication	72
4.5	Analysis of TAALES Bigrams Amongst EAP and ESP Contexts for Lexical Sophistication	75
4.6	Analysis of TAALES Trigrams Amongst EAP and ESP Contexts for Lexical Sophistication	76
4.7	Analysis of TAALES Academic Lists Amongst EAP and ESP Contexts for Lexical Sophistication	77

List of Figures

4.1	Lexical Diversity (TTR of 0.5)	37
4.2	Lexical Diversity (TTR of 0.6 – 0.65)	38
4.3	Lexical Diversity (TTR of 0.66-0.69)	39
4.4	Lexical Diversity (TTR of 0.7)	41
4.5	Lexical Diversity (TTR of 0.71-0.72)	42
4.6	Lexical Diversity (TTR of 0.73-0.74)	43
4.7	Lexical Diversity (TTR of 0.75)	45
4.8	Lexical Diversity (TTR of 0.76-0.77)	46
4.9	Lexical Diversity (TTR of 0.8-0.82)	47
4.10	Lexical Diversity (TTR of 0.83-0.88)	49
4.11	Lexical Diversity (TTR of 0.4-0.5)	52
4.12	Lexical Diversity (TTR of 0.6-0.64)	53
4.13	Lexical Diversity (TTR of 0.65-0.66)	55
4.14	Lexical Diversity (TTR of 0.67)	56
4.15	Lexical Diversity (TTR of 0.68-0.69)	57
4.16	Lexical Diversity (TTR of 0.7-0.71)	58
4.17	Lexical Diversity (TTR of 0.72)	59
4.18	Lexical Diversity (TTR of 0.75)	60
4.19	Figure (4.19): Lexical Diversity (TTR of 0.75)	61
4.20	Lexical Diversity (TTR of 0.76-0.77)	62
4.21	Lexical Diversity (TTR of 0.78-0.79)	63
4.22	Lexical Diversity (TTR of 0.8-0.81)	64
4.23	Lexical Diversity (TTR of 0.82-0.9)	66
4.24	EAP Interview Analysis - Age	79
4.25	EAP Interview Analysis – Study Duration	80
4.26	ESP Interview Analysis - Age	88
4.27	ESP Interview Analysis – Year	89

List of Abbreviations

EAP	English for Academic Purposes
EFL	English as a Foreign Language
ESP	English for Specific Purposes
LD	Lexical Density
TAALES	Tool for the Automatic Analysis of Lexical Sophistication
TTR	Type Token Ratio

Chapter One

Introduction

Chapter One

Introduction

This chapter introduces the current academic focus on the use of English as a Foreign Language learning addressed within the domain of the Kingdom of Bahrain.

1.1 Background Information

Amongst the various languages spoken in the Kingdom of Bahrain, the official language and lingua franca is Arabic. Following the Arabic language, English is also utilized by the majority of the inhabitants of the Kingdom as a second and foreign language. Other spoken languages include Farsi, Tagalog, Hindi, and Urdu (Crystal & Smith, 2023). In terms of its educational practices, public education first commenced in the year 1919, and was made available to both males and females. The levels taught include primary, intermediate, and secondary levels, with mandatory education being necessitated for all children between the ages of six up to fourteen. Higher education is also made available to all citizens, with various universities. Amongst the largest is the University of Bahrain, which is also the only national university (Crystal & Smith, 2023).

English language education in Bahrain is both fundamental and mandatory, with it being made available across all public and private schools on the island. The availability of access to the language is essential as the language is used across various domains, contexts, and for a variety of purposes, making its comprehension and usage fundamental to daily life in Bahrain. The English language is even considered a second language within all educational facilities in the country, and its status clearly expands to more than a mere second language.

English as a Foreign Language (EFL) is defined in terms of languages that are studied by learners where the native language spoken in their country is not English. As the main language used within contexts where EFL is taught is not English, educators must ensure that they develop and find models of English that are of use to their students (Gunderson, Di'Silva, & Odo, 2009). English is mandatorily taught commencing from the first grade to all inhabitants of the island. Various universities across the country offer English Mediated Instruction (EMI). For instance, various colleges within the University of Bahrain teach certain courses using English for Specific Purposes (ESP), including the Bahrain Teacher's College, and the colleges of Business Administration, Science, Engineering, and Information Technology. In addition, English is taught for academic purposes (EAP) in the English bachelor degrees in the Department of English Language and Literature of the College of Arts. Thus, the importance of the English language is highlighted not only in terms of teaching English for Academic Purposes (EAP), but also in terms of teaching in accordance to Specific Purposes (ESP), across colleges that initially teach a variety of subjects and majors. Thus, it is necessary to provide further insight into the aforementioned terms.

1.1.1 English for Academic Purposes Versus English for Specific Purposes

English for Academic Purposes, or for short, EAP, has been defined as a branch of ESP. According to Hyland (2006), EAP is mainly a form of language education that relates to the identification of particular features of language, articulatory and communicative skills, and the discourse-oriented practices of a target academic group. Thus, EAP is primarily designed to teach learners the English language, its science, and its mechanics. On the other hand, English for Specific Purposes, or ESP, can

be defined as a methodology applied to learning language that considers the specific requirements and needs of the language learner. ESP takes into consideration the learners' purpose in acquiring the foreign language, which means that the language is not taught for the sake of the language being learnt itself, but rather, is adjusted and adapted to the requirements of a specific field of study or domain. The fundamental difference between EAP and ESP, then, is that the former focuses on the description and explanation of the language, while the latter is based on orienting the language towards very specific fields and domains (Papadima, 2020).

The quality of text, or of speech, profoundly affects its reception and comprehension by an audience. There exists a variety of features of text that can influence its linguistic characteristics, with specific importance placed on the nature of the lexical items that occur within it. Thus, the term 'lexical complexity' can be posited, and is discussed in the following section.

1.1.2 Lexical Complexity

Lexical Complexity can be defined as the availability of a diverse and elaborate variety of lexical units that are both basic and advanced and readily accessible for use (Wolfe-Quintero, Inagaki & Kim, 1998). Ellis and Ellis, Rod, and Barkhuizen (2005) define complexity as the degree of complex and elaborated language utilized by a language learner. Lexical complexity itself can be defined as the variety of different vocabulary items or lexemes that are utilized in the speech or writing of a language user (Dewi, 2017).

Skehan (2009) relates that there is an association between the aspects of lexical complexity and grammar that result in the formation of both fluency and complexity. The outcome of this combination formulates 'sophisticated language'. In order for a text to be considered complex, a

single grammatical unit must be comprised of numerous distinct grammatical forms (Pallotti, 2009). According to Zeno, Ivens, Millard and Duvvuri (1995), when the vocabulary content within a text is observed, the complexity of the vocabulary is directly related to the general frequency of a lexical item's appearance in text. Nagy and Scott (2000) state that learners can elevate their ability to understand lexically complex texts when they are exposed to discourse that contains unique lexical units as they develop a significant comprehension of unfamiliar ideas and concepts. However, a text that is comprised of a high percentage of unique lexical items can become incomprehensible to readers due to the high level of precision associated with complex vocabulary (Hayes & Ahrens, 1988).

According to Bulte' and Housen (2014), lexical complexity holds considerable significance in second language learning, research, and teaching. It has been identified as an indicator, parameter, and diagnostic element of language progress, development and proficiency in both first and second languages (Bulte' & Housen, 2014).

It can be comprehended that different factors can affect the lexical complexity of a given individual's language. First, a language user's lexical complexity is subjected to change across a variety of contexts. For instance, the difference between lexical complexity in learners' L1 and L2 has been explored, as well as lexical complexity across spoken and written discourses. In addition, Lahmann, Steinkrauss, and Schmid (2015) state that education is another factor that can affect the lexical complexity of a language user and learner, with high levels of education resulting in the production of more complex language by speakers. Amongst the units that can be utilized to analyze and measure a language user's proficiency in language is their lexical complexity, which is specifically related to the use of varying vocabulary units in written or spoken discourse.

There are primarily two distinct manners of measuring lexical complexity identified as frequency and diversity. Diversity is the count of the total varying words that can be found within a text. The method applied to this calculation is the Type-Token Ratio, which both Skehan (2009) and Yu (2010) state can be calculated using the D formula. The alternative term, that being lexical frequency, measures the typologies of words that are observed when language is articulated, in comparison to the NES corpora's list of word frequencies, which is conducted to explore the level from which the word majority in either a written or spoken sample of language originates from (Skehan, 2009). The quality of text, or of speech, profoundly affects its reception and comprehension by an audience. There exists a variety of features of text that can influence its linguistic characteristics, with specific importance placed on the nature of the lexical items that occur within it.

Different measures of lexical complexity exist that can act as individual units that measure the whole complexity of a language user's lexeme choices in discourse. These measures are *lexical density*, *lexical diversity*, as well as *lexical sophistication*. The first measure to consider is lexical density. It has been defined by Halliday (1985) as "a measure of the density of information in any passage of text, according to how tightly the lexical items (content words) have been packed into the grammatical structure" (p.10). It can also be defined as the statistical analysis of language to assess the degree of lexical richness demonstrated within a text (Daller, van Hout & Treffers-Daller, 2003). Furthermore, lexical diversity is the variety of vocabulary and lexical items present within discourse (Malvern et al., 2004). The final measure, that being lexical sophistication, is defined as text comprised of content that is advanced, in which the

vocabulary is selective and does not occur frequently in language use (Tidball and Treffers-Daller, 2008).

The focus on all three areas of lexical complexity permits the researcher to reveal to which degree language learners within each individual context produce written discourse that incorporates a diverse range of vocabulary, the ratio of functional to lexical words, as well as the sophistication of the lexical words collected. As a result, it can be determined whether EAP or ESP contexts enable learners to utilize language with a higher degree of effectiveness and proficiency.

Specialization may also influence lexical complexity when considered in terms of the language used within academic contexts across EAP and ESP contexts. The term specialization in this context primarily refers to the field of English for Specific Purposes, which relates that different areas of specialization and knowledge require speakers to utilize language and jargon specific to the field of study. As a result, the lexical complexity across EAP and ESP contexts can differ profoundly on the basis of profession and field of knowledge.

1.2 Research Problem

Lexical Complexity is a feature that can be considered in language production by different people across varying contexts and backgrounds. It reflects the quality of the vocabulary and lexical units utilized in discourse production, which is generally measured in terms of *lexical density*, *lexical sophistication*, and *lexical diversity*. The degree of complexity of language is also measured in accordance with the number of unique and diverse words encountered in a single instance of written discourse, the repetition of words, and the overall depth of meaning as assessed through the percentage of functional and lexical words included.

When considering the contexts of EAP and ESP, it can be understood that certain fields of learning English as a Foreign Language (EFL) require learners to appropriately utilize specific lexical items and jargon to appropriately communicate the meaning associated with the specific fields being studied. The use of specific vocabulary and jargon is especially fundamental in the field of English for Specific Purposes, which is true for the field of medicine, business, architecture, biology, chemistry, etc. Within these disciplines, learners need to appropriately utilize language in a manner that is complex, diverse, and sophisticated enough to communicate meaning with proficiency and effectivity. Thus, it is possible that lexical complexity is directly affected by whether EFL learners are exposed to EAP or ESP English in their education. It can also be argued that since EAP and ESP contexts focus on language learning in different ways as per the ultimate purpose of each course, the ultimate degree of proficiency and hence lexical complexity acquired by learners can differ. Hence, it is important to investigate the degree of effectiveness observed in relation to each context and its effect on the lexical ability of its learners.

1.3 Research Questions

The study aims to answer the following main questions:

1. Are there any significant differences in the lexical complexity (density, diversity, and sophistication) between EAP and ESP university learners?
2. What are the perceptions of EAP and ESP university learners on their levels of lexical complexity?

1.4 Purpose of the Research

The primary aim of this study is to investigate the differences in the lexical complexity amongst university learners studying in EAP and ESP contexts. Through the research, the following objectives are addressed:

1. To explore whether EAP and ESP contexts influence the range of vocabulary used by the learners.
2. To explore whether EAP and ESP contexts influence the complexity of the vocabulary used by the learners.
3. To explore whether EAP and ESP contexts influence the number of lexical words versus functional words used by the learners.
4. To explore the opinions of learners on the effectiveness of EAP versus ESP contexts in language learning and language proficiency.

1.5 Methodology

To answer the aforementioned questions, the thesis used both quantitative and qualitative data collection methods. Specific measures of lexical complexity were addressed within the thesis to examine learners' lexical density, lexical sophistication, and lexical diversity. For each measure, specific formula and specific data analysis tools were applied to analyze the results obtained from the learners, which are detailed in Chapter Three. The results were analyzed using the SPSS software to obtain the level of significance and the independent sample T-test results.

In terms of the participants involved, samples for the research were collected from EFL learners from the Business and Arts Colleges. To obtain the qualitative data, ESP and EAP learners from Bahraini universities were also interviewed and their responses reported and analyzed.

1.6 Significance of the Research

Through this current study, the degree of effect of EAP and ESP learners' language proficiency is assessed for the first time in the local context of Bahrain Higher Education. The specific focus of proficiency of the study is EFL university students' lexical complexity, providing an insight into the inherent meaningfulness and effectivity of the written discourse produced by learners across both contexts. What adds to the significance of the study lies primarily in its methodology. While most studies have focused on the effects of either of the variables (EAP or ESP) (*see Chapter Two*) or on one of the measures of lexical complexity, namely those being lexical sophistication, lexical diversity, and lexical density, the existing study compares the effects of both EAP and ESP using all the three different areas of lexical complexity. This adds to the validity of the investigatory tools and to the reliability of their findings.

Through the outcome of the research, the significant contextual factors required to enable learners to appropriately and effectively utilize language are addressed, which can provide further opportunity for pedagogical development and enhancement of the educational experience for learners studying the English language.

Chapter Two

Literature Review

Chapter Two

Literature Review

The literature review section mainly focuses on providing a detailed insight into key research conducted on the three measures of lexical complexity across different contexts. The research addressed investigates how EAP and ESP contexts affect the overall lexical complexity and language proficiency of the learners involved in the research studies. The research has been divided according to the various measures of lexical complexity.

2.1 Conceptual Framework

2.1.1 Theoretical Background

2.1.1.1 English for Academic Purposes Versus English for Specific Purposes

2.1.1.1.1 English for Academic Purposes (EAP)

English for Academic Purposes is a branch of English language teaching and learning developed from English for Specific Purposes. Hyland and Hamp-Lyons (2011) define EAP as the description of English in sociolinguistic, psycholinguistic, and general linguistic scopes, within the contexts of scholarly exchange and academic study. According to Hyland (2006), EAP is mainly a manner of language education that relates to the identification of particular features of language, the articulatory and communicative skills, as well as the discourse-oriented practices of a target academic group. Here, the learners' expertise and subject-matter needs are taken into consideration (Hyland, 2006). The main goal of this form of English is to enable and facilitate the learning and teaching of the language, and well as its use in research by learners (Hyland & Hamp-Lyons, 2002). Through EAP, pedagogies are developed that are research-based, enabling

the assistance of learners in their research and learning of the English language (Hyland, 2006).

The focus of EAP is directed onto academic contexts, differentiating it from ESP. However, it can be noted that in English language teaching and applied linguistics, EAP is considered a sub-discipline of ESP. EAP itself is an extensive pragmatic discipline. It covers a variety of topics and fields of study, including applied linguistics, general linguistics, and educational matters, all of which can be taken into consideration through EAP (Hamp-Lyons, 2011). Starfield (2011) relates that texts provide the primary focus within this field, rendering discourse explicit to learners through the investigation of the structure of both spoken and written varieties of discourse.

EAP can also be considered as a form teaching the English language that is specialized (Hyland, 2006). According to Hyland (2006), EAP teaching is directed to research and language instruction that emphasizes the communicative requirements and practices of certain groups within academic or educational contexts. According to Rogers (2010), there is a significant difference between both General English and EAP. The latter is oriented towards specific goals and this discipline can highly affect the methods adopted in teaching as well as materials selected for teaching purposes. It is also important to relate that the student body involved in studying EAP are generally learners above the age of eighteen, meaning that this discipline is not directed towards children. The primary goal associated with EAP is to enable its learners to communicate in an effective manner, as well as succeed academically (Rogers, 2010).

The focus within EAP classrooms also differs profoundly. While General English may attend to teaching vocabulary and grammar, EAP classrooms are focused on discourse, register, and genre. Thus, it can be

noted that English for Academic Purposes is geared towards specific learners and developing their competencies (Rogers, 2010). In terms of Disciplinary Variation, it can be noted that for EAP the most significant factor to consider is a learner's knowledge acquisition.

Hyland (2006) also recognizes the necessity of adequately preparing learners and training them to engage with information in innovated ways. It is frequently observed that higher academic levels encourage student exposure to a variety of diverse genres. It can be noted that these varying disciplines are represented differently and reveal new ways of knowledge construction, teaching and learning. These factors can play a significant role in influencing student comprehension in relation to each discipline and can also encourage different methods of learning. In terms of EAP, it can be comprehended that the challenges encountered by learners are not limited to their linguistic abilities, but rather, are associated with their ability to engage with new literacies and communicative practices. Academic communication is thus, a 'social practice', which demands that the features of discourse, linguistic preferences, and communicative practices utilized across varying academic contexts must be analyzed in order to be taught to learners.

2.1.1.1.2 English for Specific Purposes (ESP)

According to Douglas (2000), ESP possesses three primary characteristics. The first is the existence of authentic material, followed by 'purpose related orientation' and finally, 'self-direction'. Purpose-related orientation refers to the teaching methods utilized, which require the administration of real-life situations or communicative tasks. Learners are thus enabled to act in situations that they are likely to encounter in actual life, in which they must utilize specialized English to communicate and accomplish a given task. In terms of self-direction, this is the quality of the ESP being used. Here, the

learners become users through self-regulated learning. The goal is to enable this form of learning to become prominent as part of the teaching strategies implemented in ESP education (Douglas, 2000).

Additional characteristics, defined as Variable characteristics, include the following:

- a. ESP caters to and is designed to accommodate certain disciplines.
- b. The methods used in teaching ESP, within certain academic contexts, can differ from those utilized to teach General English.
- c. Although ESP can be directed towards secondary school level learners, it is primarily focused on teaching adult or tertiary level learners, including learners within professional situations.
- d. Learners who are exposed to ESP are often intermediate or advanced in their proficiency level, rendering the course to be designed specifically for them.
- e. A degree of basic comprehension of the language system is assumed for the majority of ESP courses.

Fields that embrace the use of ESP include those of engineering, medicine, business management, computer science, tourism, and law. Each of the fields indicated, in addition to other professions, require a specialized use of English for effective communication to occur within professional contexts (Donesch-Jezo, 2012).

According to Donesch-Jezo (2012), one of the most significant and important branches of English Language Teaching and English as a Foreign Language is English for Specific Purposes. The development has occurred as the current professional and scientific expansions globally necessitate that speakers possess an extensive awareness of linguistic and vocabulary

structures observed in the English language to effectively communicate in the jargon related to specific fields of profession, which is further necessitated by the development of the English language into a prominent discourse utilized in communication and the conveyance of knowledge.

The growth of ESP is driven by global demands, learner-oriented focus, and linguistic revolution. Following the end of the war, global developments occurred at a rapid pace, with growths in international trade, digital technology, multinational companies, as well as satellite communication. The developments and growth rendered it necessary for learners and speakers to gain access to a lingua franca that could accommodate business, academics and trade (Nodoushan & Ali, 2020). Further significance can be placed on the development of Communicative Language Teaching and the Notional Functional teaching approach, which were feasible in teaching the English Language to learners who were not in TESOL-oriented education (Nodoushan & Ali, 2020).

English for Specific Purposes (ESP) is considered a subfield of applied linguistics (Nodoushan & Ali, 2020). It can be defined as a methodology applied to language learning that considers the specific requirements and needs of the learner. The methodology takes into consideration the factors that necessitated the learners' purpose in acquiring the foreign language. A fundamental variation between ESP and General English is primarily noted in terms of the way teaching and learning is practised, rather than theorized. However, significant changes have occurred in how language is taught across the two variants. ESP has influenced changes in the approaches adopted in teaching General English, which now highlights the significance of analysing student needs and purposes when learning language. Some of the most prominent difference can also lie in the fact that ESP is mainly taught to enable English language

users to utilize English effectively to accomplish tasks in their professional careers. On the other hand, General English can often include the teaching of grammatical structures and vocabulary that are generalized to the language itself rather than to a specific field of profession.

2.1.1.2 Measures of Lexical Complexity

2.2.1.2.1 Lexical Density

The term lexical density, also referred to as lexical richness, was defined by Halliday (1985) as “a measure of the density of information in any passage of text, according to how tightly the lexical items (content words) have been packed into the grammatical structure” (p.10). Halliday (1985) relates that the lower density of a text renders it simple and easier to comprehend by the reader. Density is considered a form of lexical complexity that revolves around the way in which words are developed (Halliday, 1985). It is also stated that the lexical items found are compared to the number of ‘running’ words in a text (Halliday, 1985).

The mentioned notion was first introduced by Ure (1971) and was later refined by Halliday (1985). The use of the term ‘lexical items’ accounts for the possibility of terms being adopted that are comprised of more than a single word such as with phrasal verbs (stand up, call off, etc.). Ure (1971) identified phrasal verbs as being two individual words, with one item being lexical, the other a preposition. Halliday (1985), however, identifies the entire phrasal verb as a single lexical unit. The development of lexical density can be observed in the sentences and analyses provided below.

1. The Trust has offered advice to local government authorities on cemetery conservation (Halliday, 1985, p.61).

Halliday (1985) analysed the sentence above as being comprised of four grammatical units and eight lexical units. Thus, a total of eight lexical items can be found amongst a collective total of twelve items. The calculation of the ratio in accordance to Ure's (1971) analysis would relate that the lexical density of the text is at sixty-seven percent. The developments contributed by Halliday (1985), however, alter the analysis. An alternate formula was developed that measured lexical density on the basis of clauses within a text. He states that the collective number of lexical items in ratio to the number of clauses within a text equate its lexical density. The following text can be utilized to clarify the idea mentioned:

2. The basic 'stuff' of living organisms is protoplasm/. There is no set composition of this /and it varies between one individual and the next/ (Halliday, 1985, p.67).

There exists a total of nine lexical items within the sentence above, as well as three individual clauses. As a result, the lexical density is calculated as 3.0 due to the presence of a ratio of nine out of three.

The measurement of lexical density necessitates the differentiation between lexical and grammatical units. The latter is comprised of items such as pronouns, conjunctions, articles, prepositions, as well as certain classifications of adverbs and finite verbs (Cindy & James, 2007). These items are presented in a closed system.

According to Thornbury and Slade (2006), the calculation of the ratio of function words to content words within a text is the text's lexical density. However, it has also been argued that lexical density refers to the number of content words such as adverbs, verbs, adjectives, and nouns, proportionate to the collective number of words found in a given text (Johansson, 2008). The degree of information presented in a text is

thus higher, when a text is comprised of a greater proportion of lexical words in comparison to function words. When a text is dense, it is comprised of more information and can possibly be challenging for readers to comprehend (Rahmansyah, 2012). However, when a text is identified as having a lower lexical density, it means that the collective number of grammatical words exceeds the lexical.

Mazgutova and Kormos (2015) claim that the measures of lexical density, sophistication and diversity are extremely significant in the assessment of an individual's English language proficiency. This is particularly so in relation to a language user's written language, which can also be determined with reference to the discourse features of cohesion and coherence, as well as text organization.

A text that is dense uses a variety of lexical words that correlate to the cumulative number of words in general. Johansson (2008), however, argues that the measure noted can lead to complications as, at times, in order to create a single lexical word, numerous function words must be reused. In these situations, the value of a TTR can be observed to be lower for a text that is relatively long rather than concise. When lexical density is considered, lexical items are measured in terms of their proportion in text (Johansson, 2008). Lexical density refers to the number of adjectives, nouns, verbs, and adverbs that can be found within the text.

2.2.1.2.2 Lexical Diversity

Lexical diversity directly relates to how diversified and varied a text is. Unlike density, it refers to the variation of word types utilized in a text. A given text can possess a high degree of lexical diversity, yet possess low lexical density (Johansson, 2008). The difference between

lexical density and diversity is because a text can be found to be comprised of a profound percentage of auxiliary words and pronouns, yet a small percentage of lexical verbs or nouns, rendering it dense, but not diverse.

According to Read (2000), terms such as 'flexibility' and 'vocabulary richness', 'lexical range, balance and verbal creativity', show that lexical diversity is related to the manner in which vocabulary is utilized as well as its size. Duran, Malvern and Richards (2004) state that the diversity of a text is related to the variety of vocabulary and lexical items present within it. If a given text is comprised of 'T+t' variation of lexical items, it is considered to possess a higher degree of diversity in comparison to one with simply 't' type vocabulary. Diversity is not, however, simply limited to the range of vocabulary used.

A simple and direct method utilized in measuring lexical diversity is to count the Number of Different Words (NDW), within a text (Duran, Malvern, Richards, & Chipere, 2004). A factor to consider in this count is related to the total number of words from which the text is comprised. This form of counting, however, is not always reliable as doubts can occur based on the size of a text. If a situation occurs where distinct transcripts vary on the collective words (tokens) used, complications can occur. For instance, a text of one hundred tokens can contain thirty-five types of words, while a text of fifty tokens can contain twenty-five different types of words.

In a case such as the aforementioned, the first text will be considered to be more diverse (Richards & Chipere, 2004). Although some researchers like Daller, van Hour and Treffers-Daller (2003) consider lexical diversity to be equivalent to lexical richness, others like Duran et al., (2004) simply consider it to only be a segment of lexical richness. A diverse text that has a high lexical density is one that is comprised of a richly varied vocabulary.

Speakers and writers are required to practise the use of a variety of words and limit their repetition of terms to produce a text that is highly lexically diverse (Johansson, 2008).

2.2.1.2.3 Lexical Sophistication

Tidball and Treffers-Daller (2008) define lexical sophistication as text comprised of content that is advanced, in which the vocabulary is selective and does not occur frequently in language use. Laufer and Nation (1995) relate that lexical sophistication is the percentage of lexical units in a text that are advanced or sophisticated. Lexical sophistication can also be defined as using terms in language that are rather technical, namely jargon, in addition to words that are uncommon and convey meaning in a manner that is concise, specific, and sophisticated.

According to Laufer and Nation (1995) words are subjected to division into varying categories formulated by taking into consideration frequency bands, which are acquired through the analysis of corpora on written language. Through a program known as Vocabfile, categorization is conducted, with frequency bands created denoting initially one thousand of the most frequently utilized word families. The initial list is augmented by the addition of a further one thousand most frequent word families. Finally, five hundred and seventy of the most frequently utilized word families common to academic texts, known as the Academic Wordlist (Laufer & Nation, 1995), are categorized. It has also been noted that through the use of the Lexical Frequency Profile, the proficiency of language users can also be assessed.

Further levels of sophistication are noted in how cognate a word is in its status (Van Hell and De Groot, 1998). A cognate word is one that exists in translation pairs, each item in the pair similar in sound and spelling. It

has been related that an item that is cognate is easier to process in comparison to one that is non-cognate, thus rendering it more sophisticated lexically (Van Hell and De Groot, 1998).

Duran et. al., (2004) relate that sophisticated use of the lexicon means that vocabulary items found in everyday language occur in low frequencies. Furthermore, in the identification of sophisticated vocabulary, the sophistication and difficulty of any given lexical unit is its frequency of repetition. Cobb and Horst (2004) also state that, when learning a language, the order of vocabulary acquired parallels the frequency of the word's occurrence, which is supported with research tools such as Laufer and Nation's (1995) Lexical Frequency Profile.

The profile operates on the notion that infrequently used vocabulary is far more difficult in comparison to lexical items that are in frequent use. Kuperman, Stadthagen-Gonzales, and Brysbaert (2012) relate that rapid processing of high-frequency words is commonly observed in comparison to low-frequency words, which are processed more gradually. According to Crossley, Cobb, and McNamara (2013), frequently utilized words in natural language are often learnt and utilized at a higher frequency than vocabulary that is articulated less often. In reference to language learning, the frequent observation of low-frequency words in a learners' written language often results in lowered proficiency scores, indicating that the learner's language quality is low (Crossley et al., 2013).

Duran et al., (2004) observes that lexical sophistication and diversity are connected. The connection between sophistication and diversity is due to the fact that when language is used extensively in any given text or speech, it is expected that the use of many different words will directly result in an increase in the diversity of the language. As a result of the

increase, the rarity of the word types utilized will also increase, rendering a text more sophisticated (Duran et al., 2004).

2.1.2 Previous Studies

2.1.2.1 Overall Lexical Complexity

A number of researchers have investigated the effects of EAP or ESP courses on the lexical complexity of learners across university contexts. Much of this research has been primarily focused on either ESP or EAP contexts rather than on a comparison between the two. The section below reflects on this research.

Dewi (2017) conducted research utilizing a quantitative corpus-based analysis to investigate the lexical complexity found in the written discourse of undergraduate learners utilizing the Web-based Lexical Complexity Analyzer. A collective total of one hundred and thirty student research projects were assessed for lexical density, complexity, variation, and sophistication. The results revealed that lexical density was calculated at 0.53. The aforementioned value was compared to several Chinese oral narratives collected for verification and contrast of the results, indicating that the written discourse yielded a higher lexical density. In terms of lexical sophistication and variation, it was noted that a higher degree of sophistication was observed in the written discourse when compared to the oral discourse of the Chinese EFL learners.

Thawarom and Singhasiri (2020) conducted a thesis to analyse the vocabulary choices and quality in the oral discourse of first-year undergraduate university learners studying science and technology. As assessed from the scores of the Ordinary National Education Test, the learners' English language proficiency was measured at an intermediate

level. The learners majored in various subjects which included Mechanical Engineering, Microbiology, Media Technology, and Media Arts. The context of data collection was an ESP classroom from three different classes focused on Technical English. The data collection and analysis were accomplished through the measurement of the learners' lexical richness, with focus placed on lexical sophistication, lexical diversity, and lexical density. The samples were collected while learners carried out a speaking task, and a total of one hundred learners were included in the research.

Three computer software programs were then utilized to analyse the data collected, which were TagAnt, D_tools, and AntWordProfiler. The outcome of the research indicated that the learners' discourse corresponded with the contents of the GSL1000 word list. Amongst the findings, it was noted that technical vocabulary was not frequently utilized by the learners. Further findings revealed that a lexical diversity of D58.8 existed amongst the learners, which is considered relatively rich. However, it was noted that repetition of words was fairly common in the speech produced by the learners. The density was calculated at 43.72%. It was also noted that nouns (20.18%) were amongst the most frequently utilized word categories, followed by verbs (13.87%), adverbs (5.52%), and adjectives (4.14%). The findings also concluded that nouns were the densest word categories utilized by the learners. The research revealed that the learners were not familiarized enough with the topic they spoke about, which can have a significant influence on lexical diversity, as per the observations of Yu (2009) and Skehan (2009). The researchers also concluded that productive vocabulary needed to be rendered more prevalent in the ESP setting to ensure that learners possess accurate and effective knowledge of their academic discipline.

Further research was conducted by Clavel-Arroitia and Pennock-Speck (2020) utilizing a Corpus Linguistic analysis in order to conduct a comparison between the lexical diversity, sophistication and density in written and spoken interactions between a total of fifty Spanish and Japanese university learners in a telecollaborative exchange. The learners were selected from the Universitat de València and Kwansei Gakuin Universities. It is also notable that the Spanish learners were EAP learners studying an English Studies course, while the Japanese learners were ESP learners enrolled in a Business Studies English course. The proficiency level to be attained at the end of the courses also differed as the Spanish learners aimed to attain the C2 level, the Japanese learners the B2 level. Three different programs were utilized for the analysis of lexical complexity which were TAALED 1.4.1, TAALES 2.2., and Text Analyzer.

The outcome of the thesis indicated that the Spanish learners utilized more complex language in oral discourse, as per the statistics found for their spoken discourse across the majority of lexical complexity indices. It was noted that these learners produced longer and inherently fewer sentences, suggesting that syntactic complexity and appropriate grammar were prevalent in discourse. In terms of lexical density, this measure was the only category of lexical complexity in which Japanese learners were noted to have performed better. In terms of written discourse, the results showed that both groups performed approximately the same across both contexts. Fundamental differences were noted in terms of spoken discourse, however. The differences were judged to be the result of the constricted timeframe permitted for spoken discourse in comparison to that for written discourse. In terms of the Spanish learners, it was also noted that the number of words in their written discourse exceeded the number of words produced in their oral discourse.

Rahayu, Utomo, and Setyowati (2021) carried out a thesis to explore the syntactic and lexical complexity noted between the written discourse of L1 and L2 speakers of English. A quantitative approach was applied to this research utilizing the Web-Based Complexity Analyser and the Web-Based Syntactic Complexity Analyser. A total of sixty-seven undergraduate learners from the English Literature study program were selected for the research, thus, the focus was placed on EAP learners of English. The learners were in their second semester at the University of Mulawarman. A collective total of one hundred and thirty-four essays were considered for the research. The languages in which the texts were written were both English and Indonesian, and the lexical density, variation, and sophistication of the texts were taken into consideration and calculated. The findings were tabulated by the researchers using the Paired Sample *T*-Test.

The findings indicated that lexical complexity was much higher in the English language (L2), in comparison to Indonesian (L1). The variation between the lexical density of the written texts across the two languages was 0.0761. In terms of lexical sophistication and variation, the values obtained were recorded as less than 0.0. Thus, the outcome indicated that higher lexical complexity was noted in English language production in comparison to Indonesian. The researchers concluded that, in their native Indonesian, the learners were prone to utilizing redundant sentences that were largely indirect. Function words were also repeated much more frequently. The researchers concluded that the production of the English essays could have been aided by the use of dictionaries, resulting in more complex language and less repetition of words.

2.1.2.2 Lexical Density

Of the three measures of lexical complexity, lexical density can be considered the most widely researched. Zareva (2019) conducted a thesis

to measure the lexical density of three varying groups of university learners that ranged from native speakers of the English language, ESL language users, and users of English as a lingua franca. The purpose of the research was to investigate how the lexical complexity of the learners' choices varied. The values calculated for the lexical density the respective groups were noted as 0.49, 0.5, and 0.51. Thus, it was noted that the results observed for members of each group were identical without any significant differences noted.

Further research was conducted by Thawarom and Singhasiri (2020), which focused on the lexical density of the vocabulary utilized within an ESP classroom. The participants were one hundred learners studying in their first year at university. Data was collected during a speaking task within the university setting. For the purpose of the research, a scale-based approach was utilized in technical identification and lexical density values were calculated using the TagAnt program. The results primarily indicated that the overall lexical density within the learners' speech was a relatively low 43.72%, indicating that adequate lexical density was not observed, and thus, language was less lexically complex. The outcome reflects that ESP contexts could provide limited exposure to English language education, which, as a result, limits the vocabulary exposure and production of learners.

Further research on lexical density was conducted by To, Fan and Thomas (2013), with focus placed on texts within English textbooks. Four varying levels of the textbooks were selected, these being elementary, pre-intermediate, intermediate, and upper-intermediate. The lexical density was calculated through the adoption of three separate measures, namely, Flesch's formula for Reading Ease (1948), and formulae for calculating lexical density as proposed by Ure (1971) and Halliday (1985).

The results indicated that three of the four texts were high in their lexically density, with the lowest lexical density noted for the upper-intermediate level document. The most lexically dense text was also identified as the pre-intermediate text.

Amelia, Silvana, and Sinar (2020) also conducted research to measure the lexical complexity of the language utilized by debaters within speeches made in the final round of the 2018 National University Debating Competition (NUDC). The speech analysis was conducted between the Opening Opposition team and the Opening Government team. Following collection, the data was transcribed into text and the lexical items and clauses were identified. The outcome of the research revealed that the Prime Minister's speech possessed a lexical density of 40.95%, while the Deputy Prime Minister's speech revealed a lexical density of 42.44%. The Leader of the Opposition, however, demonstrated a lexical density of 46.14%, while the Deputy Leader of the Opposition was 47.90%. These figures suggest that the debater's speeches contained a wider degree of information due to the presence of a greater number of lexical items over grammatical items.

The section that follows provides an in-depth insight into the varying social factors and methods of data analysis that were addressed and utilized within this thesis. The methodology section addresses the participants that took part within the research, as well as the forms of qualitative and quantitative data analysis utilized. Further information is provided on the formulae adopted to calculate measures of lexical sophistication, lexical diversity, and lexical density while details of both independent and dependent variables are also given.

In sum, much of the research conducted on lexical complexity has focused on the general richness of vocabulary demonstrated by native

speakers of English, as well as some ESL users of English. Much focus has been placed primarily on lexical density. However, research to date has not been concerned with comparing EAP and ESP learners. Thus, it is necessary for a comparative study to be conducted that highlights the difference between the various measures of lexical complexity of EFL learners, especially within the Arabic Gulf. Research on EAP versus ESP, with specific focus on their effects on lexical complexity, has not been conducted within this region prior to this study.

Chapter Three

Methodology

Chapter Three

Methodology

The thesis entails both quantitative and qualitative methods of data collection and analysis. For the purpose of the investigation, written samples were collected from learners for the analysis of lexical complexity. The lexical complexity demonstrated by learners in various EAP and ESP contexts across various universities in Bahrain were assessed and analyzed using three distinct measures on the basis of the sub-category of lexical complexity considered, namely, lexical density, lexical sophistication and lexical diversity.

3.1 Methodology

3.1.1 Participants

For the purpose of this research, written samples were collected from EAP and ESP learners in different universities in Bahrain, with each sampling containing a total of nine thousand five hundred words. The samples were collected from EFL learners from EAP and ESP contexts. The EAP learners were Humanities and Fine Arts students, while the ESP learners were university students within technical disciplines. The written samples collected from both sets of learners were then statistically analyzed to assess the degree of lexical complexity observed between the two language disciplines.

3.1.2 Quantitative Data

3.1.2.1 Lexical Density Statistical Calculation

The lexical density was calculated through a specific formula which required the total number of lexical words to be divided by the total number

of words in a select text. The value then obtained was multiplied by one hundred to attain the lexical density in percentage form for each fifty-word paragraph for both EAP and ESP learners. The formula is as follows, “The number of lexical items/The total number of words * 100”.

3.1.2.2 Lexical Diversity (Type Token Ratio)

Statistical Calculation

The lexical diversity was enumerated using the Type-Token Ratio (TTR), which measures the vocabulary range within the sample provided, which can then undergo analysis. Within a written text, the collective value of words that are distinct are referred to as the ‘Types’, while the total number of words are known as ‘Tokens’. The formula to calculate the TTR involves the total value of Types being divided by the total value of Tokens in a language segment (type-token ratio = (amount of types/number of tokens) * 100). The lexical density of the text is determined by the closeness of TTR ratio to the value of 1. Thus, the closer the ratio to 1, the higher the lexical density.

3.1.2.3 Lexical Sophistication Statistical

Calculations

A specific measure was applied to calculate the lexical sophistication, which was concluded through the division of the collective total of sophisticated word types by the collective value of general word types available within the sample collected. The analysis was conducted using the Tool for the Automatic Analysis of Lexical Sophistication (TAALES 2.0).

3.1.2.4 Qualitative Data

A qualitative approach was also adopted through interviews conducted with eight randomly selected EAP and ESP learners across universities in

Bahrain. The participants were required to share their perceptions and opinions on lexical complexity and its importance in English language production. They also communicated their views on how EAP and ESP courses and academic levels can influence the lexical complexity reflected in student language production. The interviews were conducted at each learners' respective university during their regular attendance hours.

3.1.2.5 Dependent Variable

The dependent variables considered in this research are various elements of lexical complexity. These include the lexical density, lexical diversity, and lexical sophistication of the discourse produced by the respondents. Each element will be explored with a distinct set of equations, as calculated by software designed to analyse the collective words in the samples collected from the EFL learners from both the ESP and EAP contexts.

3.1.2.6 Independent Variable

The independent variables within this thesis are the distinct EFL contexts in which the learners were being taught the English language, which specifically relates to whether the learners were educated in an EAP or ESP context during their university experience, as required by their major. The learners from the EAP contexts were Fine Arts and Humanities learners, while the ESP learners studied varying majors within Bahraini universities.

Chapter Four

Findings and Discussion

Chapter Four

Findings and Discussion

The findings section reports the outcome of detailed analysis of various measures of lexical complexity, as obtained from the analysis of written essays provided by university students from both EAP and ESP contexts. The analysis of the individual measures has been provided in a series of stages, commencing with a general report of both lexical density and type token ratio values obtained from the collective total number of words obtained from learners from each context (EAP/ESP). Following the report on general results obtained, the type token ratio calculated from the total number of words in each group was compared to the calculation of the overall lexical density as well. Following the aforementioned sub-section, the analysis of the results of an independent T-test were also provided, followed then by the results of a TAALES analysis of the texts to analyze lexical sophistication. Following the varying forms of quantitative analysis, the qualitative analysis of the interview results was also provided.

4.1 Lexical Diversity (TTR) Against Lexical Density

4.1.1 Lexical Diversity (TTR) Against Lexical Density of EAP Learners

In order to provide an in-depth analysis of the results observed amongst both EAP and ESP learners, the values obtained were arranged from the highest to the lowest TTR noted amongst both groups. The lexical density that corresponded to each TTR, per paragraph, has also been presented. It is necessary to note that a total of one hundred and sixty divisions were

created for each group, which have been presented individually in the form of histograms. Amongst the EAP learners, the lowest TTR noted was 0.5, while the highest was 0.9. It is also noted that the highest Lexical Density observed amongst the EAP respondents was 73%, while the lowest was 37%.

Furthermore, it is also crucial to note that the categories created within this section have been divided into ranges based on the Type Token Ratio calculated for each set of text divisions. The categories created have been divided into TTR ranges of 0.5, 0.6, 0.7, 0.8, and 0.9. It should be mentioned that the lexical densities were also observed in each category. Amongst the 0.5 range, the minimum lexical density calculated was 41%, while the highest was 44%. The next range to consider is the 0.6 range, within which the lowest lexical density observed was 58%, while the lowest was 38%. Moreover, within the 0.7 Type Token Ratio range, the highest lexical density observed was 65%, while the lowest was noted at 43%. Furthermore, the highest lexical density calculated within the 0.8 category was 73%, while the lowest was 37%.

4.1.1.1 Lexical Diversity (Type Token Ratio of 0.5)

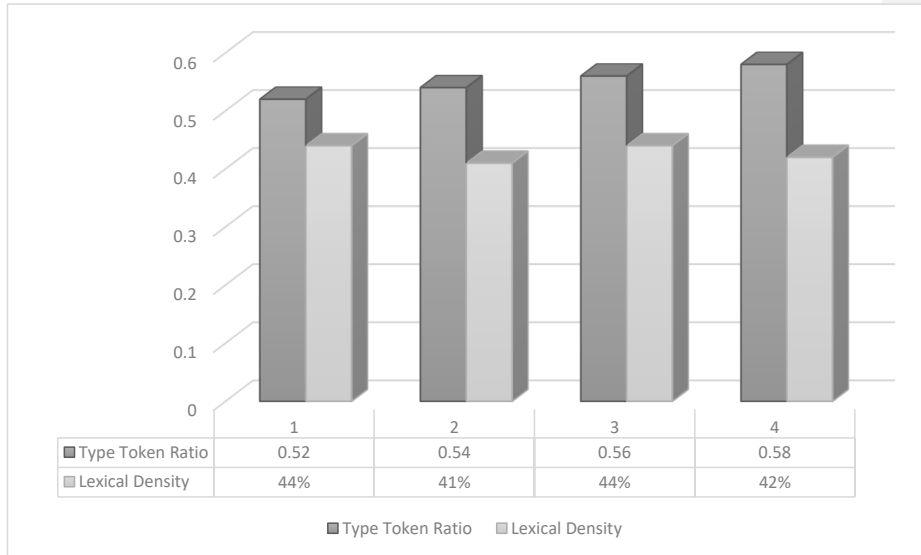


Figure (4.1): Lexical Diversity (TTR of 0.5)

The figure above demonstrates the lexical density and TTR calculated for the EAP learners amongst the EAP group. It can be observed the least obtained TTR amongst the EAP learners occurred in the 0.5 range. It can also be observed that the learners within this group demonstrated TTRs of 0.52, 0.54, 0.56, and 0.58. It is also imperative to report that the lexical density calculated within this group did not exceed the 40% range, with the minimum value occurring at 41%, and the highest being 44%. The values obtained indicate that the lexical density correlated with the TTR which, although it occurred in mid-range, means that a lower TTR resulted in a lower lexical density as well.

4.1.1.2 Lexical Diversity (Type Token Ratio of 0.6 – 0.65)

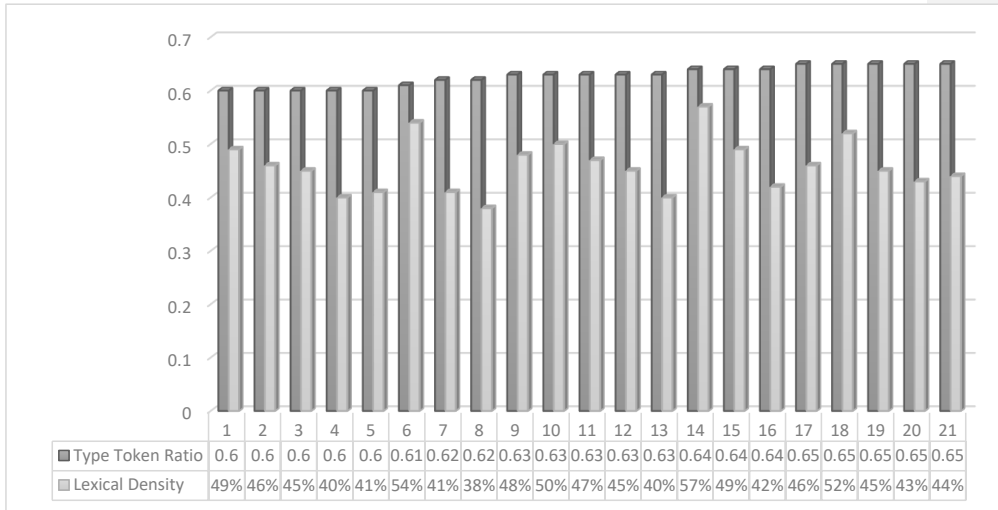


Figure (4.2): Lexical Diversity (TTR of 0.6 – 0.65)

As per the values collected, it can be noted that a total of 48 paragraphs were observed to occur within the 0.6 range. As a result, the values were divided into groups ranging from 0.6 to 0.65, and a second group ranging between 0.66 and 0.69. The graph provided above relates the results of the former. It can be noted that the values within this group include five TTR occurrences of two individual values, those being 0.6, 0.63, and 0.65. The remainder of the values demonstrates three instances of 0.63, and two of 0.62. It is important to observe that although the TTR has now increased and neared the value of one, the majority of the values in terms of lexical density remain within the same value range that was observed within the 0.5 range, that being 40%. Amongst the results noted for the EAP learners who reside within the 0.6 range, the lowest lexical density observed is 40%, while the highest is 49%.

However, the lowest lexical density found in general within this group is 38%, which was a value even lower than the collective variables calculated in the 0.5 TTR range. Furthermore, the highest lexical density observed within this group was 57%. A total of four values occurred in the 50% range in terms of lexical density, which included LDs of 57%, 52%, 54%, and 50%. The values demonstrate that although the lexical density did not increase significantly, its development has commenced, and a higher range of density is being noted amongst the learners. However, it is also important to mention that significant correlation has not been noted in reference to the increase of TTR and lexical density.

Thus, an increase in the Type Token Ratio does not necessarily determine a correlating increase in lexical density as well. As observed, the highest TTR value of 0.65 did not mean that the highest lexical density was observed within the same group, which is evident when regarding the lexical density value of 57%, the highest amongst this value range, which was noted to correspondent to a TTR value of 0.64. Similarly, the lowest lexical density of 38% was observed against a TTR of 0.62 rather than 0.6, demonstrating that a high TTR value does not necessarily indicate a high lexical density, and vice versa.

4.1.1.3 Lexical Diversity (Type Token Ratio of 0.66 – 0.69)

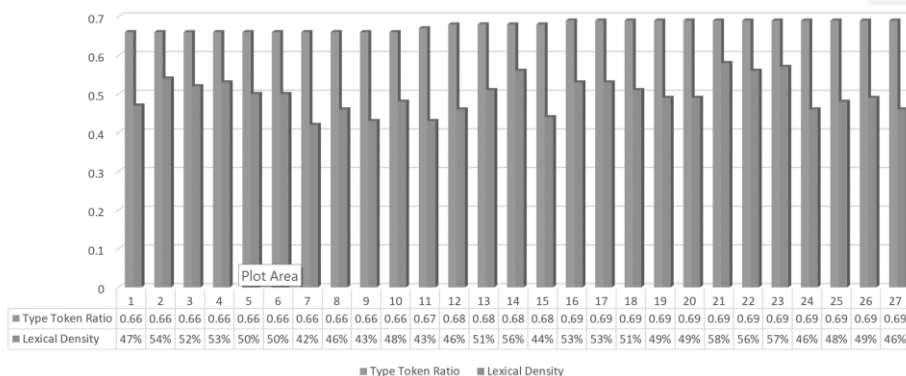


Figure (4.3): Lexical Diversity (TTR of 0.66-0.69)

The figure above reveals the value pattern observed within the TTR range of 0.66 and 0.69. Based on the calculations performed and the results noted, it is observed that fourteen of the twenty-seven paragraphs yielded a lexical density that occurred within the range of 50%, while the remainder yielded lexical densities occurring only within the 40% range. It can be observed that a lexical density below 40% was not evident in this case, nor was there a lexical density higher than the 50% percentile. In terms of TTR, it was noted that 10 of the values within this section occurred at 0.66, while the lowest frequency was noted at 0.67. A total of four occurrences were noted for 0.68, while the remainder occurred at 0.69. It is also important to note that the majority of TTR values observed, that being 44%, were all within the 0.69 range amongst this group of TTR values.

In reference to lexical density, it was noted that the values calculated ranged differently across the 50% percentile, with the lowest being 50% itself, and the highest at 58%. However, when occurring in the 40% range, the lowest value obtained was 42%, the highest being 49%. It is also important to note that once again, the lexical density observed did not definitely and strictly correlate in increase or decrease with the Type Token Ratio. However, it cannot be ignored that the values obtained for the highest Type Token Ratio, as well as the highest lexical density both occurred in the same paragraph. The values were a TTR of 0.69 and a lexical density of 58. The same pattern occurred in reference to the lowest values noted, that being a TTR of 0.66 resulting in a lexical density of 42%.

The pattern can be seen to recur amongst the values within this specific group as all the instances of high lexical densities occur against a high TTR as well. For instance, the lexical densities of fifty-seven and fifty

six percent can only be noted against a TTR of 0.69. However, amongst the higher TTR values, it can also be observed that a lexical density of a mere 46% also occurs. Similarly, across a TTR of 0. Thus, although some paragraphs yield correlated values between the Type Token Ratio and lexical density, numerous values are also observed to occur in randomized patterns, once more indicating that the TTR and lexical density may not always correlate as values increase or decrease.

4.1.1.4 Lexical Diversity (Type Token Ratio of 0.7)

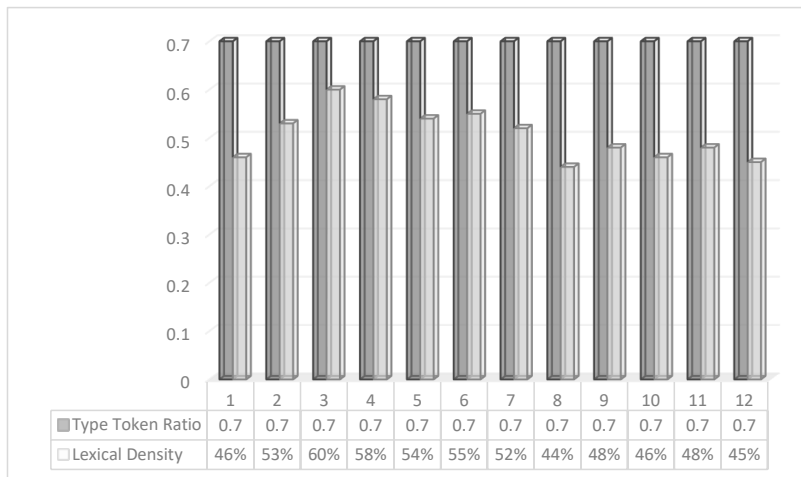


Figure (4.4): Lexical Diversity (TTR of 0.7)

The figure above demonstrates the values observed within the 0.7 range of the Type Token Ratio. As the graph reveals, a total of 12 sections were obtained within this range. It needs to be mentioned that although the TTR value has remained the same, a wide variety of changes were observed in reference to the lexical densities witnessed in this range. The lexical density has been noted to vary between the 40% range, up to 60%. The range is significant as it reveals that the Type Token Ratio can have limited effects on lexical density. Amongst the results, it can be noted that seven cases of

TTR all demonstrated a lexical density around 40%, with the highest value noted being 60%, which is the first instance of a lexical density of 60% being obtained amongst the EAP learner group.

Furthermore, the 50% range is comprised of five values, with the lower being 52%, and the highest reaching up to 58%. When noting the lowest value obtained, 44% was observed, which is higher than the lower lexical density observed within the 0.6 range at 38%. A 6% variation is demonstrated, which shows significant increase in minimum lexical density within the 0.7 range for TTR. Furthermore, the highest value observed in lexical density amongst the 40% percentile was 48%, which was common amongst two of the samples calculated.

4.1.1.5 Lexical Diversity (Type Token Ratio of 0.71 - 0.72)

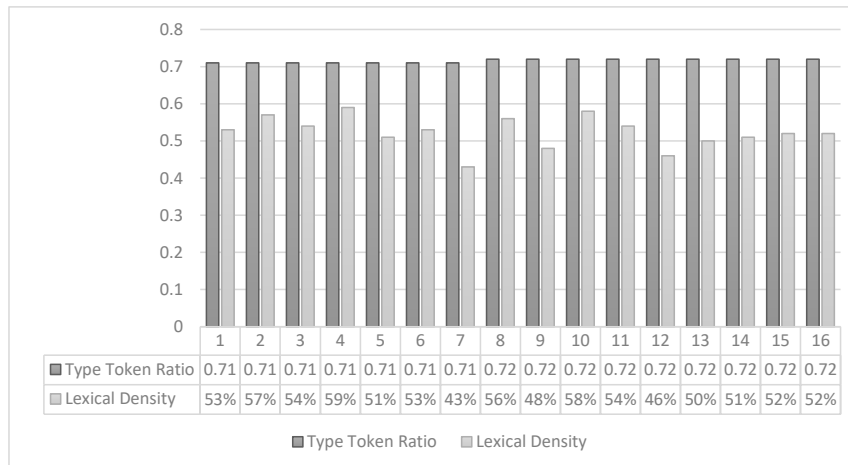


Figure (4.5): Lexical Diversity (TTR of 0.71-0.72)

The figure provided demonstrates the lexical density and Type Token Ratio noted amongst written samples collected from EAP learners, with the TTR values of 0.71 and 0.72. Seven of the samples collected all possessed a TTR

of 0.7, while the remaining nine samples represented a TTR of 0.72. A total of sixteen samples were noted within this category. The results demonstrate that the highest lexical density noted within this category is 59% of the overall text sample. The value mentioned is lower than the value noted amongst the 0.7 TTR range, but only varies by 1%. Further observation demonstrates that a total of thirteen samples possessed a lexical density within the 50% range. The development is significant as it demonstrates an increase in the collective lexical density value within this range. Furthermore, the remaining three samples within this group possessed lexical densities within the 40% range, with the highest value residing at 48%, and the lowest at 43%. The minimum value obtained demonstrated a mere 1% difference between the 0.7 range and the 0.71-0.72 range. It is also worth mentioning that, once more, lexical density is not determined by the overall TTR of the sample. For instance, the highest value obtained for the lexical density of the 0.71-0.72 range was 59%, which was noted amongst the 0.71 TTR values rather than the 0.72.

4.1.1.6 Lexical Diversity (Type Token Ratio of 0.73 – 0.74)

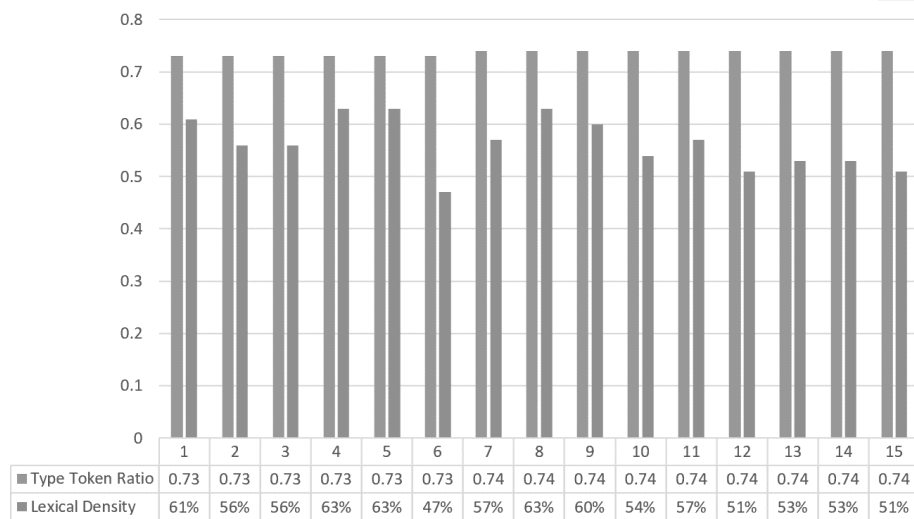


Figure (4.6): Lexical Diversity (TTR of 0.73-0.74)

The figure provided conveys the lexical densities and Type Token Ratio of samples within the 0.73 -0.74 TTR range. A total of fifteen samples were collected, amongst which six were of TTR 0.73, while the remaining nine were of 0.74 TTR. The samples demonstrated a significant increase in the total number of samples with lexical densities of 60%. Amongst these samples, a total of five reached or slightly exceeded 60%, with the highest lexical density among the group noted at 63%. The specific value was common amongst three of the samples, two of which possessed a TTR of 0.73, and one of 0.74. furthermore, another nine samples all had Type Token Ratios within 50%, of which the highest value calculated was 57%. The former percentage was noted amongst two samples from the 0.7 TTR, but none from 0.73.

The lowest in this range was 51%, which was common amongst two of the samples, both from a TTR value of 0.74. An additional two of the 50% were noted at 53% and were once more observed to correlate with a TTR of 0.73, while the two lexical densities occurring at 56% were found amongst the 0.73 TTR value. The values obtained once more make it clear that lexical density and TTR do not always increase or decrease in correlation, but rather, can vary in a randomized pattern. Finally, it can be mentioned that the minimum TTR value calculated amongst this specific group is the value of 47%, which is three percent higher than the 44% noted in the previous group range in terms of minimum value.

4.1.1.7 Lexical Diversity (Type Token Ratio of 0.75)

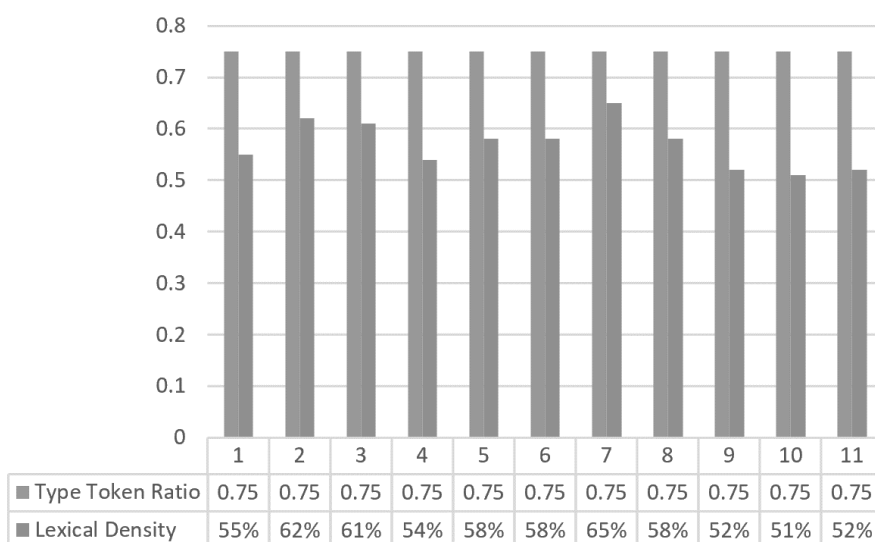


Figure (4.7): Lexical Diversity (TTR of 0.75)

The figure above represents the values obtained in the 0.75 range of TTR, as calculated from eleven written samples collected from English for Academic Purposes learners within different universities. The results obtained demonstrate a significant increase in the minimum value of the Type Token Ratio observed within the written samples. Here, eight of the eleven samples all occur within the 50% range in terms of their lexical density. The minimum lexical density noted was 51%, which was observed for a single sample. Two of the samples produced a lexical density of 52%, while fifty-four and fifty-five percent occurred just once. It was also observed that three of the eleven samples all produced a TTR of 58%. However, the highest TTR noted amongst the 0.75 group, as well as all prior groups ranging from 0.5 to 0.74, is 65%. Furthermore, a single occurrence of both sixty-two and sixty-one percentages can also be observed. In this specific TTR value category, it can be seen that the lexical

density calculated has increased significantly in comparison to prior sample categories.

4.1.1.8 Lexical Diversity (Type Token Ratio of 0.76 – 0.77)

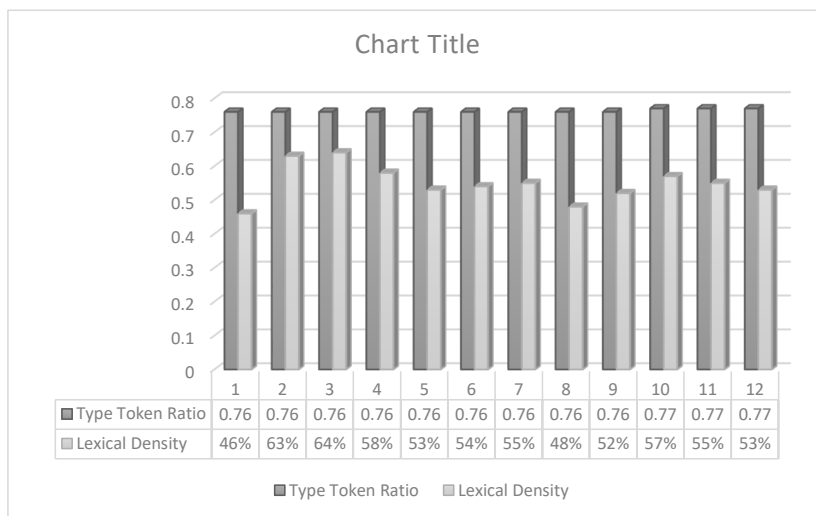


Figure (4.8): Lexical Diversity (TTR of 0.76-0.77)

The figure provided above demonstrates the TTR range of 0.76 and 0.77, as observed amongst a total of twelve of the samples analyzed from EAP learners. Amongst the samples collected, a total of nine out of the twelve possessed a Type Token Ratio of 0.76, while the remainder had a TTR value of 0.77. In terms of the lexical density calculated, the values indicated that the highest value obtained was 64% amongst the collective twelve figures. In fact, only two of the samples in this range reached a lexical density above 60%, with the second sample calculated at 63%. It should be mentioned that both of these samples belonged to the 0.76 range. The remainder of the values ranged between fifty-two and fifty-eight percent as far as the fiftieth percentile is concerned. Two of the samples occurred at

fifty-five and fifty-three percent, while only one sample for each of fifty-eight, fifty-seven, and fifty-two percent was noted. The lowest values belonged to the fortieth percentile, with the lowest of these being forty-six percent in terms of lexical density, the highest being forty-eight percent. The difference, however, between the highest and lowest values within this specific value range of TTR is 20% in terms of lexical density.

4.1.1.9 Lexical Diversity (Type Token Ratio of 0.8 – 0.82)

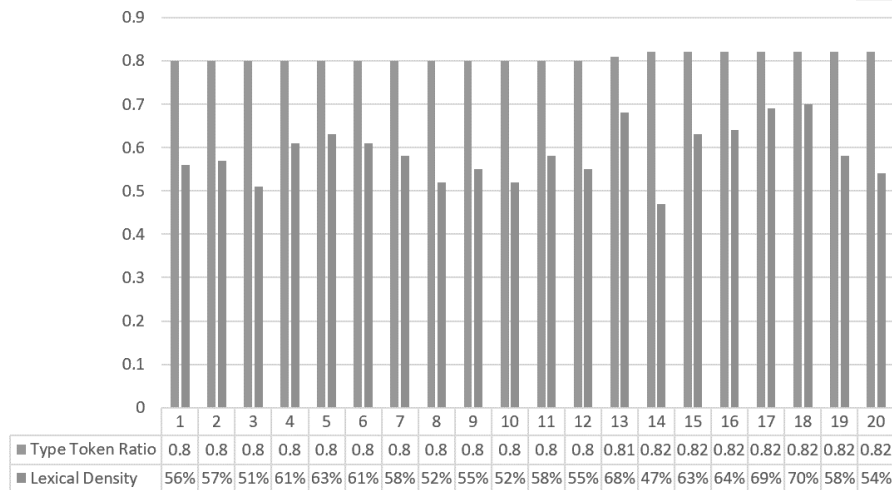


Figure (4.9): Lexical Diversity (TTR of 0.8-0.82)

The figure above represents the overall values observed from twenty samples that possessed a Type Token Ratio that ranged between 0.8 and 0.82. The values observed demonstrate that a total of twelve out of the twenty samples all obtained a TTR value of 0.8, while only one was calculated to possess a total TTR of 0.81. The remaining seven samples showed a value of 0.82 when their TTR values were calculated. Amongst these responses, the highest lexical density that was calculated was noted to be seventy percent. Amongst the TTR values calculated between the range of 0.5 to 0.7, the highest lexical density obtained can be found in the

0.8 range of Type Token Ratios. However, only a single lexical density has been noted at 70% in this value range.

The next highest lexical density is noted at sixty-nine percent, followed by sixty-eight percent. It can be important to note that a total of seven values amongst the total twenty were calculated as lexical densities within the 60% range. The most frequently observed lexical density was found amongst thirteen samples, all of which occurred within the fiftieth percentile. Amongst these values, the highest lexical density was fifty-eight percent, which was found in three samples. Only one sample had a lexical density of fifty-seven, with yet another single instance showing fifty-six percent. A further two samples occurred at fifty-two, while the lowest in the fiftieth percentile was at fifty-one percent, again for just one sample. Within these samples, the lowest value obtained in lexical density was forty-seven percent, with a Type Token Ratio of 0.82.

Furthermore, in terms of the correlation between TTR and lexical density, it can be noted that the highest lexical density (70%) within this range occurred at a TTR of 0.82. However, the lowest lexical density also corresponded to a Type Token Ratio of 0.82. The majority of values for lexical density that occurred against the 0.82 Type Token Ratio were mostly noted within the fifty, sixty, and seventy percent ranges, while the 0.8 TTR corresponded to lexical densities within the fifty and sixty ranges. Thus, a fundamental difference could not be observed between the lexical densities calculated for the samples provided at Type Token Ratio values between 0.8 and 0.82. However, a general steady increase in lexical density has been common as the TTR ranges increased from 0.5 to 0.8.

4.1.1.10 Lexical Diversity (Type Token Ratio of 0.83 – 0.88)

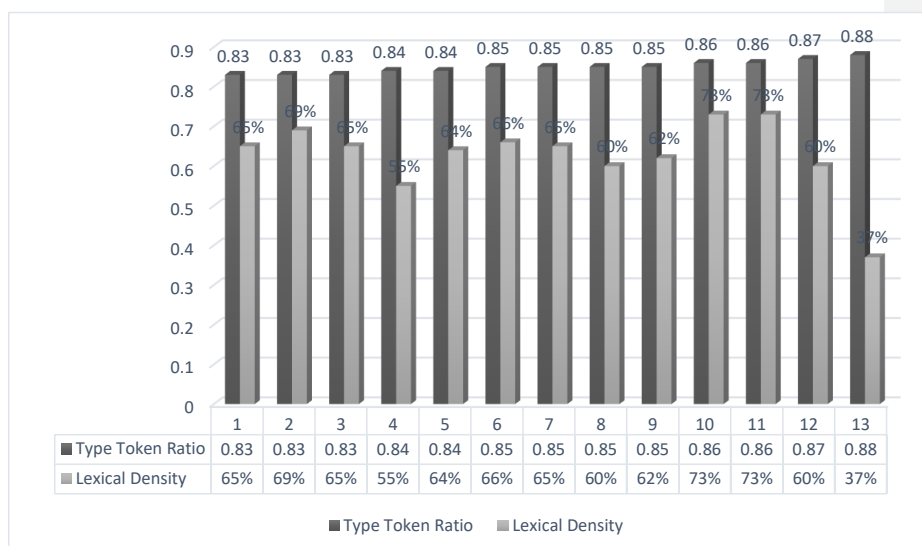


Figure (4.10): Lexical Diversity (TTR of 0.83-0.88)

The graph provided demonstrates the results obtained from an additional thirteen samples collected from Fine Arts and Humanities university students. The TTR range observed is the highest noted amongst the total number of samples that were collected. The graph above demonstrates the Type Token Ratio range between 0.83 to 0.88. It can be observed that three of the samples were calculated to possess the TTR value of 0.83, two of 0.84, four of 0.85, and two of 0.86. The least frequently observed TTRs in this range were those of 0.87 and 0.88. It can also be observed that the Type Token Ratio of 0.88 is the highest amongst the collective values calculated for EAP learners.

In terms of lexical density, the highest percentage obtained within the entire EAP learner samples was calculated at 73%, which occurred in two instances, both of which correlated against 0.86 Type Token Ratio. The

remaining results returned nine lexical density samples within the sixtieth percentile, with the highest value obtained being 69% in a single sample. The second highest value within this range occurred once at 66%, while three instances of 65% in lexical density were noted. There was only one instance of a lexical density of sixty-four and also one only of sixty-two percent, while the values of sixty percent occurred twice.

The two lowest values observed in this range of TTR were fifty-five percent and thirty-seven percent. It can be noted that the highest Type Token Ratio value obtained across the entire EAP samples also possessed the lowest range of lexical density, that being in the thirty percent range, at thirty-seven percent. It is also necessary to state that this is the lowest lexical density calculated amongst the entire samples analyzed from the EAP learners within Bahraini universities. The analysis provides the greatest evidence to demonstrate that the lexical density of a text does not increase with an increase in Type Token Ratio.

4.2 Type Token Ratio of ESP Learners Across Divisions

The analysis of the written samples collected from the English for Specific Purposes learners from Bahraini universities demonstrated a number of factors in reference to both lexical density and Type Token Ratio. First and foremost, the values obtained revealed that the lower value of TTR recorded amongst this group of language learners was 0.48. In contrast, the highest Type Token Ratio observed was 0.9, which was 0.42 values higher. Furthermore, only a single value was noted to occur in the 0.4 range in terms of TTR. However, there were a total of eight samples that resided in the value category of 0.5. Amongst these numbers, the highest lexical density calculated was forty-three percent, while the lowest was twenty-six

percent. The values in between the peaks all occurred within the thirtieth percentile, against the high, mid, and low levels.

Furthermore, the results also demonstrated that the lowest TTR in the 0.6 range was 0.6 itself, while the highest was 0.69. When lexical density was considered, the highest value obtained was sixty-three percent, while the lowest was thirty percent. The lowest value obtained for the 0.6 range in lexical density was also observed to be four points higher than the lowest lexical density within the 0.5 TTR range. The majority of the values in this group were also observed to occur within the thirtieth and fortieth percentile ranges, with thirty percent occurring more frequently.

In terms of the Type Token Ratio of 0.7, the lowest value was 0.7 itself, while the highest was 0.79. This range is the one with the most values noted amongst all the TTR results gathered from the ESP learners. When lexical density was considered, the highest value calculated was fifty-five percent. On the other hand, the lowest value obtained was thirty-four percent. The majority of the lexical density values calculated amongst this group was within the forty percent range. When lexical density is considered, the highest value calculated was sixty percent, while the lowest was thirty-nine percent. It can once more be noted that the majority of the lexical density percentages obtained all fell into the forty-percent range.

4.2.1 Lexical Diversity Analysis: Type Token Ration Against Lexical Density

4.2.1.1 Lexical Diversity (Type Token Ratio of 0.4 - 0.5)

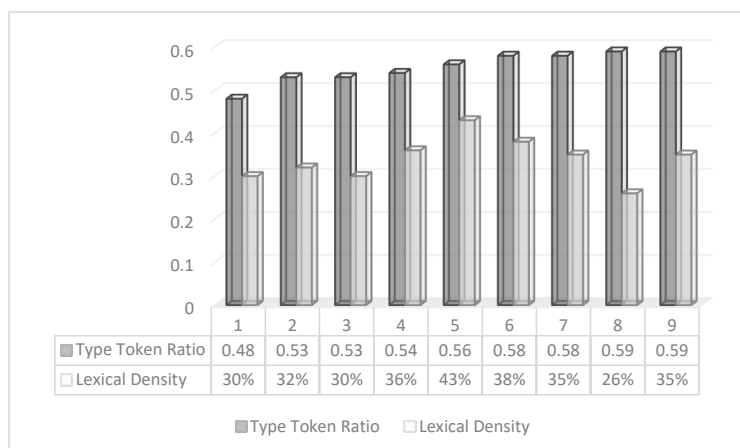


Figure (4.11): Lexical Diversity (TTR of 0.4-0.5)

The figure above reveals the values calculated for nine sections of the text analyzed from EAP learners. Amongst these values, the lowest TTR calculated was 0.25, and the highest was 0.59. Within this category range, it was also noted that two type token ratios of 0.53 were obtained, while only one occurrence of TTRs 0.54 and 0.56 were obtained. The analysis also demonstrated that two of the text divisions possessed TTR values of both 0.58 and 0.59.

In terms of lexical density, the highest value calculated was 43%, while the lowest was 26%. Two instances yielded a lexical density of thirty percent, as well as thirty-five percent. Values of thirty-two percent, thirty-eight percent, and thirty-six percent each occurred just once within the TTR range of 0.4 to 0.5.

The values calculated also demonstrated that the highest range within the 0.5 category correlated with a lexical density value of 26%, the lowest noted in this category. However, 0.59 was amongst the highest type token ratio figures obtained in this category, demonstrating that the increase in

TTR did not mean there was an increase in lexical density. The lack of correlation between TTR and lexical density is further supported with the occurrence of the TTR of 0.48, the lowest within the entire ESP group, which yielded a lexical density of 30%. This value is four percent higher than the lexical density observed for the highest TTR value calculated within this group range.

4.2.1.2 Lexical Diversity (Type Token Ratio of 0.6 – 0.64)

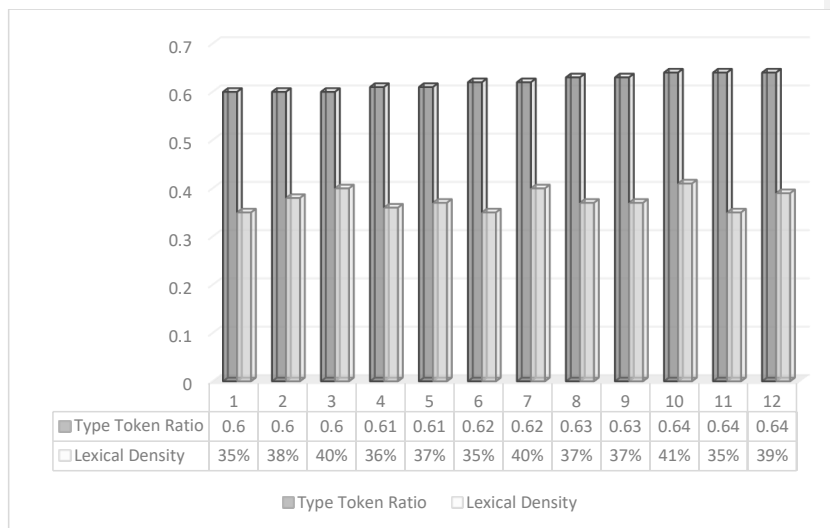


Figure (4.12): Lexical Diversity (TTR of 0.6-0.64)

Figure 12 demonstrates the responses collected from twelve of the total number of samples within the ESP group of learners. The current category being represented in Figure 12 ranges between the Type Token Ratio values of 0.6 to 0.64. Amongst these results, it can be noted that a total of three values occurred at 0.66 and 0.64, while two stances were noted of all the values obtained between 0.61 to 0.63. Amongst these results the lexical density has been noted to range between the fortieth and thirtieth

percentiles. A total of three instances of lexical density have been noted in the range of 40%, those being 41% and two instances of 40%. The remaining nine sample values all witnessed lexical densities occurring in the thirtieth percentile.

The highest value calculated in terms of lexical density amongst the thirtieth category was thirty-nine percent, while the lowest was thirty-five percent. Amongst the lowest value, three varying occurrences were noted. The varying occurrences were also observed amongst the value of 37%. Furthermore, only one instance of thirty-six percent was noted, as well as a single instance of thirty-eight percent. In terms of their correlation with the Type Token Ratio, the highest lexical density observed was, in fact, noted against the highest TTR within this category, that being 0.64.

The lowest lexical density, however, was also observed against a TTR of 0.64, demonstrating that a significant pattern of correlation does not exist between TTR and lexical density in the case of ESP learners. The lack of correlation is also supported with the additional values noted as the second highest case of lexical density (40%) in this group was also noted against the lower TTR value within this specific group range.

4.2.1.3 Lexical Diversity 9Type Token Ratio of (0.65 – 0.66)

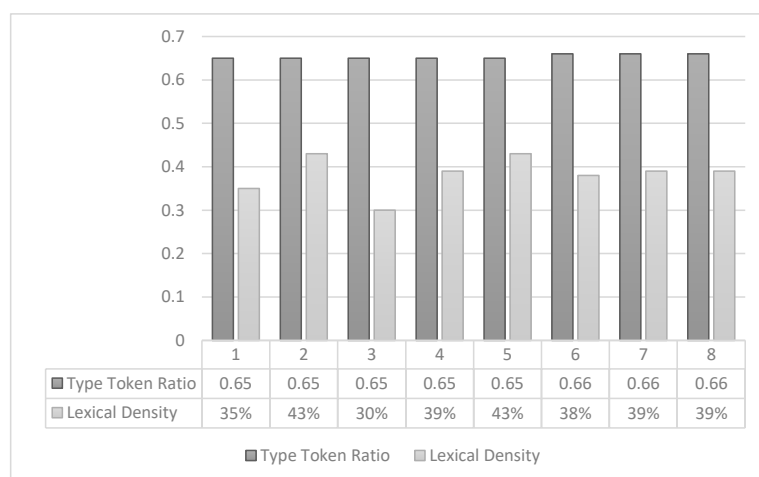


Figure (4.13): Lexical Diversity (TTR of 0.65-0.66)

The figure above demonstrates the lexical densities calculated within the TTR range of 0.65 to 0.66. According to Figure 4.13, a total of eight occurrences were observed within this range. Five of these instances reflected a Type Token Ratio of 0.65, while the remaining three belonged to the range of 0.66. The responses indicated that the highest lexical density within this category occurred at forty-three percent against a TTR value of 0.65. It is also imperative to note that the lowest lexical density observed was at thirty percent, which was five percent less than the lower lexical density value calculated in the 0.6 to 0.64 value group. Moreover, amongst the values calculated, three instances of lexical density were shown to occur at thirty-nine percent, two at forty-three percent, and one instance each of thirty-eight, thirty-five, and thirty percent.

4.2.1.4 Lexical Diversity (Type Token Ratio of 0.67)

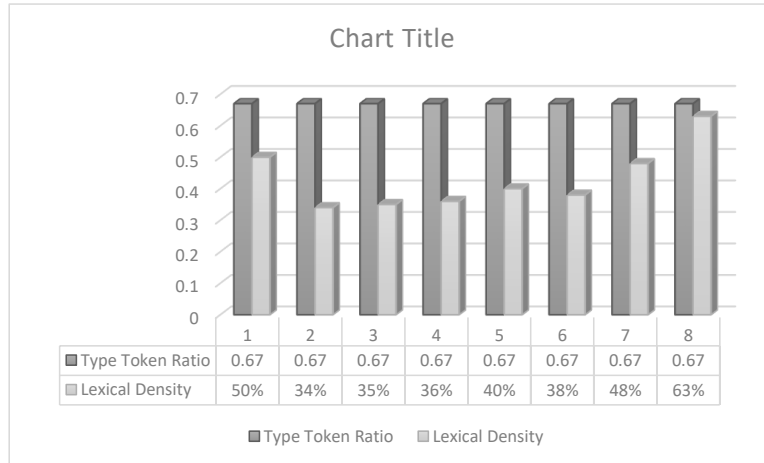


Figure (4.14): Lexical Diversity (TTR of 0.67)

Figure 4.14 shows a TTR of 0.67 amongst the written samples collected from ESP learners, revealing a total of eight samples calculated with a Type Token Ratio of 0.67. Within this range, it was observed that the lowest lexical density was 34%, while the highest value was sixty-three percent in terms of lexical density. It can also be seen that the highest lexical density observed within the ranges of 0.4, 0.5, and 0.6 has been noted at a TTR of 0.67.

Amongst the values calculated, there is only a single instance of 63%. The pattern has been common amongst the remainder of the lexical densities calculated within this TTR range. It need be mentioned that a total of four values were calculated in the thirtieth percentile, two in the fortieth, and only one instance of 50%.

4.2.1.5 Lexical Diversity (Type Token Ratio of 0.68 – 0.69)

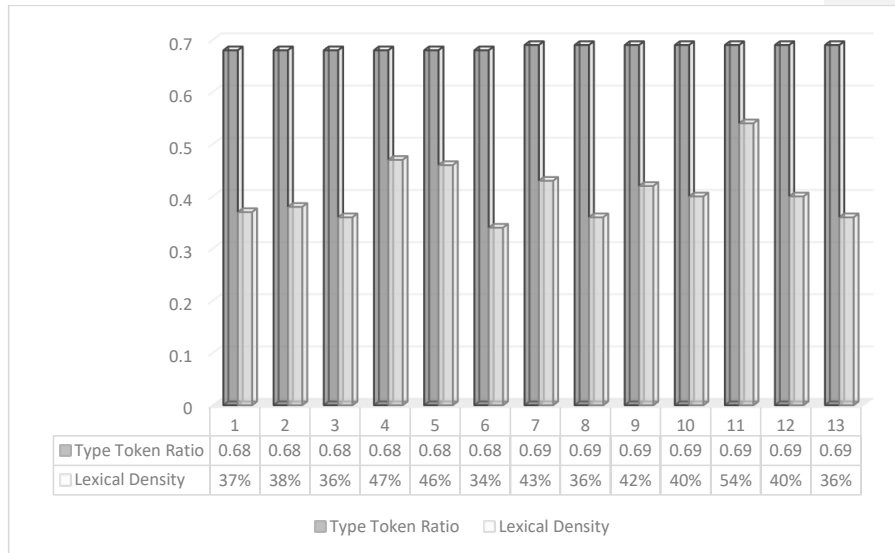


Figure (4.15): Lexical Diversity (TTR of 0.68-0.69)

Figure 4.15 demonstrates the TTR range between 0.68 and 0.69. A total of thirteen values were noted to occur within this specific range, amongst which six samples demonstrated a Type Token Ratio of 0.68, while the remaining seven occurred at a TTR of 0.69. Within this range, it can be noted that the maximum lexical density was at fifty-four percent, while the minimum value occurred at thirty-four percent. Six of the lexical density values calculated occurred within the fortieth percentile, while six of them occurred within the thirtieth. The highest value noted amongst the 40% range was forty-seven, while the lowest was at forty percent, of which there are two instances.

Moreover, the lowest value noted amongst the thirty percent range was thirty-four percent, while the highest was at thirty-eight percent. Three instances were observed to occur at thirty-six percent of lexical density,

while the remainder of the values were only noted at one occurrence each. It must be observed that the highest lexical density noted in this category corresponded to the highest value calculated in terms of TTR. Similarly, the lowest lexical density value was calculated to parallel the lowest value of the Type Token Ratio. Within this specific category of TTR, it has also been noted that the highest lexical density values all occurred against the highest TTR values, while the collective total of the lowest lexical density values also corresponded with the lowest TTR.

4.2.1.6 Lexical Diversity (Type Token Ratio of 0.7 – 0.71)

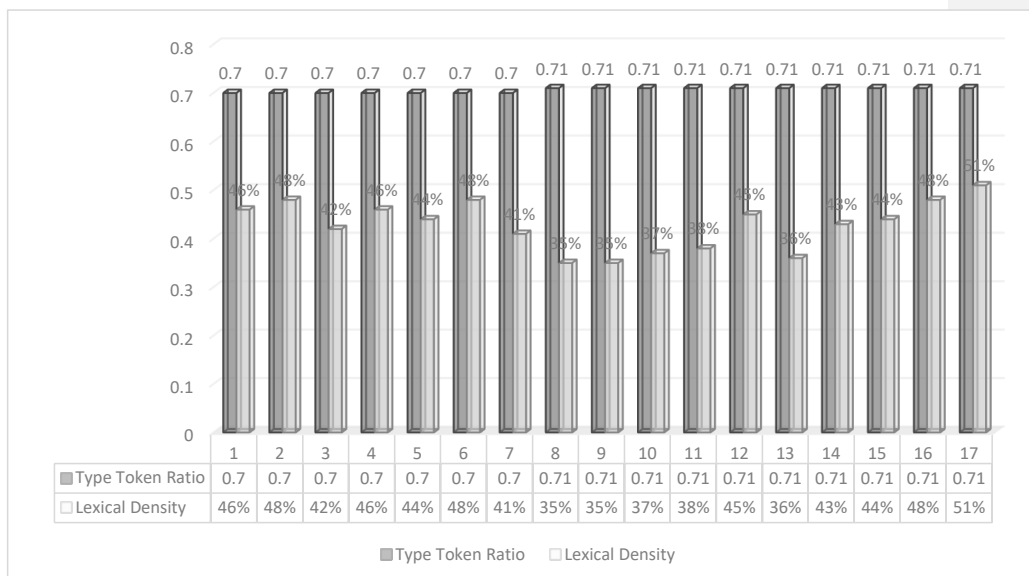


Figure (4.16): Lexical Diversity (TTR of 0.7-0.71)

The figure above demonstrates the values calculated against a TTR of 0.7-0.71. A total of seventeen samples were noted within this range. Amongst the values noted, there were seven instances of 0.7 TTR, while the remaining ten were all 0.71. In terms of lexical density, the lowest value calculated within this category was thirty-five percent, while the highest was fifty-one percent. Both the highest and lowest values corresponded to

a TTR of 0.71. The analysis also revealed that a total of five instances occurred within the thirtieth percentile, while eleven were noted amongst the fortieth. Lexical densities of forty-four and forty-eight percent occurred twice, while the remainder had only a single instance each.

It can also be observed that the fortieth percentile of lexical density was common amongst both TTRs of 0.7 and 0.71. However, the majority within the thirty percent range were all found against Type Token Ratios of 0.71. The results reveal that the lowest lexical density values of the collective ESP category were within a relatively high Type Token Ratio, that being the 0.7 range.

4.2.1.7 Lexical Diversity (Type Token Ratio of 0.72)

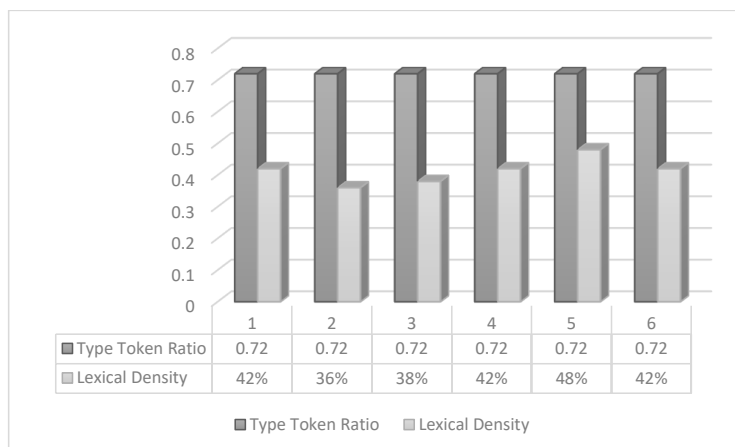


Figure (4.17): Lexical Diversity (TTR of 0.72)

The figure above demonstrates the values obtained with a Type Token Ratio of 0.72. A collective total of six samples were noted with this TTR and the lexical densities occurred between thirty and forty percent. Amongst these samples, the majority, that being four, were within the fortieth percentile, with the highest being forty-eight percent. The remaining values all occurred at 42%, which is also the lowest lexical

density within the entire 0.72 range. The maximum value observed within this percentage range is thirty-eight percent, while the lowest is thirty-six percent. It is also important to state that the highest overall lexical density within this value range is forty-eight percent, while the lowest is thirty-six percent.

4.2.1.8 Lexical Diversity (Type Token Ratio of 0.73 – 0.74)

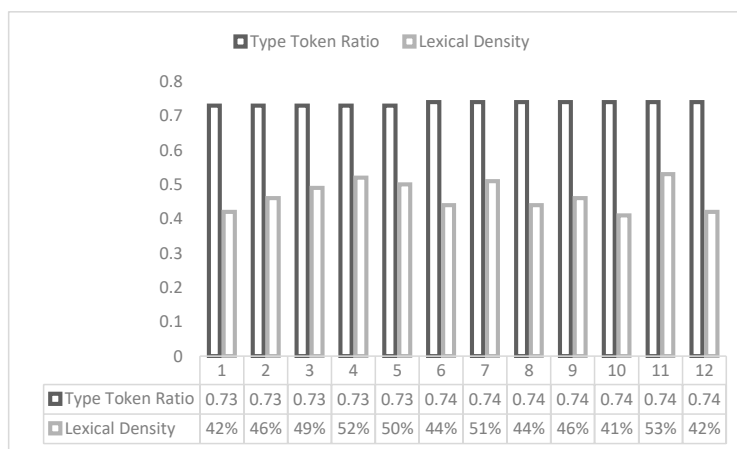


Figure (4.18): Lexical Diversity (TTR of 0.73-0.74)

Figure 4.18 demonstrates the results calculated from twelve samples of ESP learners. The Type Token Ratios observed were calculated at 0.73 and 0.74. Amongst these values, five of them were all calculated at 0.73, while the rest occurred at 0.74. In terms of lexical density, the highest value calculated was fifty-three percent. The fiftieth percentile also shows three occurrences, of which fifty percent occurred with a TTR of 0.74, and the remainder at a TTR of 0.73. Each instance of the fiftieth percentile occurs once, with the overall values being fifty-three, fifty-two, fifty-one, and fifty percent. In addition, the lowest value noted within this category is forty-one percent. The outcome proves significant as it is the highest value observed amongst the other lower values noted in other TTR value

divisions. Within the fortieth percentile, two instances were noted at 42%, 44%, and 46%.

4.2.1.9 Lexical Diversity (Type Token Ratio of 0.75)

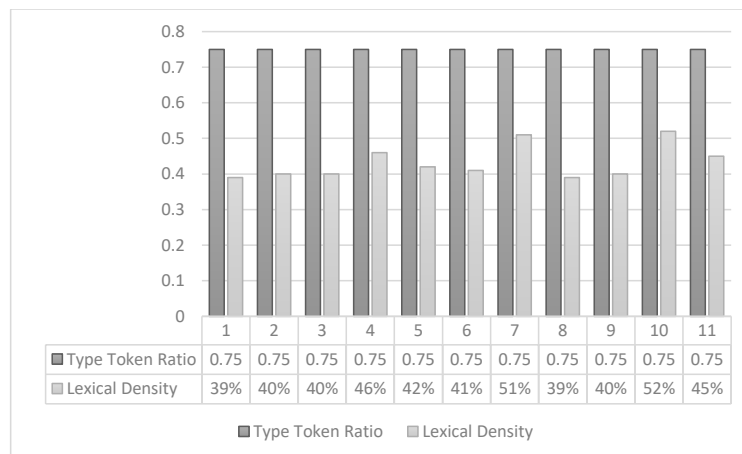


Figure (4.19): Lexical Diversity (TTR of 0.75)

The figure above shows the TTR calculated for eleven written samples collected from ESP learners. The highest lexical density calculated in this section was fifty-two percent, while the highest was thirty-nine percent. Amongst the percentages of lexical density that occurred against a Type Token Ratio of 0.75, there are instances of fifty, forty, and thirty percent. The highest value noted amongst the fifty percent range is fifty-two and fifty-one percent. The majority of occurrences were at forty percent, which was observed in seven of the samples. Three of these occurrences were at exactly forty percent, while the remainder occurred only once each. These were values of forty-one, forty-two, forty-five, and forty-six percent. Furthermore, in terms of the lowest values that occurred, two samples possessed a lexical density of thirty-nine percent.

4.2.1.10 Lexical Diversity (Type Token Ratio of 0.76 – 0.77)

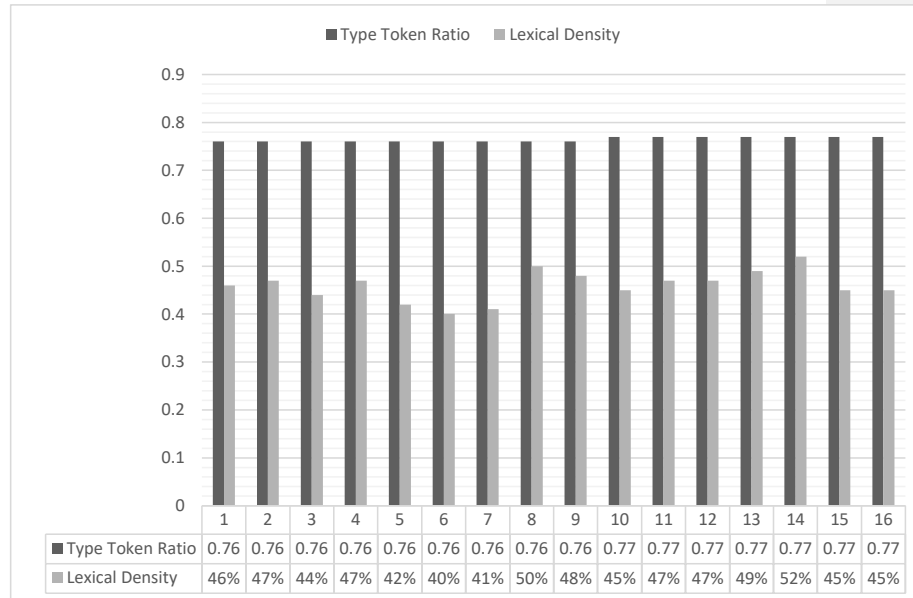


Figure (4.20): Lexical Diversity (TTR of 0.76-0.77)

Sixteen samples were analyzed, and the results obtained have TTRs of 0.76 and 0.77. A total of nine of the samples occurred at 0.76, while the remainder all occurred at 0.77. Amongst these samples, 52% was the highest figure calculated in terms of lexical density. The second highest value was fifty percent, as noted against TTR values within this category. No additional calculations yielded results within the fiftieth percentile. The remainder of lexical densities calculated all occurred within the fortieth percentile. The highest value obtained was forty-nine percent, while the lowest was forty percent. There were four instances of forty-seven percent, and three at forty-five percent. The mentioned values are the most repeated ones calculated for lexical density for Type Token Ratios 0.76 and 0.77.

4.2.1.11 Lexical Diversity (Type Token Ratio of 0.78 – 0.79)

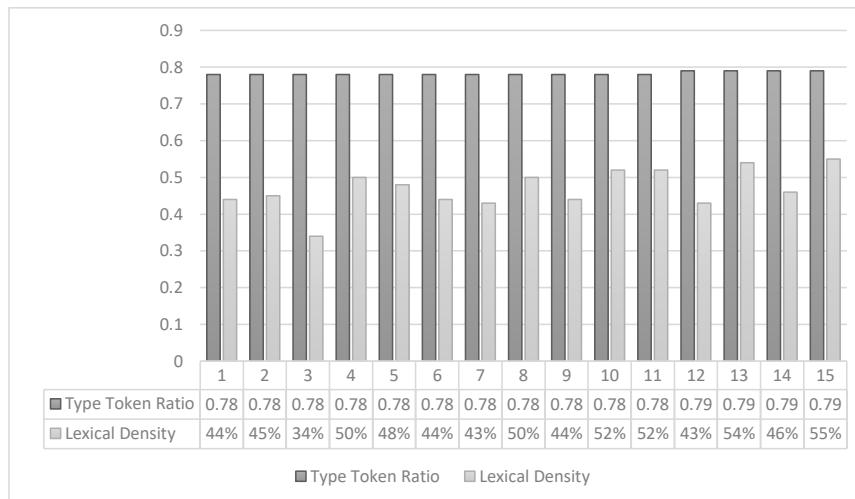


Figure (4.21): Lexical Diversity (TTR of 0.78-0.79)

The values noted in the figure above demonstrate the results obtained for a total of fifteen samples following their analysis of TTRs and lexical density. Figure 4.21 illustrates the samples that possess TTR values of 0.78 and 0.79. Amongst the samples, ten yielded TTRs of 0.78, while only two reached type token ratios of 0.79. When lexical density was considered, six of the samples were within the fiftieth percentile. Eight of the samples were all found to have densities within the forty percent range, while only one case was calculated at thirty-four percent.

Within the fifty percent range, two sample texts were calculated to have lexical densities of fifty percent, two at fifty-two percent, one at fifty-four percent, and one at fifty-five percent. The highest lexical density amongst all the samples in this TTR range was fifty-five percent, while the lowest was thirty-four. Furthermore, when considering the forty percent range, the highest value obtained was forty-eight percent. The lowest

occurred at forty-three percent. A total of three samples were all calculated to have a lexical density of forty-four percent.

With regard to the correlation between TTR and lexical density, the highest occurred against the highest TTR of the group, that being 0.79 paralleled to fifty-five percent. In addition, the lowest TTR, that being 0.78, occurred against a lexical density of a mere thirty-four percent. However, this does not indicate that lexical density will increase or decrease with a similar fluctuation in TTR. In fact, as the values indicated, the majority of the higher percentile instances of lexical density all occurred against the lower TTR within this division.

4.2.1.12 Lexical Diversity (Type Token Ratio of 0.8 – 0.81)

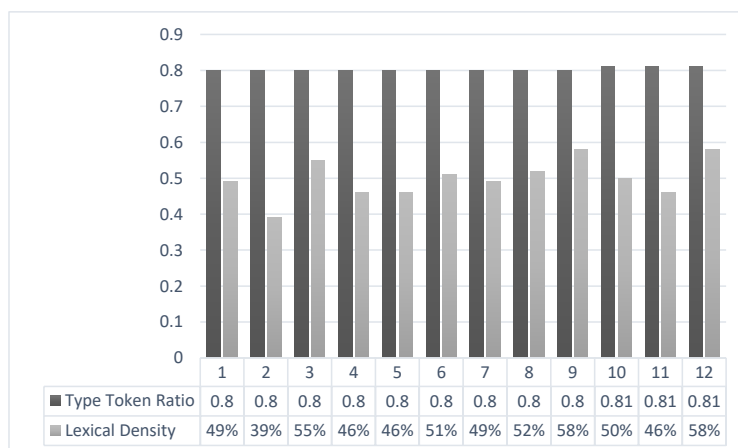


Figure (4.22): Lexical Diversity (TTR of 0.8-0.81)

The figure above illustrates the values obtained in reference to the Type Token Ratios 0.8 and 0.81, from a total of twelve items from the written samples collected from English for Specific Purposes learners. These results reveal that a total of nine out of twelve samples all possessed TTRs of 0.8, while only three were 0.81. Amongst these results, the lexical

density was also calculated, the highest value calculated being fifty-eight percent, the lowest thirty-nine percent. In terms of percentiles, fifty percent of the samples within this category all possessed lexical densities that occurred within the fifty percent range. The lowest value was fifty percent. The remaining values, except for thirty-nine percent, were all within the fortieth percentile. The highest number calculated here was forty-nine percent, while the lowest was forty-six percent, both of which occurred twice.

In terms of the correspondence between TTR and lexical density, the highest lexical density within this group actually occurred against a TTR of 0.8, which was the same value noted against a lexical density of thirty-nine percent. Similar fluctuation between lexical density and TTR was not determined as, for instance, lexical densities within the fortieth percentile could exist against TTRs of 0.81, while densities of 58% also existed against a type token ratio of 0.81.

4.2.1.13 Lexical Diversity (Type Token Ratio of 0.82 – 0.9)

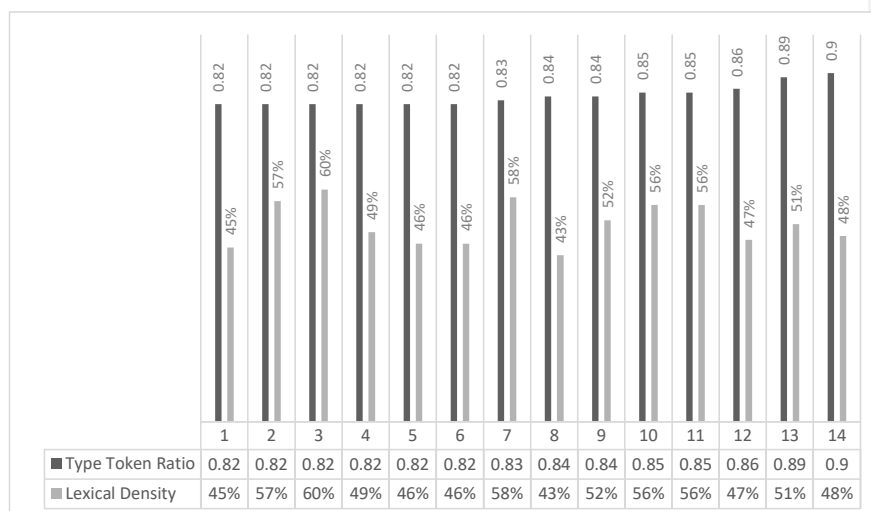


Figure (4.23): Lexical Diversity (TTR of 0.82-0.9)

It can be seen from the figure above that TTR values range between 0.82 and 0.9. The results demonstrate that six values occurred as TTRs of 0.82, while only one existed as a TTR of 0.83. An additional two values were calculated at 0.84 and 0.85 each, and a single case was noted for each TTR value of 0.86, 0.89, and 0.9. The highest lexical density amongst the values in this category was sixty percent, which occurred against a type token ratio of 0.82, while the second highest was fifty-eight percent against a TTR of 0.83. The lowest lexical density in this TTR range was forty-three percent, which was calculated to correspond to the Type Token Ratio of 0.84.

The total number of lexical densities against their Type Token Ratios were organized into three sets of percentage values. The first, with only a single value was sixty percent, which was the highest lexical density observed. The second set of values were in the fifty percent range, in which six values occurred. The highest among these was fifty-eight percent,

followed by fifty-seven percent. Two instances were observed at fifty-six, while only one for each value could be noted at fifty-two and fifty-one percent.

Furthermore, when considering the lexical densities calculated within the forty percent range, seven items were found. The highest amongst them was forty-nine percent, followed by forty-eight, forty-seven, and two cases of forty-six percent. Only a single value occurred for each of forty-five and forty-three.

4.3 Comparative Analysis of EAP and ESP Samples

4.3.1 Statistical Analysis of EAP and ESP Samples

a. Lexical Density

Table (4.1): Statistical Comparison of EAP and ESP Lexical Density

Group Statistics					
Group		N	Mean	Std. Deviation	Std. Error Mean
LD	EAP	160	0.5339	0.07342	0.00580
	ESP	153	0.4376	0.06764	0.00547

From the results recorded in the table above, it can be seen that the higher mean value belongs to the EAP learners, who obtained an average of 0.5339. This value shows a 0.0963 variance between EAP and ESP groups. The ESP group, however, obtained a mean of 0.4376. As a result, it has been noted that the EAP group, on average, had the majority of their values occurring around 0.5, which indicates that EAP lexical density was higher, and thus, the overall lexical complexity of the EAP group was greater than the ESP group. Further analysis also demonstrates that the standard deviation of the two groups differed slightly, with the EAP group receiving

a standard deviation of 0.07342, the ESP group at 0.06764, lower by 0.00578.

The standard error mean relates to the possibility of error occurring with further tests of alternate samples. In other words, if the test was reconducted with other samples from the population, the mean obtained might differ from that of the current test. However, since the test was conducted utilizing the entire population, alternative samples cannot be taken, which means that the standard error mean is irrelevant to this test.

Table (4.2): Levene's Test for Equality of Variances on Lexical Density

		Levene's Test for Equality of Variances		t-test for Equality of Means							
		F	Sig.	t	df	Significance		Mean Difference	Std. Error Difference	95% Confidence	
						One-Sided p	Two-Sided p			Lower	Upper
LD	Equal variances assumed	0.748	0.388	12.045	311	0.000	0.000	0.09623	0.00799	0.08051	0.11195
	Equal variances not assumed			12.067	310.573	0.000	0.000	0.09623	0.00797	0.08054	0.11192

The table above demonstrates the results of Levene's Test for Equality of Variances as conducted for lexical density. The value of significance assigned for this test was 0.05. As per the rules of the test, a value obtained higher or lower than 0.05 in terms of significance, as represented by the abbreviation Sig, indicates that a significance is noted between the outcomes obtained, which means that the difference between the values of the ESP and EAP varied enough to indicate that the two groups differed in the degree of lexical diversity demonstrated in their written discourse.

Table (4.3): Statistical Comparison of EAP and ESP Type Token Ratio

Group Statistics					
Group		N	Mean	Std. Deviation	Std. Error Mean
TTR	EAP	160	0.7260	0.07125	0.00563
	ESP	153	0.7235	0.07515	0.00608

The table above depicts the results of the SPSS analysis carried out for ESP and EAP students in reference to their Type Token Ratio. In terms of the mean, the test reveals that the difference between the two groups was extremely minute. The results show that a value of merely 0.0025 difference is observed. Regarding the two means, the EAP students obtained a slightly higher Type Token Ratio. The difference, however, is not significant enough. Furthermore, in terms of standard deviation, only a minor difference is also observed. The EAP learners, once more, had a higher Type Token Ratio in comparison to the ESP students. The standard deviation was also close in value, with the ESP students obtaining a higher standard deviation in this regard. The standard error mean is once more irrelevant in this context as the entire population of the study was utilized in this research rather than samples from a percentage of the population.

The results indicate that the EAP and ESP learners did not vary profoundly in terms of their Type Token Ratio, which suggests that their lexical diversities were rather similar to a degree. The same is noted in terms of the standard deviation. In this regard, the possibility is raised that the different forms of English language education may yield an identical degree of lexical complexity ultimately. Thus, it may not make a significant difference to learners should they choose either form of English language education when the strength of their vocabulary choices is involved.

4.4 Lexical Sophistication

In accordance with Kyle and Crossley (2015), the TAALES program utilizes a number of frequency logarithms to provide its specific indices. These logarithms include the Kucera-Francis (KF) Logarithm, the Thorndike-Lorge (TL) Logarithm, and the Brown Logarithm. The results of these logarithms can then be collected together to show factors such as register range (Ncats), word range (Nsamp), and frequency (Freq). The frequency value can be obtained by dividing the total frequency scores of the tokens found in the text by the tokens in the text that actually obtained frequency scores. The calculation is achieved through the obtention of frequency lists, including calculations of content words (CW), function words (FW), and the collective scores of all words (AW). The frequency lists utilized in this program are as follows:

1. British National Corpus: This corpus includes one hundred million words of written English, and ten million spoken words of English, all of the British variety. The version adopted for this corpus reference is version 3. This corpus includes language collected from a variety of newspapers, journals, as well as specialist periodicals. Further texts included are memoranda and academic essays. The information within the corpus includes a vast variety of genres and registers ranging across numerous fields.

2. Brown verbal frequency: This corpus is based on the Long-Loud Corpus of English Conversation, which includes a million words. (Svartvik & Quirk, 1980). Vocabulary within this collection relates to the fields of statistics, sociology, linguistics, as well as psychology. Thus, academic language is frequently noted within the content of this specific corpus.

3. N-gram indices: the N-gram indices are bigram and trigram frequencies of spoken and written subcorpora that have been calculated by the BNC. There are about eighty million written words, and ten million spoken subcorpora included in these indices (Crossley et al., 2012). Through TAALES, a total of five indices are calculated in this category. The first type is comprised of non-normalized logarithm-transformed frequency counts. Next, the bigram or trigram frequency totals are divided by the number of bigrams or trigrams actually present in the text as represented in the reference corpus. Another type of index requires the division of the total bigram or trigram frequency values with the number of words in the target text. Furthermore, the calculation of unique bigram or trigram type counts are also conducted. The division of unique bigrams or trigrams in the text by the number of words in the entire text also results in the obtention of proportional scores.

The following indices of lexical sophistication have been included in this study:

- a. Word Range

**Table (4.4): Analysis of TAALES Word Range Amongst EAP and ESP Contexts
for Lexical Sophistication**

<u>Category</u>	<u>EAP</u>	<u>ESP</u>	<u>Variation</u>
KF_Freq_AW	10937.74	10438.97	
KF_Freq_AW_Log	2.984606	3.007129	
KF_Ncats_AW	13.19642	13.26744	0.07102
KF_Nsamp_AW	293.6297	288.0562	5.5735
BNC_Written_Freq_AW	9.972025	9.343832	
BNC_Written_Freq_AW_Log	-0.08769	-0.08457	
BNC_Written_Range_AW	73.68025	73.14383	0.53642
BNC_Written_Freq_CW	1.071123	1.045953	
BNC_Written_Freq_CW_Log	-0.8795	-0.80708	
BNC_Written_Range_CW	56.72952	57.37282	

The values demonstrated in the table above reflect the range of words utilized by the EAP and ESP learners in accordance with Brown's Corpus and the British National Corpus. When the values are observed, it is noted that in terms of word range, both groups yielded almost similar results with only minimum variations between the scores obtained. The differences observed between the values can be noted at 0.53642 when the British National Corpus is considered. Brown's Corpus also demonstrates that the EAP learners obtained a higher value in the range of words they utilized in their texts amongst all the words analysed. The specific corpora selected for this analysis are both comprised of words found in a variety of scholarly

and academic work, as well as words collected from newspapers, articles, and journals.

When contrasting the British National Corpora and Brown's Corpus, the EAP students can be judged to have obtained higher scores than those of the ESP learners. The difference in scores might indicate that the EAP students possessed a more diverse range of vocabulary in their language use. As a result, the written language they produced resulted in higher scores in comparison to the ESP learners. The higher scores could also be attributed to the fact that EAP learners are not limited to diversity of lexical knowledge and vocabulary in their field. They could be exposed to a higher number of lexical units and vocabulary items in their university courses, and thus, utilize more words from a range of corpora in comparison to the ESP learners. Additionally, it can be recognized that ESP language learners are primarily exposed to specific professional jargon associated with their fields of study when they are educated in the English language at university level. The exposure means that the range of vocabulary they are exposed to, and thus utilize in their written work, could be limited to their specific field of study.

However, it can be noted that amongst the collective totals found in the analysis, the overall number of words obtained were limited to begin with. For example, the word range in accordance with the British National Corpus was only 73.68025 for EAP learners, while the ESP learners had a total of 73.1383. The same dynamic is noted for Brown's Corpus, although these values are slightly higher. For instance, the EAP and ESP learners both yielded approximately three hundred words from the corpus. However, the words were amongst a total of nine thousand five hundred words, which indicates that they are a merely a small percentage of the overall total. The values indicate that the general lexical sophistication of

the two groups were significantly low, although the EAP learners possessed a slightly higher lexical sophistication in comparison to the ESP learners.

b. N-Gram Frequency

N-grams refer to combinations of words such as bigrams and trigrams. Based on the results of the TAALES analysis, the following N-gram frequency values were found:

Table (4.5): Analysis of TAALES Bigrams Amongst EAP and ESP Contexts for Lexical Sophistication

<u>Category</u>	<u>EAP</u>	<u>ESP</u>	<u>Variatio</u> <u>n</u>
Brysaert_Concreteness_Combined_AW	2.384283	2.383411	0.000872
Brysaert_Concreteness_Combined_CW	2.753625	2.764189	0.01056
Brysaert_Concreteness_Combined_FW	1.912577	1.854661	0.05791

			<u>Variation</u>
BNC_Written_Bigram_Freq_Normed	0.325181	0.321237	0.00394
BNC_Written_Bigram_Freq_Normed_Log	-1.42126	-1.52107	0.09981

The table above reflects the use of bigrams amongst both EAP and ESP learners in the texts produced and analyzed through the TAALES program. As per the results indicated, the total number of bigrams produced in the texts were limited in general. The analysis was conducted on all words from all texts for each of the learner categories (AW), the content words (CW), and the function words (FW). As per the results, it is noted that the EAP learners utilized slightly more bigrams amongst the total number of words and the function words. On the other hand, when the utilization of the content words was considered, the ESP group obtained a slightly higher value with a mere variation of 0.01056 between the two groups in total. Amongst the remainder of the results, all obtained a

variation of less than 0.6 between the values of bigrams for both EAP and ESP learners. Thus, the same degree of lexical sophistication is observed between the two learner categories.

Similar outcomes can be observed when considering the British National Corpus in terms of the bigrams noted. A slight difference that is less than one is observed between the bigrams found in EAP and ESP learner texts, demonstrating that the degree of lexical sophistication between the two groups are almost identical.

- **Trigrams**

Table (4.6): Analysis of TAALES Trigrams Amongst EAP and ESP Contexts for Lexical Sophistication

<u>Category</u>	<u>EAP</u>	<u>ESP</u>	<u>Variation</u>
BNC_Written_Trigram_Freq_Normed	0.01527	0.017532	0.002262
BNC_Written_Trigram_Freq_Normed_Log	-2.21138	-2.2302	0.01882

The results in the table above demonstrate the occurrence of trigrams within the texts collected and analyzed from both EAP and ESP learners. It was noted that the frequency of trigram occurrences was extremely limited. The results also indicated, in this case, that the ESP group possessed a slightly higher count of trigram occurrences in their language than those in the language of the EAP group. The overall difference between the two groups with their respective values of 0.0022662 and 0.01882, indicates that, based on trigram analysis, the overall lexical sophistication of both groups was limited. Furthermore, the same degree of lexical sophistication can be said to occur between the EAP and ESP groups based on the limited variation between the values obtained for each. The shared degree of lexical sophistication indicates that overall, both learner

groups did not differ in this matter, which means their academic exposure to the English language in either course type was not influential in enhancing their lexical sophistication, with specific reference to trigram value and occurrence.

- **Academic Lists**

Table (4.7): Analysis of TAALES Academic Lists Amongst EAP and ESP Contexts for Lexical Sophistication

<u>Category</u>	<u>EAP</u>	<u>ESP</u>	<u>Variation</u>
COCA_Academic_Range_AW	0.61723	0.595149	0.022081
COCA_Academic_Frequency_AW	10098.25	9499.843	598.407
COCA_Academic_Range_Log_AW	-0.46523	-0.42963	0.0356
COCA_Academic_Frequency_Log_AW	2.94094	2.954067	0.013127
COCA_Academic_Range_CW	0.387416	0.369994	0.017422
COCA_Academic_Frequency_CW	1058.374	1053.968	4.406
COCA_Academic_Range_Log_CW	-0.76574	-0.68067	0.08507
COCA_Academic_Frequency_Log_CW	2.185924	2.267681	0.081757
COCA_Academic_Range_FW	0.953482	0.959843	0.006361
COCA_Academic_Frequency_FW	23324.92	23180.03	144.89
COCA_Academic_Range_Log_FW	-0.02554	-0.02299	0.00255
COCA_Academic_Frequency_Log_FW	4.045638	4.065839	0.020201

The table above reflects the occurrence of academic words in terms of both range and frequency within the texts produced by both EAP and ESP learners. It can be noted that, overall, the frequency of academic words from the Corpus of Contemporary American English differed between EAP and ESP language learners. The frequency noted amongst the EAP learners (10098) was higher than that of the ESP learners (9499) by an approximate

value of 598. The range was calculated for all the words utilized within the texts provided by the learners during data collection.

When specifically analyzing the frequency of content words, a very slight difference can be noted between EAP and ESP values. The difference is limited to 4.406, with EAP learners utilizing approximately 1058.374, ESP students utilizing 1053.968. The values demonstrate that, once more, the lexical sophistication between the two groups is almost identical. The difference observed is not sufficient for the variation to be considered significant.

Commented [MP1]: Inserted comma

When function words are considered, both groups once more utilized function words at the same frequency, with very limited variations noted between the two. In general, both groups utilized a frequency of approximately two thousand three hundred function words. When comparing between the content and function words, it can be noted that a greater number of function words were utilized compared to the content words, which can indicate that the learners generally demonstrated a less degree of lexical sophistication in their language as content words are generally associated with more complex and sophisticated forms of vocabulary.

4.5 Interview Analysis

4.5.1 EAP

Demographic Information:

The results obtained from the respondents indicated that all of the four individuals interviewed were native speakers of the Arabic language. They were also all EAP students studying English language and literature at the University of Bahrain. Based on the responses obtained, all the respondents shared a similarity in terms of their academic progress, which was that they

took the project writing course, ENGL450, during their final year at university.

Age:

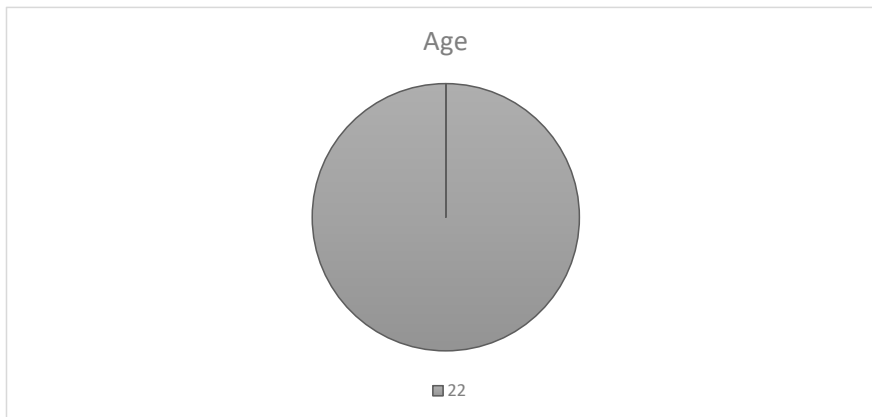
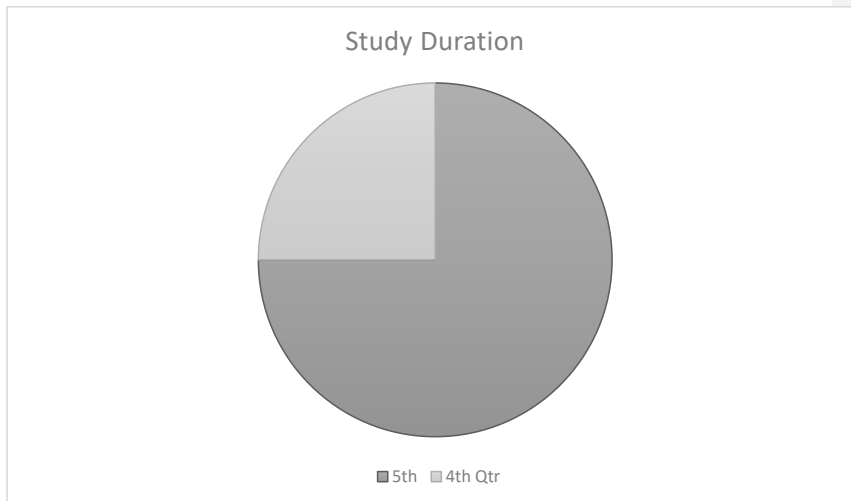


Figure (4.24): EAP Interview Analysis - Age

The graph above demonstrates that all the interviewees involved in this thesis were twenty-two years old. The uniform age of the respondents permits a fair comparison to be made in the responses provided by the interviewees as it explores the same age range, allowing a comparison to be made specifically in relation to this age group. Thus, age is prevented from acting as a determining factor in relation to the opinions presented, consistency being maintained as a result.

Period of Study

Figure (4.25): EAP Interview Analysis – Study Duration



All but one of the interviewees that took part in this research were in their fifth and final year of undergraduate study. Only one of the respondents was a fourth-year student. This student, however, was also studying in their final year at university level.

GPA

Interviewee 1: 2.47

Interviewee 2: 2.64

Interviewee 3: 3.11

Interviewee 4: 3.5

Based on the data provided, it can be observed that the learners chosen for the purpose of the interview were all at varying ranges of academic

ability, which was noted from their overall GPAs. This specific selection was done in order to ensure that responses were collected from learners from a variety of academic and mental capabilities, which could also affect the linguistic abilities of the learners as well.

The following information was gathered in reference to respondents' views on EAP and ESP in their university education.

•*How do you feel these English courses affect your fluency in the English language?*

The responses obtained for this question conveyed varying opinions. One of the respondents communicated that they found the grammar courses to be extremely easy when they commenced their language studies. His language did not change profoundly, and the interviewee mentioned that he actually found his proficiency level to decrease following these courses due to their ease. However, the linguistic courses such as morphology and phonology significantly affected the interviewee's proficiency in English. The individual found that these courses affected their GPA negatively but was then aided through the varying literature courses provided to them. The effects were due to the extent of reading involved in the courses, which also permitted them additional language exposure.

Another respondent also communicated that the majority of the support they obtained from their university courses was from literature courses. These courses enabled them to learn more vocabulary and improved their language as a result. The respondent also claimed that they became more conscious of their vocabulary choices, which especially affected their choice of words in speech. Improvements in lexical complexity were most significant for the respondent as a result of the literary exposure.

An additional response from one of the interviewees actually conveyed that the English courses did not affect the participant's language, but rather, their own knowledge of the language was enhanced during their academic career rather than having been acquired from their courses.

• *Which courses have had the most significant effect on your lexical complexity (vocabulary richness) and why?*

In response to this question, two of the respondents once more conveyed that literature courses affected them most profoundly and aided in the improvement of their language skills. One interviewee mentioned that, as they consistently came across new vocabulary terms while reading literary texts, they would research words and translate them. These words then became a part of their own vocabulary, resulting in significant improvements in their vocabulary. The new vocabulary also became a prominent factor in the respondent's spoken and written language as they would utilize the learnt vocabulary items in their own essays and practise their use further. The conclusion was drawn based on the respondent's own statement, *"If I found words I liked I would write them down and use them more, which was why my language improved in literature and I used to write essays a lot, so the more I used them the more I learnt."*

One of the respondents indicated that the most profound difference was elicited by the American literature courses that they took at the university, while another conveyed the following, *"they gave us a variety of novels and we were exposed to these texts from different authors and different centuries, so we were exposed to many authors, and this was much more than the technical knowledge in grammar books. So lexical knowledge came mainly from the literary books."*

• *Which aspect of your language improved when you took these courses?*

The interviewees provided varying answers to this question. One of the responses indicated that the most significant factor that improved was the interviewee's comprehension of language as well as their vocabulary. They also greatly improved in terms of their exposure to culture and contexts, developing the ability to analyze culture and sarcasm. Further improvements the respondent noted was in their pronunciation of language, as well as their ability to write using the English language.

Another student who was interviewed mentioned that their ability to accurately use tenses was affected by the courses taken. They especially improved in their use of the perfect tense, which was not expected of them until they took grammar courses at higher levels. Additionally, the interviewee mentioned that their syntactic awareness of language was greatly affected by the courses taken, as well as their linguistic creativity and ability to express themselves. As stated by the interviewee, *"I believe literature also really improved my linguistic creativity by encouraging me to express myself in different ways."* This individual also stated that they would initially commence with writing simple sentences, which then developed into the ability to produce more complex structures with varying figures of speech. Hence, cultural awareness was also achieved.

Another interviewee mentioned that their pronunciation and punctuation were significantly affected literature courses. They mentioned the following, *"I also think my punctuation and my pronunciation really improved, and courses like phonology really helped me understand how to pronounce certain words based on their positioning in words."*

• *Do you feel you also require taking an ESP course to improve the lexical sophistication of your language?*

Based on the responses collected during the interview, it was noted that one of the respondents did believe it would be more advisable to take an ESP course in addition to her own major in EAP as it would enable her to perform better in the workplace. The interviewee stated that they were especially interested in taking English for law as it would aid them in their career as a translator later on in life.

Another response was obtained from an interviewee who mentioned that they would also prefer being able to take an ESP course in English for business as it would enable them to perform better in future prospects of their possible professional career choice, while yet another added that taking ESP courses would be beneficial to them as well.

• *Would you recommend language learners taking ESP courses to improve the richness of their vocabulary and lexical choices?*

The interviewees provided varying responses to this question. One of them offered the response that the choice of taking ESP courses greatly depended on the actual field a person would choose later on. This interviewee continued by stating that English major students especially do not require being exposed to ESP courses as they already obtain a wide variety of vocabulary and language exposure. Thus, studying another form of English is not mandatory nor beneficial to them.

Another interviewee offered an opinion that resembles the former by stating that ESP courses are not as helpful as EAP courses. However, one of the interviewees stated that all learners should actually take any opportunity provided to them to work with the language and gain additional exposure to it. The interviewee stated, *"I would advise people to take EAP*

courses, but I do know not everyone follows its career paths. I would though advise them to dip their toes into the EAP ocean.”

•Why would you recommend that language learners take ESP/EAP courses to improve the richness of their vocabulary and lexical choices?

Varying justifications were provided by those interviewed in reference to the matter. One of them stated that it is best for all language learners to gain exposure to language from varying sides. Another communicated that it is best for learners to focus on their individual fields rather than expand to varying jargon from other fields. Yet another actually stated that ESP learners should actually be exposed to EAP courses instead as some ESP learners may not progress in language proficiency and skills beyond knowledge of the specific field’s jargon. This interviewee stated, *“From what I noticed, for example, business students are good with the terms related to business but can’t construct good sentences or write even paragraphs.”*

Another interviewee stated that they felt that if one entered the field of translation especially, it would be advisable for them to take ESP courses as well to become more familiar with the jargon associated with those fields in order to reduce time and effort when researching terms.

•What are the gains of teaching L2 learners either, or both, EAP/ESP courses?

Varying advantages were given for both EAP and ESP courses. According to one of the interviewees, it was mentioned that the advantages of EAP included the opportunity to be exposed to varying branches of the language, with an in-depth knowledge of theory, literature and linguistics. The same interviewee mentioned that, in terms of ESP, the advantages

include the obtention of a unique edge in future professions due to exposure to professional jargon. As stated by the interviewee, *“You really need a good grip on your English language to know the list of needed vocabs to excel at your job as the course is extremely specific and you immerse yourself in the field you choose. It is very specific and serves your specific needs.”*

According to one interviewee, EAP serves to provide learners with a very advanced knowledge of language across varying aspects. Learners are also exposed to the language specifically and in different registers. Another advantage is that learners will be able to then differentiate significantly between colloquial and formal use of the language. In terms of ESP, learners will know how to use the language in their specific fields.

Yet another interviewee stated that EAP is very advantageous to learn because it exposes learners to day-to-day use of language, which cannot be developed through profession and needs to be studied formally. ESP is advantageous as it exposes learners to the language’s vocabulary, but this respondent did not find any further advantages in ESP courses.

• *What are the drawbacks of teaching L2 learners either, or both, EAP/ESP courses?*

One interviewee found that ESP was rather limited and did not really teach learners how to use the language. In terms of EAP, the respondent found that drawbacks were dependent on the course being taken. Some of the courses are rich enough to be beneficial, while others do not provide real-life exposure and actual language use as per actual contexts. This individual states, *“If it does not offer real-life exposure by making them speak and use it in real life contexts and not academic contexts, it is useless.”*

Another interviewee stated that EAP courses can be disadvantageous at times as research is not conducted actively by students, and that students may take courses that they are not actually interested in. When ESP is concerned, the individual stated, *“You are not really learning English as a whole language. You are just zooming in on one small corner of the language and do not get the full experience.”* Thus, the interviewee was of the opinion that ESP courses provide limited exposure.

From one interviewee came the idea that learners exposed to EAP are not exposed to specific information related to other fields. Also, cultural boundaries exist. For instance, the interviewee related the following situation, *“I know someone who only studied EAP and that person does not understand slang words. So, when he goes to a group of people exposed to slang words and double negation, he does not understand them at all.”* In terms of ESP, it was stated that language knowledge is limited to specific connotations and different layers and vocabulary comprehension are not achieved. As stated, *“my representation of the world depends on my field. Thus, it can make people refer to vocabulary only in terms of their specific purposes and fields.”*

4.5.2 ESP

Demographic Information

Age

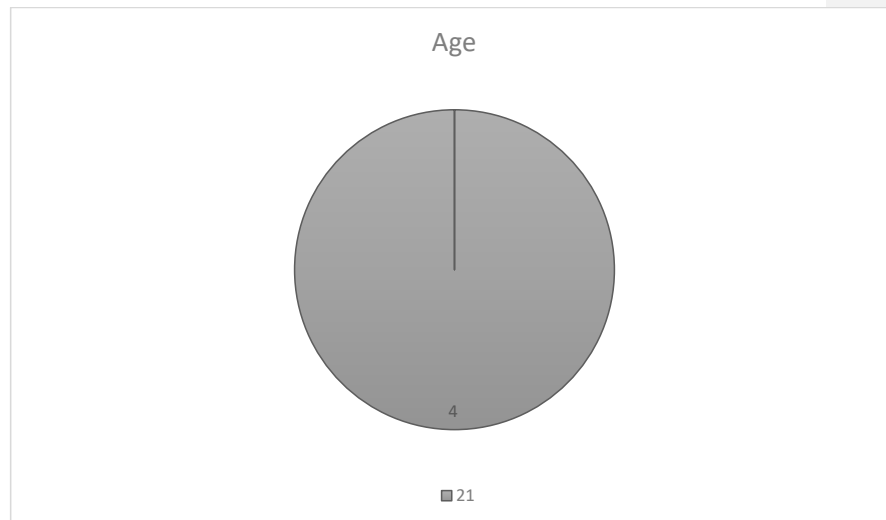


Figure (4.26): ESP Interview Analysis - Age

As indicated by the diagram above, all the learners interviewed within the ESP field were twenty-one years old. Once more, the maintenance of all interviewees within a single age group ensures that alternative social factors do not influence the responses provided. Rather, all the learners come from both a similar academic background in terms of their exposure to the English language at university level, as well as being of the same age, which permits more reliable results and opinions to be communicated.

Year of Study

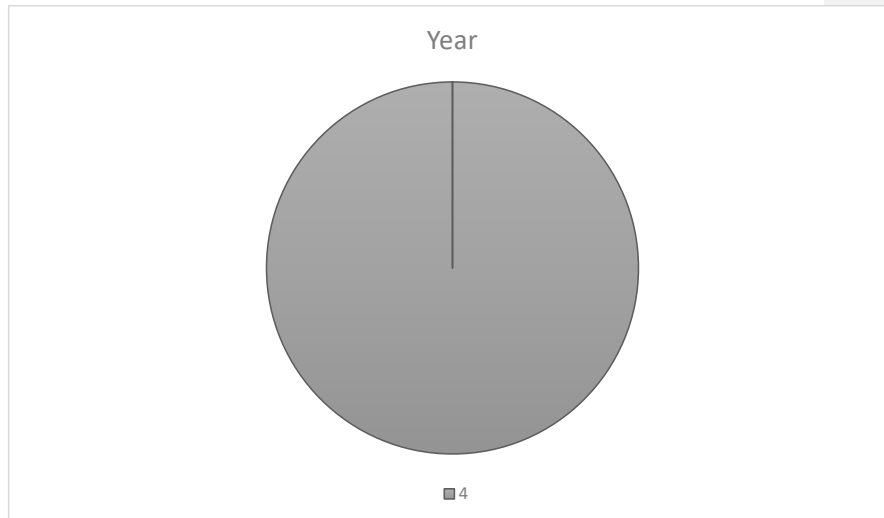


Figure (4.27): ESP Interview Analysis – Year

The figure above demonstrates that all the learners were in their fourth year of study at university level. As they all study the same major, it also indicates that they have been exposed to the same English language ESP courses, at the same period of time, which would eliminate other external factors from influencing the opinions provided by the interviewees.

GPA

Interviewee 1: 3.13

Interviewee 2: 3.11

Interviewee 3: 3.53

Interviewee 4: 2.6

The results obtained in terms of the GPAs showed that three of the interviewees all possess a GPA within the 3.0 range. The lowest range found was a GPA of 2.6.

Interview Analysis

• *What type of English courses are you currently taking at university?*

All the interviewees mentioned that they were currently taking an advanced English course in project writing. They were primarily exposed to business related vocabulary, as well as made to conduct research, from which they are expected to submit a research paper.

• *How do you feel these English courses affect your fluency in the English language?*

Respondents provided different responses to this question. One mentioned that the most affected skill was their writing skill. The interviewee mentioned that they previously possessed weak writing skills, especially in terms of their sentence structure. Taking ENGL341 improved the interviewee's writing abilities. Another participant indicated that their general fluency in English greatly improved, especially since the course was taken on campus rather than online. A third respondent mentioned that their fluency was not affected by initial English courses provided in their field of study, but that ENGL341 made the most difference to their proficiency. Their speaking and writing both improved as well.

• *Which courses have had the most significant effect on your lexical complexity and why?*

The majority of the respondents provided approximately similar answers to this question. The majority of the participants conveyed that ENGL341

affected their skills the most. One of the participants mentioned that their writing skills improved with this course, while others mentioned that the additional challenge pushed them to learn more about the language. One of the participants mentioned that ENGL341 possessed difficult and diverse vocabulary, which was the reason why it was more complex and improved their language proficiency,

• *Which aspect of your language improved when you took these courses?*

One respondent indicated that their writing and communication skills were enhanced when taking the ESP courses mentioned. Further improvements were noted in their sentence structure and their ability to paraphrase. Another interviewee mentioned that their improvements in language were most prevalent in their speaking skills. In both writing and speaking, their syntactic structures and accuracy of their sentences were most affected.

• *Do you feel you also require taking an EAP course to improve the lexical sophistication of your language?*

The respondents primarily had a positive reaction towards the notion of taking EAP courses to support their ESP courses. One interviewee stated the following: *“I think I definitely need to take an EAP course as well. I think it will be more difficult, but I will need it and it will benefit me.”* One of the other interviewees mentioned that exposure to EAP courses will enable them to improve their proficiency and skills to an extent that will be able to support them in their future profession as well. Another respondent mentioned the following: *“I think I do need to take an EAP course as well because we have a lack of exposure to English as a language. We only take specific vocabulary related to business and have learnt about creating*

business documents, but we only focus mostly on report writing. It is too limited.”

•Would you recommend language learners taking ESP/EAP courses to improve the richness of their vocabulary and lexical choices?

Most of the respondents stated that EAP courses would be highly beneficial for them to improve their academic level in terms of language proficiency. A respondent stated the following in support: *“EAP courses would be better because they are exposed to more of English as a language. Students will focus on the language entirely.”*

•Why would you recommend language learners to take EAP courses to improve the richness of their vocabulary and lexical choices?

The main benefit communicated by the respondents was that their knowledge base will increase. They also opted for L2 learners to take a mixture of ESP and EAP courses. One of the participants also mentioned that they prefer ESP courses as it would positively affect their knowledge of their field of choice. As stated by the interviewee, *“For us, it helps us in business a lot because we need to know these terms and vocabulary for our field.”*

•What are the gains of teaching L2 learners, either, or both, EAP/ESP courses?

Different interviewees provided varying responses to this question. One only focused on taking a mixture of the two types of language courses, mentioning that this will help them both learn about their own specialization, as well as their English language skills. However, other opinions were also stated. A participant stated that the two courses should

not be studied simultaneously as confusion can be caused. One skill learnt through the ESP courses, which the participant opted for, was writing reports specific to their own field of work and profession of choice.

• *What could be the drawbacks of teaching L2 learners, either, or both, EAP/ESP courses?*

The majority of the respondents conveyed the same response, which indicated that EAP courses would be challenging to commence with. However, they also highlighted the positive factor, being that early exposure would yield extremely positive results. One of the participants indicated that both courses are fundamental to their language progression by stating the following: *“ESP would be too limited to learn the actual language, but EAP courses also would not expose me too much to actual English. As a student, I need to have more than just vocabulary, and not just business terms. However, also, with EAP, I won’t know what I need to know for business purposes, which is another issue as well. We need a mixture to be balanced.”*

Chapter Five

Conclusion

Chapter Five

Conclusion

This chapter summarizes the results obtained and analyzed from the research. The conclusion divides the results based on the three measures of lexical complexity, as well as the outcome of the interviews. Further recommendations and suggestions for future research were also included within this section.

5.1 Conclusion

The research reported in this thesis aimed to provide an insight into the effects of English for Academic purposes and English for Specific Purposes on the language proficiency of EFL university learners. Specific focus was placed on vocabulary use, with all three measures of lexical complexity being taken into consideration to measure the overall lexical complexity of the learners' written discourse. University students' writing was collected from both disciplines, with approximately 9500 words analyzed from learners across each discipline mentioned.

Various forms of data analysis were conducted for the purpose of this thesis. In terms of lexical density, the learners' written samples were analyzed utilizing the lexical density formula provided in the methodology section. According to the quantitative analysis of the written samples, it was noted that a value of seventy-three percent was the highest lexical density observed amongst all EAP learners, while the ESP learners managed to attain a lexical density of sixty-three percent as the highest value. Based on the outcome noted, it can be comprehended that the EAP learners were capable of utilizing a greater number of lexical words in comparison to function words. Thus, it can be comprehended that the proficiency and vocabulary usage of the EAP students exceeded that of the

ESP students when lexical density is considered. The results, thus indicate that EAP could be the better form of language study that results in learners adopting better choices of vocabulary in their written discourse.

Research question 2 was addressed through the analysis of the TTR values demonstrated by the learners when their written discourse was analyzed. Where TTR is concerned, the ESP learners scored a maximum value of 0.9, while the lowest value obtained was 0.48. EAP in contrast demonstrated that the highest TTR attained was 0.88, which was slightly lower than that of the ESP students. The outcome indicates that learners of ESP were able to utilize a greater degree of diverse vocabulary in a single written production of discourse in comparison to the EAP learners. Additionally, the lowest TTR value for the EAP learners was 0.52, while the ESP learners returned a minimum of 0.48. The results thus demonstrate that even the lowest EAP diversity in discourse was still higher than the lowest in ESP. The results in terms of lexical diversity indicate that EAP learners could possess a great diversity in their use of the English language, although exceptional cases can be noted.

When the mean of the two disciplines were assessed in terms of lexical density, it was found that the EAP learners possessed a higher mean at a value of 0.5339, in comparison to the 0.4376 obtained by the ESP learners. Once more, the EAP learners' lexical complexity was measured as being higher in comparison to the latter. Further consideration can also be placed on the mean of lexical diversity. Both EAP and ESP students attained a value that was nearly equivalent, with the former's mean being 0.7260, and the latter being 0.7235. Thus, the lexical diversity amongst the two groups did not differ significantly. Further statistical analysis of the results utilizing the independent sample T-test revealed that the variations between the results obtained were indeed significant. This is clear when considering

the value obtained from Leven's Test for Equality of Variances. Thus, in reference to research question 1(a), it can be stated that significant differences in the lexical density of EAP and ESP students can indeed be noted.

The indices addressed in the lexical sophistication measure included word range, N-gram frequency, trigrams and bigrams. In terms of word range, different corpora were selected as part of the analysis, which include Brown's Corpus and the British National Corpus. Both groups demonstrated approximately the same results, with the EAP learners obtaining slightly higher results in comparison to the ESP learners. The results thus convey that EAP learners possessed a much greater range of vocabulary in their written discourse. Further results on lexical sophistication were in terms of both bigrams and trigrams. In total, the number of bigrams produced were limited amongst both groups, with EAP learners producing a slightly higher number of bigrams in comparison to the ESP learners. As a result, it is noted that both groups possessed the same level of lexical sophistication based on bigram numbers.

Furthermore, in terms of trigrams, the occurrence was also limited, as was noted with bigrams. The outcome showed that both EAP and ESP groups demonstrated the same degree of lexical sophistication as well. Thus, the results of the analyses showed that the difference between EAP and ESP learners across different measures of lexical complexity overall yielded similar outcomes, that being that EAP learners possessed a slightly higher level of lexical complexity, with the exception of sophistication showing a similar degree. Therefore, in reference to research question 1(b), it is noted that EAP learners generally possessed slightly greater lexical sophistication in comparison to the ESP students, demonstrating that extensive exposure to the English language through an EAP course could

result in higher exposure to a diverse and unique set of lexemes and vocabulary.

Interviews were conducted to address research question number three which inquired into the perceptions of EAP and ESP university learners on their lexical complexity levels. When the interview results were analyzed, it was noted that the ESP students mainly noticed their language improving when they took a research course during their final year, while the EAP students were more influenced by literature courses which aided them in enhancing their lexical complexity. In terms of taking either ESP or EAP courses to improve language and vocabulary, the majority of the EAP students mentioned that they preferred their own EAP courses to improve their language as it exposed them to a wider range of vocabulary. However, the ESP students mentioned that EAP courses could be beneficial to them, but adding that the courses need to be introduced earlier during their academic experience to be more beneficial. Overall, both quantitative and qualitative means of analysis demonstrated that EAP learning had a greater influence on improving the learners' lexical complexity in comparison to ESP exposure. However, further research can be conducted to better support the results, which can be achieved through collecting and analyzing data from more EAP and ESP learners from additional universities across the Kingdom of Bahrain.

5.2 Limitations and Delimitations

The limitations of the research were primarily defined in terms of the overall number of samples that were analyzed for the purpose of the thesis. Due to time restrictions and challenges accessing both EAP and ESP contexts, obtaining additional samples was made difficult. Furthermore, problems were also encountered when attempting to collect samples and interview learners from various universities across the Kingdom of

Bahrain, which were due to specific regulations enforced on data collection within those university contexts.

Chapter 6

References

Chapter 6

References

References

- Amelia, P., Silvana, T., & Sinar, T. (2020). Lexical density and grammatical intricacy in debaters' speeches. *Language Literacy: Journal of Linguistics, Literature, and Language Teaching*, 4(1), 168-184. doi:10.30743/ll.v4i1.2519
- Bulté, B., & Housen, A. (2014). Conceptualizing and measuring short-term changes in L2 writing complexity. *Journal Of Second Language Writing*, 26(4), 42-65.
- Cindy, C., & James, P. (2007). *The Psychological Functions of Function Words*. New York: Psychology Press.
- Cobb, T., & Horst, M. (2004). Is there room for an academic wordlist in French? In 2549558149 1718089079 P. Boogards & 2549558150 1718089079 B. Laufer (Authors), *Vocabulary in a second language: Selection, acquisition, and testing* (pp. 15-38). Amsterdam: Bejnamins.
- Crossley, S. A., Cobb, T., & McNamara, D. S. (2013). Comparing count-based and band-based indices of word frequency: Implications for active vocabulary research and pedagogical applications. *System*, 41(4), 965-981. doi:10.1016/j.system.2013.08.002
- Crystal, D. (1982). *Profiling Linguistic Disability*. London: Edward Arnold.
- Crystal, J. A., & Smith, C. G. (2023, April 24). Bahrain. Retrieved April 29, 2023, from <https://www.britannica.com/place/Bahrain>
- Daller, H., van Hout, R., & Treffers-Daller, J. (2003). Lexical richness in the spontaneous speech of bilinguals. *Applied Linguistics*, 24(2), 197-222.

- Dewi, R. (2017). Lexical and syntactic complexities in undergraduate students' research articles and their correlations to their quality. *Jurnal Pendidikan Humaniora*, 5(3), 123-133. doi:10.17977/um030v5i32017p123
- Donesch-Jezo, E. (2012). *English for Specific Purposes: What does it mean and why is it different from teaching General English?* Nagpur. Retrieved 2012, from https://www.researchgate.net/publication/308914571_English_for_Specific_Purposes_What_does_it_mean_and_why_is_it_different_from_teaching_General_English.
- Donesch-Jezo, E. (2012). *English for Specific Purposes: What does it mean and why is it different from teaching General English?* Retrieved from https://www.researchgate.net/publication/308914571_English_for_Specific_Purposes_What_does_it_mean_and_why_is_it_different_from_teaching_General_English
- Douglas, D. (2000). *Assessing English for Specific Purposes*. Cambridge University Press.
- Duran, P., Malvern, D., Richards, B. J., & Chipere, N. (2004). Developmental Trends in Lexical Diversity. *Applied Linguistics*, 25(2). doi:10.1093/applin/25.2.220
- Durán, P., Malvern, D., Richards, B., & Chipere, N. (2004). Developmental trends in lexical diversity. *Applied Linguistics*, 25(2), 220–242.
- Education*. bahrain.bh. (2019)
- Ellis, R., & Barkhuizen, G. (2005). *Analysing accuracy, complexity, and fluency; Analyzing learner language*. Oxford University Press.
- Gregori-Signes, C., & Clavel-Arroitia, B. (2021). Analysing lexical density and lexical diversity in university students' written discourse. *Procedia - Social and Behavioral Sciences*, 198, 546-556. doi:10.1016/j.sbspro.2015.07.477
- Gunderson, L. (2008). *ESL (ELL) literacy instruction: A guidebook to theory and practice* (2nd ed.). doi:10.4324/9780203894217.
- Gunderson, L., D'Silva, R., & Odo, D. (2009). *ESL (ELL) Literacy Instruction: A Guidebook to Theory and Practice* (2nd ed.). Routledge.

- Halliday, M. (1985). *Spoken and Written Language*. Deakin University.
- Hamp-Lyons, L. (2011). English for academic purposes: 2011 and beyond. *Journal of English for Academic Purposes*, 10(1), 2-4. doi:10.1016/j.jeap.2011.01.001
- Hayes, D. P., & Ahrens, M. (1988). Vocabulary simplification for children: A special case of 'motherese.'. *Journal of Child Language*, 15, 395-410.
- Hyland, K. (2006). *English for Academic Purposes*. Routledge.
- Hyland, K., & Hamp-Lyons, L. (2002). EAP: Issues and Directions. *Journal of English for Academic Purposes*, 1, 1-12.
- Johansson, V. (2008). Lexical diversity and lexical density in speech and writing: A developmental perspective. *Department Of Linguistics And Phonetics Working Papers*, 53, 61-79.
- Kuperman, V., Stadthagen-Gonzales, H., & Brysbaert, M. (2012). Age-of-acquisition ratings for 30,000 english words. *Behavior Research Methods*, 44(4), 978-990.
- Lahmann, C., Steinkrauss, R., & Schmid, M. (2015). Factors Affecting Grammatical and Lexical Complexity of Long-Term L2 Speakers' Oral Proficiency. *Language Learning*, 66(2).
- Laufer, B., & Nation, P. (1995). Vocabulary Size and Use: Lexical Richness in L2 Written Production. Retrieved from https://www.lexutor.ca/vp/laufer_nation_95.pdf
- Laufer, B., & Nation, P. (1995). Vocabulary size: Lexical richness in L2 written production. *Applied Linguistics*, 16, 307-322.
- Ilis, Rod & Gary Barkhuizen. 2005. *Analysing learner language*. Oxford: Oxford University Press.
- Malvern, D., Chipere, N., & Duran, P. (n.d.). Lexical Diversity and Lexical Sophistication in First Language Writing. In 2549613298 1718122276 B. Richards (Ed.), *Lexical Diversity and Language Development*. Palgrave Macmillan.
- Malvern, D., Richards, B., Chipere, N., & Duran, P. (2004). *Lexical diversity and language development: Quantification and assessment*. Palgrave Macmillan.

- Mazgutova, D., & Kormos, J. (2015). Syntactic and lexical development in an intensive English for Academic Purposes programme. *Journal of Second Language Writing*, 29, 3-15.
doi:10.1016/j.jslw.2015.06.004
- McNamara, D., Louwerse, M., McCarthy, P., & Graesser, A. (2010). Cohmetrix: Capturing linguistic features of cohesion. *Discourse Processes*, 47, 292-330.
- Nagy, W. E., & Scott, J. A. (2000). Vocabulary processes. In 2549598942 1718113614 M. L. Kamil, 2549598943 1718113614 P. B. Mosenthal, 2549598944 1718113614 P. D. Pearson, & 2549598945 1718113614 R. Barr (Authors), *Handbook of Reading Research* (Vol. 3, pp. 269-284). Mahwah, NJ: LEA.
- Nodoushan, S., & Ali, M. (2020). English for Specific Purposes: Traditions, Trends, Directions. *Studies in English Language and Education*, 1(1), 247–268.
- Pallotti G (2009) CAF: Defining, refining and differentiating constructs. *Applied Linguistics* 30(4): 590–601.
- Papadima, O. (2020). An overview of ESP and EAP in language pedagogy. In 2549544894 1718081217 J. Badio (Author), *Categories and Units in Language and Linguistics* (pp. 115-127).
- Pennock Speck, B., & Clavel-Arroitia, B. (2020). Analysing Lexical Density, Diversity, and Sophistication in Written and Spoken Telecollaborative Exchanges. *Computer Assisted Language Learning Electronic Journal* , 22(3), 230–250.
- Pennock-Speck, B., & Clavel-Arroitia, B. (2021). Analysing Lexical Density, Diversity, and Sophistication in Written and Spoken Telecollaborative Exchanges. *CALL-EJ*, 22(3), 230-250.
- Rahayu, F., Utomo, A., & Setyowati, R. (2021). Syntactic and Lexical Complexity of Undergraduate Students' Essays: A Comparison Study between L1 and L2 Writings. *Indonesian Journal of English Language Teaching and Applied Linguistics*, 5(2), 251-263.
- Rahmansyah, H. (2012). *Grammatical Intricacy and Lexical Density of the SMA Student's Textbooks* (Master's thesis, University of Medan). Medan.
- Read, J. (2000). *Assessing Vocabulary*. Cambridge: Cambridge University Press.

- Rogers, L. (2010). An all-round challenge 1. *English Teaching Professional*, 69.
- Salmani-Nodoushan, M. A. (2020). English for specific purposes: Traditions, trends, directions. *Studies in English Language and Education*, 7(1), 247-268. doi:10.24815/siele.v7i1.16342
- Skehan, P. (2009). Article Navigation JOURNAL ARTICLE Modelling Second Language Performance: Integrating Complexity, Accuracy, Fluency, and Lexis. *Applied Linguistics*, 30(4), 510–532. <https://doi.org/10.1093/applin/amp047>
- Skehan, P. (2021). Modelling Second Language Performance: Integrating Complexity, Accuracy, Fluency, and Lexis. *Applied Linguistics*, 30(4), 510-532.
- Starfield, S. (2011). Doing critical ethnographic research into academic writing: The theory of the methodology. In Belcher, D., Johns, A. & Paltridge, B. (eds.), *New directions in English for specific purposes research*, Ann Arbor, University of Michigan Press, 174 -196.
- Thawarom, T., & Singhasiri, W. (2020). Lexical richness of one-minute speaking task by science and Technology University students. *The Journal of AsiaTEFL*, 17(1), 70-86. doi:10.18823/asiatefl.2020.17.1.5.70
- Thornbury, S., & Slade, D. (2006). *Conversation: From Description to Pedagogy*. Cambridge: Cambridge University Press.
- Tidball, F., & Treffers-Daller, J. (2008). Analysing lexical richness in French learner language: What frequency lists and teacher judgements can tell us about basic and advanced words. *Journal Of French Language Studies*, 299-313.
- To, V., Fan, S., & Thomas, D. (2013). Lexical Density and Readability: A Case Study of English Textbooks. *Internet Journal of Language, Culture and Society*, (37).
- Ure, J. (1969). 'Lexical density and register differentiation. In G. Perren & J. Trim, *Applications of linguistics. Selected papers of the Second International Congress of Applied Linguistics*. Cambridge University Press.
- Ure, J. (1971). Lexical density and register differentiation. *Applications of Linguistics*, 23(7), 443-452.

- V., Fan, S., & Thomas, D. (2013). Lexical Density and Readability: A Case Study of English Textbooks. *Internet Journal of Language, Culture and Society*.
- Van Hell, J., & De Groot, A. (1998). Conceptual representation in bilingual memory: Effects of concreteness and cognate status in word association. *Bilingualism: Language and Cognition*, 1(3), 193-211.
- Wolfe-Quintero, K., Inagaki, S., & Kim, H. (2001). SECOND LANGUAGE DEVELOPMENT IN WRITING: MEASURES OF FLUENCY, ACCURACY, AND COMPLEXITY. *Studies in Second Language Acquisition*, 23(3), 423-425.
- Wolf-Quintero, K., Inagaki, S., & Kim, H. (1998). *Second language development in writing: Measures of fluency, accuracy, and complexity*. University of Hawaii Press.
- Yu, G. (2010). Lexical Diversity in Writing and Speaking Tasks Performances. *Applied Linguistics*, 31(2), 236-259.
- Zareva, A. (2019). Lexical complexity of academic presentations. *Journal of Second Language Studies*, 2(1), 71-92. doi:10.1075/jsls.18003.zar
- Zeno, S. M., Ivens, S. H., Millard, R. T., & Duvvuri, R. (1995). *The Educator's Word Frequency Guide*. Brewster: Touchstone Applied Science Associates.

الملخص

استكشفت الأطروحة وقارنت بين درجة التعقيد المعجمي التي ترد في الخطاب المكتوب لمتعلمي اللغة الانجليزية لأغراض أكاديمية (EAP) ومتعلمي اللغة الانجليزية لأغراض محددة (ESP) بين طلاب الجامعات في مملكة البحرين. تم تحليل ما مجموعه تسعة آلاف وخمسمائة كلمة من المتعلمين حول كل سياق، كما تم تطبيق طرق مختلفة للتحليل الكمي لقياس الكثافة المعجمية والتنوع والإلمام في الخطاب الناتج. تم حساب الكثافة المعجمية من خلال إيجاد النسبة المئوية من العدد الإجمالي للكلمات المعجمية في مقابل عدد الكلمات في خطاب المتعلمين من كل سياق. تم تحليل النتائج بعد ذلك باستخدام برنامج "SPSS"، والذي أظهر أن النتائج دلالات هامة بالفعل، حيث حقق متعلمي اللغة الانجليزية لأغراض أكاديمية (EAP) كثافة معجمية أعلى بقليل مقارنة بمتعلمي اللغة الانجليزية لأغراض محددة (ESP). تم أيضاً حساب قيمة "TTR" ومعالجتها بعد ذلك عبر برنامج "SPSS" لإثبات وجود تباين ملحوظ بين القيم التي حصل عليها طلاب اللغة الانجليزية لأغراض أكاديمية (EAP) في مقابل طلاب اللغة الانجليزية لأغراض محددة (ESP). عند تقصي الإلمام المعجمي، تم استخدام برنامج "TAALLES 2.0" لمقارنة الخطاب المكتوب للمتعلمين في مقابل المتون النصية لـ "بريتيش ناشيونال كوربس" (British National Corpus)، و التردد اللفظي للمتون النصية لـ "براون" (Brown Verbal Frequency)، و أيضاً مؤشرات مقاطع الكلمات "N-gram". تم تقصي نطاق الكلمات ومقاطع الكلمات الثلاثية (trigram) والثنائية (bigram)، وأظهرت النتائج أن متعلمي اللغة الانجليزية لأغراض أكاديمية (EAP) يمتلكون مفردات أكثر إلماماً عند النظر في القوائم الأكاديمية ونطاق الكلمات. ومع ذلك، فيما يتعلق بمقاطع الكلمات الثلاثية (trigram) والثنائية (bigram)، كان لدى متعلمي اللغة الانجليزية لأغراض محددة (ESP) نسبة أعلى قليلاً من مقاطع الكلمات "N-gram"، على الرغم من أن التردد العام كان محدوداً في كلا السياقين. كما تم تطبيق المزيد من التحليل الكمي لاستكشاف وجهات نظر المتعلمين من الجانبين إزاء التعقيد المعجمي، مما أظهر أن متعلمي اللغة الانجليزية لأغراض أكاديمية (EAP) يؤمنون بأن التعقيد المعجمي الخاص بهم يتأثر بشكل إيجابي بمجال دراستهم. عندما تمت مقابلة متعلمي اللغة الانجليزية لأغراض محددة (ESP)، لفت الرأي المشترك عموماً إلى أن دورات اللغة الانجليزية لأغراض محددة (ESP) حظيت بالأفضلية عند النظر في التطوير المهني. ومع ذلك، اقترح المتعلمون أن التعرض المبكر إلى درجة من دورات اللغة الانجليزية لأغراض أكاديمية (EAP) موصى به أيضاً لتعزيز المهارات اللغوية والكفاءة. يمكن أن يتيح إجراء المزيد من البحوث امكانية لجمع عدد أكبر من الخطابات المكتوبة واستخدام نطاق أوسع من أدوات تحليل البيانات لتحقيق نتائج أكثر تفصيلاً.

الكلمات المفتاحية: اللغة الانجليزية لأغراض أكاديمية (EAP)، اللغة الانجليزية لأغراض محددة (ESP)، الكثافة المعجمية، التنوع المعجمي، الإلمام المعجمي



جامعة البحرين
كلية الآداب
قسم اللغة الانجليزية وآدابها

قياس التعقيد المعجمي لطلاب الجامعات من متعلمي اللغة الانجليزية لأغراض أكاديمية (EAP)
ومتعلمي اللغة الانجليزية لأغراض محددة (ESP): دراسة مقارنة

بحث تكميلي بمتطلبات درجة الماجستير في دراسات اللغة الإنجليزية التطبيقية

من إعداد:

زينب نادر محمود

20135181

بإشراف:

د. ديانا عبدالكريم الجهري

استاذ مساعد في اللغويات

جامعة البحرين

مملكة البحرين

أبريل 2023

