

**A LONGITUDINAL STUDY OF THE USE OF
HIGH FREQUENCY NOUNS AMONG
MALAYSIAN SECONDARY SCHOOL
STUDENTS**

ZHANG XINXIN

**FACULTY OF LANGUAGES AND LINGUISTICS
UNIVERSITY OF MALAYA
KUALA LUMPUR**

2017

**A LONGITUDINAL STUDY OF THE USE OF HIGH
FREQUENCY NOUNS AMONG MALAYSIAN
SECONDARY SCHOOL STUDENTS**

ZHANG XINXIN

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

**FACULTY OF LANGUAGES AND LINGUISTICS
UNIVERSITY OF MALAYA
KUALA LUMPUR**

2017

UNIVERSITY OF MALAYA
ORIGINAL LITERARY WORK DECLARATION

Name of Candidate: Zhang Xinxin

Registration/Matric No: TGB130028

Name of Degree: Master of English as a Second Language

Title of Project Paper/Research Report/Dissertation/Thesis (“this Work”):

A LONGITUDINAL STUDY OF THE USE OF HIGH FREQUENCY NOUNS
AMONG MALAYSIAN SECONDARY SCHOOL STUDENTS

Field of Study: Corpus Linguistics and Second Language Acquisition

I do solemnly and sincerely declare that:

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

Candidate’s Signature

Date:

Subscribed and solemnly declared before,

Witness’s Signature

Date:

Name:

Designation:

ABSTRACT

This study investigates how second language development takes place over time. It takes nouns as a starting point, and explores how selected nouns are used differently in the texts taken from the Longitudinal Corpus of Languaculturer Narrative Texts (Chau, 2015). The texts (based on the same prompt) consist of four parts written at four different points in time by the same group of students. Five most frequent nouns *river*, *girl*, *flowers*, *lake*, and *friend* together with their respective singular or plural forms were selected for the investigation. Pattern Grammar (Hunston & Francis, 1999) was adopted to explore how these selected words were used over time. The findings show both a constructive process and a reductionist process (Chau, 2015) operate at the same time in learner data, and indicate that language development is a non-linear and complex process (e.g., Larsen-Freeman, 2006, 2009; Ortega, 2009; Ellis, 2011; Chau, 2015). Meanwhile, it is found that students are able to use more adjectives and nouns to modify the words in focus at a later period in time, which shows students' lexical repertoire has expanded across time. Implications of the study are provided.

Key words: nouns, second language development, Pattern Grammar, corpus linguistics

ABSTRAK

Kajian ini menyiasat bagaimana perkembangan bahasa kedua berlaku dari masa ke masa. Petikan teks yang digunakan dalam kajian ini diambil daripada Longitudinal Corpus of Languaculturer Narrative Texts (Chau, 2015). Petikan teks dalam korpus ini ditulis oleh pelajar sekolah menengah dalam empat bahagian dan ditulis pada empat masa yang berbeza. Lima kata nama yang paling kerap digunakan adalah, river, girl, flowers, lake, dan friend dan digunakan dalam pengkajian ini. Pattern Grammar (Hunston & Francis, 1999) dan kaedah linguistic corpus telah digunakan untuk mengkaji cara lima perkataan ini digunakan dalam tempoh 24 bulan. Hasil kajian menunjukkan bahawa terdapat konstruktif dan reduktif process berlaku serentak di kalangan para pelajar. Hasil kajian ini membuktikan bahawa perkembangan bahasa merupakan non-linear dan kompleks (Larsen-Freeman, 2006, 2009; Ortega, 2009; Ellis, 2011; Chau, 2015). Selain itu, para pelajar dalam korpus ini menunjukkan bahawa mereka boleh kerap menggunakan adjektik dan kata nama untuk mengubah suai kata kunci pada perikat akhir. Ini membuktikan bahawa himpunan leksikal para pelajar berkembang dengan masa. Implikasi kajian telah dinyatakan.

Kata kunci: kata nama, perkembangan bahasa kedua, Pattern Grammar, corpus linguistik

ACKNOWLEDGMENTS

I would like to express my sincere thanks to many people, without whom the completion of this dissertation would not be possible. First and foremost, I would like to express my heartfelt and profound gratitude to my supervisor Dr. Chau Meng Huat whose invaluable insight and constructive comments guided this work from its beginning to the end, and whose patience and support helped me to get through all the difficulties in this academic journey.

I would also like to extend my appreciation to the lecturers of Faculty of Languages and Linguistics of University Malaya for sharing their knowledge with me.

Last but not least, my eternal gratitude goes to my family who have always loved me unconditionally. I dedicate this dissertation to them.

TABLE OF CONTENTS

ABSTRACT.....	iii
ABSTRAK.....	iv
ACKNOWLEDGMENTS.....	v
TABLE OF CONTENTS.....	vi
LIST OF FIGURES.....	x
LIST OF TABLES.....	xi
CHAPTER 1: INTRODUCTION.....	1
1.1 Background of the study.....	1
1.2 Contextual information.....	3
1.3 Problem statement.....	4
1.4 Research objective.....	4
1.5 Scope and limitations.....	5
1.6 Significance of the study.....	6
1.7 Conclusion.....	6
CHAPTER 2: LITERATURE REVIEW.....	8
2.1 Learner language.....	8
2.2 Analysing learner language: methods from L2 acquisition.....	9
2.3 Corpus linguistics.....	10
2.3.1 Learner corpora.....	12
2.3.2 Pattern Grammar: noun patterns.....	13
2.4 Previous studies on nouns.....	16
2.5 Conclusion.....	18

CHAPTER 3: METHODOLOGY.....	19
3.1 Corpus used in this research.....	19
3.2 Tools for analysis.....	20
3.2.1 The concordancer	20
3.2.2 The word list tool.....	22
3.3 Data analysis.....	24
3.4 Conclusion.....	25
CHAPTER 4: RESULTS.....	26
4.1 RIVER.....	26
4.1.1 Patterns of use of RIVER.....	27
4.1.2 Observation on instances under each pattern.....	29
4.1.2.1 Instances under each pattern of use of RIVER.....	29
4.1.2.1.1 det N.....	29
4.1.2.1.2 adj N.....	31
4.1.2.1.3 N prep n.....	34
4.1.2.1.4 N clause.....	36
4.1.2.2 Emergent lexis with RIVER.....	39
4.2 GIRL.....	40
4.2.1 Patterns of use of GIRL.....	41
4.2.2 Observation on instances under each pattern.....	43
4.2.2.1 Instances under each pattern of use of GIRL.....	43
4.2.2.1.1 det N.....	43
4.2.2.1.2 N prep n.....	49
4.2.2.1.3 N clause.....	51
4.2.2.2 Emergent lexis with GIRL.....	53
4.3 FLOWER.....	55

4.3.1	Patterns of use of FLOWER.....	56
4.3.2	Observation on instances under each pattern.....	57
4.3.2.1	Instances under each pattern of use of FLOWER.....	57
4.3.2.1.1	det N.....	57
4.3.2.1.2	N prep n.....	63
4.3.2.1.3	N clause.....	66
4.3.2.2	Emergent lexis with FLOWER.....	67
4.4	LAKE.....	69
4.4.1	Patterns of use of <i>lake</i>	70
4.4.2	Observation on instances under each pattern.....	71
4.4.2.1	Instances under each pattern of use of <i>lake</i>	71
4.4.2.1.1	det N.....	71
4.4.2.1.2	N prep n.....	74
4.4.2.2	Emergent lexis with <i>lake</i>	75
4.5	FRIEND.....	76
4.5.1	Patterns of use of FRIEND.....	77
4.5.2	Observation on instances under each pattern.....	78
4.5.2.1	Instances under each pattern of use of FRIEND.....	79
4.5.2.1.1	det N.....	79
4.5.2.1.2	N clause.....	84
4.5.2.2	Emergent lexis with FRIEND.....	84
4.6	Conclusion.....	86
 CHAPTER 5: DISCUSSION AND CONCLUSION.....		87
5.1	The research questions.....	87
5.2	Implications of the study.....	94
5.3	Recommendations for future study.....	95

5.4 Conclusion.....	96
References.....	98
Appendix A.....	104
Appendix B.....	105
Appendix C.....	106

University of Malaya

LIST OF FIGURES

Figure 3.1: A fragment from concordances of <i>river</i> in the subset of LoCLaNT.....	21
Figure 3.2: A fragment from concordances of <i>river</i> to its left.....	21
Figure 3.3: A fragment from concordances of <i>river</i> to its right.....	22
Figure 3.4: A frequency list of words from the subset of LoCLaNT.....	22
Figure 5.1: The number of prepositional phrases of the selected nouns.....	91
Figure 5.2: The types of clauses of the selected nouns.....	92

University of Malaya

LIST OF TABLES

Table 3.1: Top 10 nouns at Time 1.....	23
Table 4.1: Raw and normalized frequency of use of RIVER.....	27
Table 4.2: Patterns of use of RIVER.....	28
Table 4.3: Determiners of the word RIVER.....	29
Table 4.4: Adjectives of the word RIVER.....	32
Table 4.5: Prepositional phrases of the word RIVER.....	34
Table 4.6: Clauses of the word RIVER.....	37
Table 4.7: Emergent lexical items with RIVER.....	39
Table 4.8: Frequency information of lexical items with RIVER.....	39
Table 4.9: Raw and normalized frequency of use of GIRL.....	40
Table 4.10: Patterns of use of GIRL.....	42
Table 4.11: Determiners of the word GIRL.....	44
Table 4.12: The innovative language use of determiners of GIRL.....	49
Table 4.13: Prepositional phrases of the word GIRL.....	50
Table 4.14: Clauses of the word GIRL.....	51
Table 4.15: Emergent lexical items with GIRL.....	54
Table 4.16: Frequency information of lexical items with GIRL.....	54
Table 4.17: Raw and normalized frequency of use of FLOWER.....	55
Table 4.18: Patterns of use of FLOWER.....	56
Table 4.19: Determiners of the word FLOWER.....	58
Table 4.20: The innovative language use of determiners of FLOWER.....	63
Table 4.21: Prepositional phrases of the word FLOWER.....	63
Table 4.22: Clauses of the word FLOWER.....	67

Table 4.23: Emergent lexical items with FLOWER.....	68
Table 4.24: Frequency information of lexical items with FLOWER.....	68
Table 4.25: Raw and normalized frequency of use of <i>lake</i>	69
Table 4.26: Patterns of use of <i>lake</i>	70
Table 4.27: Determiners of the word <i>lake</i>	71
Table 4.28: Prepositional phrases of the word <i>lake</i>	74
Table 4.29: Emergent lexical items with <i>lake</i>	75
Table 4.30: Frequency information of lexical items with <i>lake</i>	76
Table 4.31: Raw and normalized frequency of use of FRIEND.....	76
Table 4.32: Patterns of use of FRIEND.....	77
Table 4.33: Determiners of the word FRIEND.....	79
Table 4.34: Clauses of the word FRIEND.....	84
Table 4.35: Emergent lexical items with FRIEND.....	85
Table 4.36: Frequency information of lexical items with FRIEND.....	85
Table 5.1: Patterns of use of the selected nouns across time.....	88
Table 5.2: The innovative language use of determiners.....	90
Table 5.3: Emergent lexis used with the selected nouns.....	92

CHAPTER 1

INTRODUCTION

This study attempts to explore how a second language (L2) develops across time, with the tools and methods from corpus linguistics. It begins with an overview of this study. The current chapter includes background of the study, contextual information, problem statement, research objective, scope and limitations, and significance of the study.

1.1 Background of the study

Second Language Acquisition (SLA) is the study of how students learn a language subsequent to the learning of their first language (L1) (Ortega, 2009; Saville-Troike, 2012). The additional language is known as L2, even though it can be the third, fourth, or eleventh language to be acquired (Ellis, 2008; Stefansson, 2013).

The purpose of research in this field has always been to better understand the nature of learner language and its development (Ellis, 2008). The early research shows immense interest in Contrastive Analysis (CA) which refers to the study of a pair of languages (Gast, 2013). CA is based on assumptions that 1) learners' first and target languages hold structural similarities and differences, and 2) the similarities facilitate learning while differences cause problems (Rustipa, 2011). The purpose is to pinpoint the differences between learners' L1 and L2, predict the problems learners may encounter, and design efficient teaching materials for language programmes (Gast, 2013).

It is gradually realized, however, the predictions made by CA seem not to be borne out in practice (Mitchell & Myles, 2004). Not all problems predicted by CA always appear to be difficult for learners; on the other hand, many predicted errors that do turn up are not predicted by CA (Rustipa, 2011). This leads to a change of emphasis from the comparison between learners' L1 and L2 to investigating the nature of the language produced by learners in its own right (Selinker, 1972; Myles, 2015).

With the emergence of the field of corpus linguistics, large data sets have been used in the study of language, and research on learner language has benefited from this development. "There is nothing new in the idea of collecting learner data. Both FLT and SLA researchers have been collecting learner output for descriptive and/or theory building purposes since the discipline emerged" (Granger, 2004, pp. 123-124). In this case, a learner corpus seems ideally to be suited to the investigation of L2 development.

Gass and Selinker (2008, p. 173) observe from L2 studies that "language learning is largely lexical learning". Corpus linguistics allows researchers to study how a word is used and explore what truly occurs in language (Sinclair, 1991; Hunston, 2002; Mahlberg, 2005), as it enables quick and efficient searches for words or strings of words in a large amount of data. Generally, the techniques from this field can be used to reveal and quantify recurrent patterns of learner productions (Barlow, 2005).

As Sinclair (1998, p. iv) suggests, a "pattern" is a way of "representing the regularities of usage of a word". Informed by Sinclair, Hunston and Francis (1999) propose a new approach Pattern Grammar to language description, and suggest using "pattern" as a descriptive tool to investigate the behaviour of individual words. Following Hunston and Francis (1999), this study examines how a word is used by L2 learners over time,

based on the observation on the changes of recurrent patterns of certain words across time.

1.2 Contextual information

English is the second language in Malaysia, and it is also considered one of the most widely used languages in the region (Yunus, Sulaiman, Kamarulzaman, & Ishak, 2013). In general, Malaysian students would learn English for eleven years in the primary and secondary schools (Darus & Subramaniam, 2009). There is a necessity to investigate how Malaysian students develop their L2 competence over time.

A learner corpus, Longitudinal Corpus of Languaculturer Narrative Texts (LoCLaNT) (Chau, 2015), is chosen for this research. As the name implies, this corpus provides longitudinal data which can afford the strongest evidence of language development (Ellis, 2008). It is well known that following and observing the same group of learners over time as they develop their L2 capacity is an ideal way for researchers to track language development (Ortega & Byrnes, 2008). A subset of the LoCLaNT data is taken to form a database for this study (see Section 3.1 later for more discussion).

This study takes nouns as a starting point to explore how learner language develops over time. Nouns are one of the most crucial elements of a text, and play an important role in building up the text (Hmouma, 2014), and make texts cohesive (Mahlberg, 2005). Further, nouns are proven to be easiest for children to learn in bilingual acquisition (Snedeker, Geren & Shafto, 2012), which may provide more evidence on language use.

For this study, the top five high frequency nouns, together with their respective singular or plural form in the subset of LoCLaNT, are selected. These are high frequency words,

which means a relatively large amount of concordance lines of the words would be available for the purpose of data analysis.

1.3 Problem statement

It is a goal, as noted earlier, for many researchers to describe how people learn a second language (L2) over time (Ellis, 2008). One common way to achieve that is to compare L2 productions with the norms by native speakers, which leads to a bias researchers held towards both the learner language and the language learners. Learner language is considered as an inferior version of L1 (Mamelina, 2013), and L2 learners are viewed as failed monolinguals (Cook, 2013).

The linguistic competence of bi/multilinguals is fundamentally different from that of monolinguals (Cook, 2013). Thus, it is not fair to judge L2 productions by the monolingual standards. This study respects the L2 developing system in its own right (Chau, 2015), and will not compare L2 productions with L1. To investigate language development, instead, this study will focus on the change of language use by students over time.

1.4 Research Objective

This research aims to track L2 development through the study of the use of nouns in the subset of LoCLaNT. It intends to answer the following two questions:

- 1) What are the recurrent patterns of use of the selected nouns used by Malaysian secondary school students at four different points in time in the subset of LoCLaNT?

To answer the first research question, the texts from Time 1 were used as the base to make a frequency list through the word list tool of AntConc. Top five frequency nouns were chosen for the investigation. The concordance lines of these words were generated via the concordancer, and then the relevant patterns were identified according to emergence criteria (See section 2.3.2).

2) How do the patterns of use of these nouns change over time in the subset of LoCLaNT?

To answer the second research question, comparison of use of each word was made among different data sets. Both patterns and instances under each pattern were identified and compared.

1.5 Scope and limitations

There are three aspects to be noted in this research. First of all, the corpus to be used in this study is only part of the LoCLaNT data. This study considers 200 narrative texts written by 50 Malaysian secondary school students, from the larger group of 124 students who contributed their texts to the development of LoCLaNT (see Chapter 3 for more details).

Secondly, this research focuses on the use of top five high frequency nouns in the subset of LoCLaNT. Five is admittedly an arbitrary number, and future research might establish whether and to what extent it is the number of words or quality of observation (e.g., in terms of details) that should drive the scope of analysis for a dissertation.

Lastly, the research takes noun lemmas instead of noun forms as units of analysis. A lemma is comprised of a headword and its inflected and reduced forms (Nation, 2001), such as its plural. Sinclair (1991) argues that each word form has its own individual patterning, so using lemmas may cause the loss of important meaning. However, using lemmas can be a helpful option (Granger & Paquot, 2009). Future research could focus on word forms, which might lead to different results.

1.6 Significance of the study

This study aims to investigate how the use of selected nouns changes over time. The linguistic feature, the noun, has so far received little attention in the field of SLA research. Hence, the findings from this corpus study provide empirical evidence of how learners use nouns over time. Meanwhile, longitudinal studies, especially those with data collected over a number of years, are rather rare in the field. The study helps scholars and teachers develop a better understanding of how learner language development takes place over a period of 24 months.

The study is based on the Pattern Grammar approach (Hunston & Francis, 1999) which was originally designed to study language produced by native speakers. The application of this approach to the investigation of learner language provides a new perspective of how L2 develops over time.

Further, this study is in line with the view that learner language is a system in its own right (Selinker, 1972; Cook, 2010). The results provide evidence that language learning is a non-linear and complex process (e.g., Larsen-Freeman, 2006, 2009; Ortega, 2009; Ellis, 2011; Chau, 2015). It may help scholars and teachers to look at 'learner language' and 'language learner' in a more respectful way (see Chau, 2015).

1.7 Conclusion

This chapter provides an overview of the present research. The following chapters are to contextualize the inquiry and investigation, that is, to give the theoretical and methodological background, present the results, and discuss the implications of this study.

Specifically, Chapter 2 presents a literature review on learner language and the studies of nouns. Chapter 3 describes the methodology of this study, including the details of the corpus and also the tools and techniques used for data analysis. Chapter 4 presents the results for the investigation of patterns of use of nouns in the subset of LoCLaNT. The final chapter will conclude the dissertation by summarizing and discussing the main findings and providing implications for future research and language teaching.

CHAPTER 2

LITERATURE REVIEW

This chapter reviews previous studies on learner language and English nouns from the perspectives of SLA and corpus linguistics. First, Section 2.1 provides an account of the views on learner language over time. Second, Section 2.2 presents the methods of analysing learner language in the field of SLA. Then, Section 2.3 focuses on the methods and theories from corpus linguistics. Last, Section 2.4 gives an overview of previous studies on English nouns.

2.1 Learner language

The study of learner language as “bad language” in language learning and teaching has a long history (Ellis, 2008, p. 45). Learner language is perceived as an incorrect and deficient version of the target language (Mamelina, 2013), and learner errors are regarded as setbacks in the language development (Hendrickson, 1978; Kalmos, 2011). Numerous publications have been concerned with learner errors, for instance, *Common Mistakes in English* (Fitikides, 1989) and *Common Errors in English Usage* (Brians, 2008).

The perspective on learner language and errors has changed considerably since 1960s. Corder (1967) argues that errors made by learners can provide evidence of a language system that a learner is having at a particular time in a learning process. He (1981) further states that learner errors are an integral part of language development. Errors then are conceptualized as steps in the language learning process (Edge, 1989) and come to be treated as learners’ attempts at using new structures, thereby improving their language competence (Kalmos, 2011). Chau (2015) maintains that instances of ‘errors’

should be treated equally as valid features of language development. It means the so-called ‘errors’ are as important as correct usages in terms of learner language development. Errors are gradually transformed from harmful influences into natural occurrence and valuable resources.

Selinker (1972) stresses that learners’ developing L2 competence is a language system in its own right, and not merely a defective copy of the native speakers’. Cook (2013) further suggests to look at learner language from a multicompetence perspective, that is, people who speak more than one language should not be treated as a deficient monolingual.

This study treats learner language as a system in its own right. Following Chau (2015, p. 3), the so-called ‘errors’ are considered as instances of “innovative language use” in this research, while the error-free usages are termed as instances of “conventional language use”.

2.2 Analysing learner language: methods from L2 acquisition

There are a number of different methods researchers adopt to investigate learner language development.

One of the first approaches is Error Analysis (EA). EA seeks to identify and describe different types of errors in order to understand how learners develop their language competence. There are five steps to carry out an EA, that is, 1) to collect a sample of learner language, and then 2) to identify, 3) to describe, 4) to explain, 5) to evaluate the errors (Corder, 1974; Zhao, 2015). This method achieved considerable popularity in the 1970s, while the interest in EA declined afterwards. It fails to offer a complete picture

of learner language, since it only focuses on what a learner does wrong and ignores what a learner does right (Ellis, 2008). Also, it attempts to explain learner language with the target language as the sole standard, which leads to what Bley-Vroman (1983) calls the comparative fallacy.

Another common method is Obligatory Occasion Analysis. This method is to identify the obligatory occasions for the use of specific target language features in the samples of learner language (Ellis, 2008). In other words, the procedure is to count both the tokens of the target language use and the obligatory occasions, and then, the percentage of accurate use of the feature is calculated (Rif'ah & Sulisty, 2004).

The method Target-like Use Analysis proposed by Pica (1983) takes one more step than Obligatory Occasion Analysis. This method not only takes account of accurate use of the feature, but also the over-uses and misuses (Ellis, 2008).

All these methods from SLA are target-language-based. It means they set the target language as the norms and standards to explain the process of L2 learning, and overlook the fact that learners have their own unique language systems in the learning process, which run counter to the core idea the present study holds that L2 language is a system in its own right (Selinker, 1972).

2.3 Corpus linguistics

It is well believed that “language looks rather different when you look at a lot of it at once” (Sinclair, 1991, p. 100). The application of linguistic corpora can provide empirical basis for the inquiry of language use and assist to discover latent language features which are unnoticeable (Biber & Reppen, 1998; Alsagoff, 2016). Kennedy

(1998) stated that the domain of corpus linguistics research is the bodies of texts which are also the evidence for linguistic description and interpretation.

The word corpus, or corpora (in plural), derives from Latin for 'body' (Hardie, 2016). In modern linguistics, a corpus means "a collection of naturally occurring language text, chosen to characterize a state of variety of a language" (Sinclair, 1991, p. 171).

The development of corpus linguistics started in 1950s (Mahlberg, 2005). Pioneering work was done by Nelson Francis and Henry Kucera who built the first English modern corpus, the Brown University Standard Corpus of Present-Day American English (also known as the Brown Corpus) in 1960s. This corpus encompasses 500 samples of language texts with approximately one million words (Dash, 2005).

The studies of corpus linguistics have changed the way of interpreting language. Sinclair (1991) proposes two different principles to explain the way of the meaning which arises from language, that is, "the open-choice principle" and "the idiom principle" (pp. 109-110). Based on the former, language texts are regarded

As the result of a very large number of complex choices. At each point where a unit is completed (a word or phrase or clause), a large range of choice opens up and the only restraint is grammaticalness (Sinclair, 1991, p. 109)

The latter states that

A language user has available to him or her a large number of semi-preconstructed phrases that constitute single choices, even though they might appear to be analysable into segments (Sinclair, 1991, p. 110)

Sinclair (1991) points out that conventional grammars are likely to follow the open-choice principle, while the evidence from corpora suggests that idiom principle dominates the operating of language (Mahlberg, 2005).

Sinclair (1991) further argues that language comprises “units of meaning”, which indicate that meaning is not carried by a word in and of itself, but made through several words in a sequence. Combinations of words in texts reflect the “pattern” of co-selection that contributes to the creation of meaning (Sinclair, 1991; Mahlberg, 2005). Sinclair (1998) points out that a “pattern” can represent the regularities of the use of a word.

2.3.1 Learner corpora

Linguistic research has developed many types of corpora based on its purpose, for instance, specialised corpus, general corpus, and learner corpus (Hunston, 2002). In the field of SLA, learner corpora received increasing interest nowadays.

A learner corpus is an electronic collection of spoken or written production made by L2 learners (Granger, 1998; McEnery & Hardie, 2011), and the analysis of learner corpora involve methods and assumptions from corpus linguistics (Barlow, 2005). Learner corpora are compiled in order to gather objective data, thereby describing learner language (Granger, 1998), which can enhance the understanding of the development of second language learning process. The best known English learner corpus in the world is the International Corpus of Learner English (ICLE) which comprises several corpora with essays written by learners from French, German, etc.

In Malaysia, there are a number of learner corpora. Corpus Archive of Learner English in Sabah/Sarawak, also known as CALES, consists of argumentative essays collected from university students in three public universities in Sarawak and Sabah. The English of Malaysian School Students corpus (EMAS), built by a group of researchers at Universiti Putra Malaysia, involves both written and oral productions based on students' essays and interviews. The corpus adopted in this study as a part of LoCLaNT was a longitudinal corpus collected over a period of 24 months by Chau (2015). A series of studies have been reported based on this corpus in Chau (2015). He concentrates on the use of three function words, *that*, *to* and *of*, and the changing structure of selected individual texts in the narration-writing practice over time. What Chau (2015, p. 209) finds is that “complexity in the developing language is dynamic rather than static, thereby challenging the notion that language development is a process towards greater complexity in language use” .

2.3.2 Pattern Grammar: noun patterns

Hunston and Francis (1999) put forward a new approach Pattern Grammar to language description. This approach “maintains the generalizing characteristics of grammatical descriptions while prioritizing the behaviour of individual lexical items” (Hunston, 2002, p. 167). The “pattern” of a word is “all the words and structures which are regularly associated with the word and which contribute to its meaning” (Hunston & Francis, 1999), whereas “grammar” is a set of generalizations about the use of words in texts (Mahlberg, 2005). A “pattern” can be seen as a language phenomenon which builds a connection between lexis and grammar (Hunston & Francis, 1999; Li, 2015).

A pattern is defined by Hunston (2002) as “a sequence of grammar words, word types or clause types which co-occur with a given lexical item” (p. 169), from which a similar definition of a pattern in the current study derives. One difference, however, is that the term “pattern” is used more broadly in this research than Hunston and Francis (1999), and, hence, more sequences which describe lexicogrammatical occurrence of a given word would be included. This is largely because of the different priorities in each research. For Hunston and Francis (1999), the aim is to describe English language with typical noun patterns. It means they only consider the patterns which are typical or special to a particular noun, and the words and structures in each pattern really contribute to the meaning of the noun. That is to say, they deal with the word-governing situation. However, the present study is concerned with learner language development, and aims to provide a picture of how words used by students in patterns develop or change over time. Due to different purposes in these two studies, more instances of patterns are considered in this study than in Hunston and Francis (1999). One example would be N which. Obviously, clauses with *which* can be placed after any nouns, which means it is not strictly a word-governing structure. Hence, Hunston and Francis (1999) did not take into account such language uses. Nevertheless, they are included in this study to illustrate how learners communicate their message through wh-clauses rather than strictly following a pattern grammar approach like in Hunston and Francis (1999).

Regardless of the difference, this work is informed by Pattern Grammar in the sense that most patterns observed in this study are identified with reference to Hunston and Francis (1999).

In some cases, moreover, the interpretation of a sequence as a pattern or not is much more subjective and open to debate (Hunston & Francis, 1999). For instance, the prepositional phrase *near the lake* as in *He saw a girl near the lake* (*girl* here is the focus) can be interpreted as a post-modifier of *girl*, or an adverbial to the whole sentence. Hence, it may be considered as a pattern **N *near* n** (*a girl near the lake*) for some researchers, but not for other researchers. For this study, cases like this would be considered as a sequence of a pattern.

To determine whether a structure or a pattern has emerged, emergence criteria are established to set a minimum amount of evidence. Following Mackey (1999), this study adopts the criterion by Pienemann and Johnston (1987) which illustrates a feature is considered as a structure or a pattern, when it is used at least twice in the learner production. For the current study, more stringent criteria are imposed, that is, a sequence is considered a pattern, when there are at least two instances of the use of that pattern in at least two different learners' productions (following Chau, 2015, p. 36).

This study adopts the coding system by Hunston and Francis (1999) that involves a number of abbreviations, for instance, n (noun group) and adj (adjective group), and the elements in each pattern are set out in the order in which they occur. A list of the abbreviations is given in Appendix A. The identification of the types of word classes is based on three grammar books: *A Comprehensive Grammar of the English Language* (Quirk, Greenbaum, Leech & Svartvik, 1985), *Longman Grammar of Spoken and Written English* (Biber, Johansson, Leech, Conrad & Finegan, 1999), and *Structure and Meaning in English: A Guide for Teachers* (Kennedy, 2003).

For instance, the coding ‘**N *near* n**’ represents a pattern, that is, a noun is followed by the preposition *near* and a noun group. The upper-case N indicates the word class whose pattern is focused on, italics mark specific words in contrast to coding symbols, in this case *near*, and n symbolizes the noun group. Thus, **N *near* n** represents, for instance, *the small river near our house* as in *najaa go to the small river near our house*. Another example of a pattern is ‘**N *which***’ which represents *the river which is near their houses* as in *They go to the river which is near their houses*.

2.4 Previous studies on nouns

Numerous studies have focused on nouns. Biber, Gray and Poonpon (2011) proposed a developmental progression index of noun phrase complexity. They found that nominal modification starts with adjectives; it progresses through more complex noun modification by nouns, prepositional phrases with concrete meaning and relative clauses; further progression involves nouns modified by -ed and -ing clauses as well as prepositional phrases with abstract meaning; final stage includes appositive noun phrases as noun modifiers as well as complement clauses.

Parkinson and Musgrave (2014), using Biber, Gray and Poonpon’s (2011) model on noun phrase complexity, examines the nouns in a corpus of 37 essays written by L2 students. The finding shows that students with lower proficiency used more attributive adjectives than students with higher proficiency. In addition, the use of noun modifiers by higher proficiency students was much closer to published frequencies for academic prose than by lower proficiency students.

Studies mentioned above are concerned with how noun phrases develop over time, while the following research is more relevant to the present work. Pioneering works on noun patterns are done by Hornby (1954). Four noun patterns are revealed in his work, that is, **N to-inf**, **N prep n**, **N that**, and **N prep conjunctive clause**. A more comprehensive list of noun patterns is done by Francis, Hunston, and Manning (1998), based on the COBUILD project (see Appendix B). It is noted that these two pieces of work, although focusing on noun patterns, are based on language produced by English native speakers. Nevertheless, the noun patterns provide a picture of what may be expected in learner language.

Mousavi and Rauof Moini (2014) studied the relationship between functions and lexico-grammatical patterns of 18 high frequency shell nouns (e.g. *change*, *process*, *form* and *characterization*) in 239 educational articles. They identified six patterns in their corpus: **N of n**, **N that**, **N verb**, **this/that N**, **N in which/by which**, **N n**. The most frequent pattern was **N of n** which indicates the node word is followed by a prepositional phrase introduced by *of* (as in *a process of change*), and the function of this pattern was to provide the details of information. Pattern **N that** which indicates the node word is followed by clauses with *that* occupies the second place, and shows a cataphoric function. Pattern **N in which/by which** indicates the node word is followed by clauses introduced by *in which* and *by which*, with a function of characterization. Pattern **that/this N** indicates the node word is preceded by determiners *that* and *this*, and shows a function of linking. Pattern **N n** indicates that the node word is modified by another noun, and has a characterization function.

Martin (2009) explored native speakers' lexico-grammatical patterns around an abstract noun contribution in medical English. He found that the use of the abstract noun contribution tends to be associated with specific grammatical patterns which are **v N**, **adj N**, **N n**, and **N prep**.

These studies use lexico-grammatical patterns as a tool to investigate the use of nouns, but none of them are concerned with how the use of certain nouns changes over time in learner data. The present study aims to find out whether and to what extent the use of high frequency nouns would change over time by 50 students considered in this study.

2.5 Conclusion

This chapter has provided the theoretical background for the present research on the investigation of the L2 development. Section 2.1 discussed the changing perspectives researchers held towards both the learner language and language learners over time. Section 2.2 introduced the methods in the field of SLA of analysing learner data. Section 2.3 presented the theories and methods from corpus linguistics. Section 2.4 dealt with the previous studies on nouns and noun patterns.

CHAPTER 3

METHODOLOGY

This chapter deals with the methodological details of the current research. Section 3.1 provides information on the corpus used for this study. Section 3.2 discusses the procedures of data analysis, which includes the tools used for this research. Lastly, Section 3.3 concludes this chapter.

3.1 Corpus used in this research

As noted earlier, LoCLaNT is adopted for this research. This corpus is developed by Chau (2015) with a purpose to explore learner language development over time. It consists of 496 narrative texts written by the 124 students (77 Female/47 Male) in a period of 24 months. The participants, 13-year-old at the first data collection, were English learners with Malay being their mother tongue. Data collections were carried out at four different points in time, that is, May 2007, November 2007, November 2008 and June 2009. The prompt used for the writing is a series of pictures about a near-drowning accident. It shows that a girl picks flowers with her friend, falls into the river, and finally is saved by three boys (see Appendix C).

As noted in Chapter 1, data from 50 students of LoCLaNT were used in this study. This dataset consists of 200 texts (43,790 words), with four texts contributed by each student. The texts were categorized into four sub-sets labelled as 'Time 1', 'Time 2', 'Time 3',

and 'Time 4', according to the dates of the data collection. To identify each text, a coding system was adopted. The 'coding' range (from 001 to 050) reflected the number of 50 students. The four sub-sets of texts were further given the label of 'A', 'B', 'C' and 'D', respectively. For example, 001A refers to the text written by the student ('1') in May 2007 (or Time 1). 006C refers to the text written by the student ('6') in November 2008 (or Time 3).

3.2 Tools for analysis

The software, AntConc by Laurence Anthony, Version 3.4.3, 2014, was adopted to analyze data. It is a freeware toolkit for corpus analysis, which provides a set of tools, such as, a concordancer, a word list generator, and a word distribution plot.

To answer research questions of this study, the concordancer and the word list generator are adopted.

3.2.1 The concordancer

The concordance lines are the lines which can provide information on the company the key word keeps in a corpus. A concordancer can produce and display all the occurrences of a particular word (or a keyword, or a node word) in a corpus. It can be generated in different formats. The form that AntConc produces is the Key Word in Context (KWIC) Concordance which is the most usual form. Figure 3.1 is an example

of a KWIC concordance for the word *river* from the subset of LoCLaNT. The keyword *river* is displayed in the center of the lines.

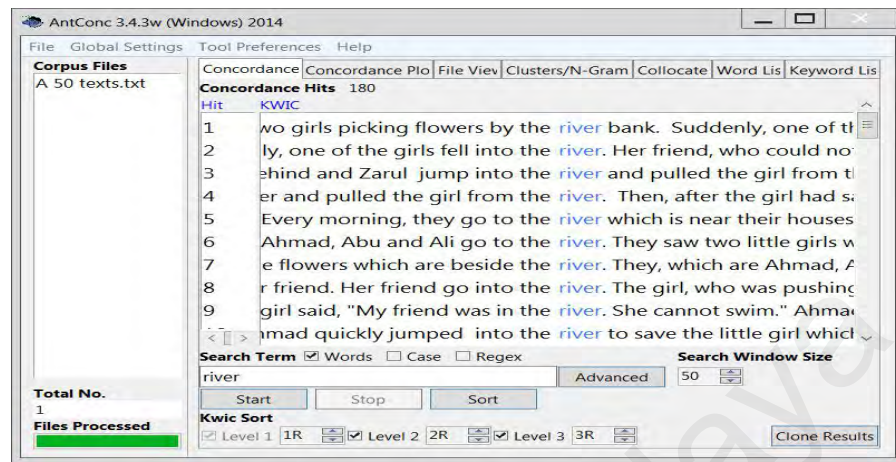


Figure 3.1: A fragment from concordances of *river* in the subset of LoCLaNT

In order to get a full picture of the environments of a node word, the concordance lines can be sorted to the left or right of the keyword as in Figure 3.2 and 3.3.

In Figure 3.2, the context to the left of the node word *river* in each line has been the basis for the sorting of the whole line by a simple alphabetical ordering. Consequently, recurring word sequences show up together.

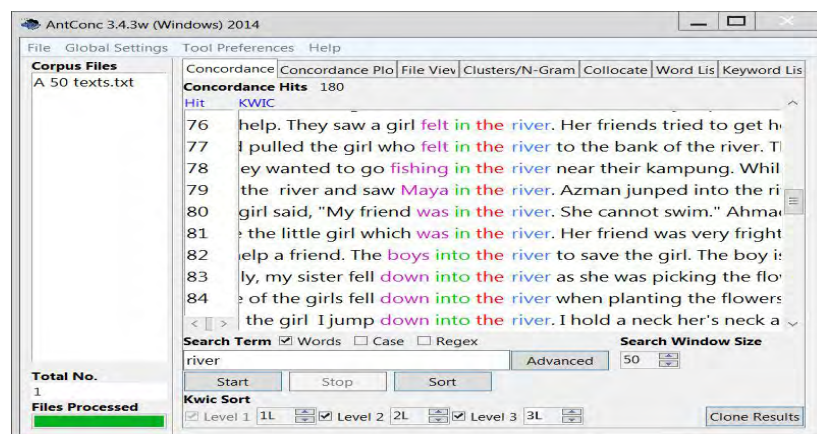


Figure 3.2: A fragment from concordances of *river* to its left

The context to the left of the node word is the basis for the sorting which results in the format in Figure 3.3.



Figure 3.3: A fragment from concordances of *river* to its right

3.2.2 The word list tool

The word list generator can sort the words into frequency order, which is helpful to highlight the focused areas in a corpus and suggest problem areas (see Figure 3.4).

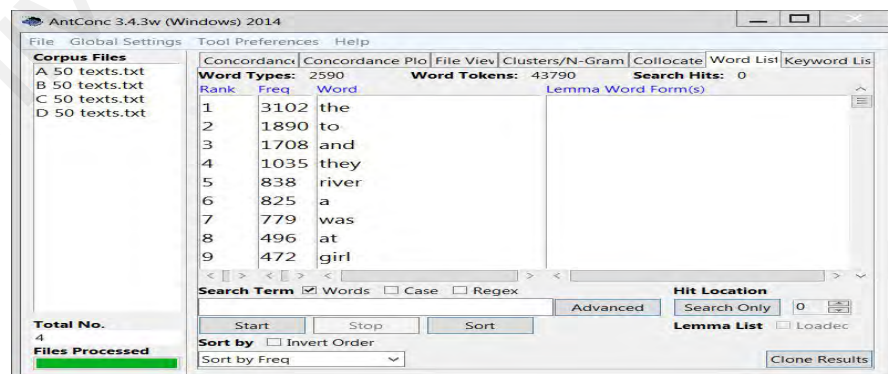


Figure 3.4: A frequency list of words from the subset of LoCLaNT

In order to answer the research question for this study, the top five high frequency nouns were first identified. Table 3.1 presents the result of top 10 high frequency nouns at Time 1. Words *river*, *girl*, *flowers*, *lake*, and *friend* occupy the first five places. These are also the five words, together with their respective singular or plural form, which form the focus of this study. Language development is examined in this study through these 5 high-frequency words and how they change over time in the 24-months period.

Table 3.1: Top 10 nouns at Time 1

Rank	Freq.	Noun
1	180	river
2	113	girl
3	65	flowers
4	63	lake
5	60	friend
6	58	friends
7	56	girls
8	51	boys
9	32	parents
10	29	flower

Regarding to the different forms of the key words, this study follows the tradition of corpus linguistics in which when a word is spelt in upper case (e.g. RIVER), it refers to the idea of lemma which consists of both singular and plural forms of the word (e.g. river, rivers); when a word is spelt in italics and in lower case without the ending -s (e.g. *river*), it refers to the singular form of the word (e.g. river); and when a word is spelt in italics and in lower case with the ending -s (e.g. *rivers*), it refers to the plural form of the word (e.g. rivers).

3.3 Data analysis

In general, the study was both quantitative and qualitative in nature. The aim was to find out the change of use of five specific words in the essays written by the same group of students over a period of 24 months.

In order to answer the first research question, “What are the recurrent patterns of use of the selected nouns used by Malaysian secondary school students at four different points in time in the subset of LoCLaNT?”, the concordancer was used to produce all the concordance lines of the top five high frequency words in both singular and plural forms. The concordance lines then were sorted to the left or right of the node by a simple alphabetical ordering. Recurring word sequences were identified according to the emergence criteria.

In terms of the second research question, “How do the patterns of use of these nouns used by Malaysian secondary school students change over time in the subset of LoCLaNT? ”, the comparison on the patterns of the word from different data sets was made. The types of patterns of the word in each data set were identified and compared, as well as the instances under each pattern in different data sets. The emergent lexis which was found to occur together with the high- frequency words in focus at Time 1, Time 2, Time 3 and Time 4 was identified to see how students change the the language use over time.

3.4 Conclusion

This chapter has shown the methodology of this work with the discussion of the adopted corpus and the research techniques for data analysis. It is seen that this study includes two steps. The first step concerns the identification of recurrent patterns of the selected nouns at four different points in time in the subset of LoCLaNT. The second step focuses on the comparison of the patterns of use of the nouns over time. During these two stages, two research tools, such as the word list generator and the concordance, are used to facilitate the investigation.

University of Malaysia

CHAPTER 4

RESULTS

As indicated in Chapter 3, the nouns *river*, *girl*, *flowers*, *lake*, and *friend* occupy the top five places in the data. This chapter reports the findings of these five words with their respective singular or plural form in separate sections. Each section is organized into two parts. Firstly, the patterns of use of these words in different data sets will be presented and briefly compared. Secondly, the instances under each pattern of use of these selected nouns will be analysed.

4.1 RIVER

A total of 845 instances of RIVER were identified in the learner data: 181 instances (sing.:180/pl.:1) at Time 1, 207 instances (sing.:204/pl.:3) at Time 2, 228 instances (sing.:226/pl.:2) at Time 3, and 229 (sing.:228/pl.:1) instances at Time 4. To control length variation and make a possible comparison between the corpora of unequal sizes, normalized frequency (Biber et al., 1998) which reflects relative frequency is calculated to see how the word distributed in different data sets. Counts are normed to a basis of 10,000 words since four data sets are about this long. Take Time 1 for example. The normalized frequency is 209 instances per 10,000 words $[(181 \text{ instances} / 8627 \text{ words}) \times 10,000 = 209 \text{ instances per } 10,000 \text{ words}]$. The raw and normalized frequency of use of RIVER across the data is summarized in Table 4.1.

Table 4.1: Raw and normalized frequency of use of RIVER

	Time 1	Time 2	Time 3	Time 4
Size of sub-corpus (tokens)	8627	11205	11649	12309
Raw frequency	181	207	228	229
Normalized frequency	210	185	196	186

201 instances of RIVER per 10,000 words are found at Time 1, 185 at Time 2, 196 at Time 3, and 186 at Time 4. Generally, the normalized frequencies descend across time, although the data sizes show an ascending trend.

To sum up, a total of 845 concordance lines of RIVER were identified in the data. The normalized frequency of the word in each data set shows a decrease over time. Concordance lines then were examined to study the patterns of use of RIVER.

4.1.1 Patterns of use of RIVER

A total of five patterns of use of RIVER are observed to occur in the data. They are **N n**, **det N**, **adj N**, **N prep n**, and **N clause** (Table 4.2). Pattern **N n** indicates that the noun RIVER is followed by another noun (as in *the river water* and *the river current*). Pattern **det N** indicates that the noun RIVER is preceded by determiners (as in *the rivers* and *a river*). Pattern **adj N** indicates that the noun RIVER is preceded by adjectives (as in *the small river* and *the long river*). Pattern **N prep n** indicates that the noun RIVER is followed by prepositional phrases (as in *the river near their house* and *the river behind their house*). Pattern **N clause** indicates that the noun RIVER is followed by clauses (as in *the river which is located near their house* and *a river not far from our house*).

Table 4.2: Patterns of use of RIVER

Patterns	Time 1		Time 2		Time 3		Time 4	
	Freq	%	Freq	%	Freq	%	Freq	%
det N	148	91%	174	93%	202	96%	181	88%
N n	19	10%	20	10%	18	8%	23	10%
adj N	9	6%	7	4%	4	2%	17	8%
N prep n	10	6%	13	7%	9	4%	13	6%
N clause	-		5	3%	2	1%	4	2%

The number of patterns of use of RIVER suggests a slight increase over time. Four patterns (**N n**, **det N**, **adj N**, and **N prep n**) occur at Time 1, while five patterns (**N n**, **det N**, **adj N**, **N prep n** and **N clause**) appear at last three data sets. Pattern **N clause** is only found to be used at later period in time, which may indicate that students started to use more different structures to modify the word RIVER over time.

It is noticed that the most frequent pattern is **det N**, with a percentage of over 88% at each data set. Pattern **N n** comes to the second place, with a percentage of over 8%. The frequency of the rest of the three patterns (**adj N**, **N prep n** and **N clause**) is far behind with a percentage of less than 8% across the corpus (see Section 4.1.2.1 for examples of each of these patterns). In terms of each pattern, there is no constant increase or decrease in the percentage to be found. It is thus necessary to zoom in on each pattern, and find the change and difference of use of patterns at different time.

4.1.2 Observation on instances under each pattern

As mentioned earlier, five patterns of use of RIVER, i.e., **N n**, **det N**, **adj N**, **N prep n**, and **N clause** appear in the data. A close look is then made on the instances under each pattern of use of the word. Examples of the use of RIVER in the learner production are given. In addition, the type of function words (such as determiners, adjective, prepositions and clauses in patterns **det N**, **adj N**, **N prep n**, and **N clause**) is discussed in Section 4.1.2.1. Meanwhile, the variety of content words (like nouns and adjectives in patterns **N n**, **N prep n**, and **adj N**) is considered in Section 4.1.2.2.

4.1.2.1 Instances under each pattern of use of RIVER

4.1.2.1.1 det N

When RIVER is used in pattern **det N**, two types of determiners are identified, i.e., *the* and *a* (Table 4.3). The definite article *the* is the most frequent determiner in the data, which accounts for over 90% of the total usage of the determiners. On the other hand, the indefinite article *a* is far less used by students, with a percentage of 10% or less in the learner data.

Table 4.3: Determiners of the word RIVER

Types	Time 1		Time 2		Time 3		Time 4	
	Freq	%	Freq	%	Freq	%	Freq	%
the	137	93%	163	94%	182	90%	171	93%
a	11	7%	11	6%	20	10%	12	7%

Examples of the definite article *the* from students' texts are as follow:

Time 1

(1) My friend was in *the river*. (004A)

Suddenly, while plucking flowers by the river, Aina fell into *the river*. (010A)

Suddenly, while they were collecting flowers Nina fell into the river. (031A)

Time 2

(2) One in the boys fall in *the river* for save the Ana's life. (014B)

Last Saturday, Samad and his friends went to *the river*. (018B)

Syariff, who is a good swimmer, jumped into *the river*. (051B)

Time 3

(3) They saw one of the girls had slipped into *the river*. (004C)

Last Sunday, I, Ajis and Ramlee went a fishing at *the river*. (030C)

At the same time, Raju and his friends want to go fishing at *the river*. (036C)

Time 4

(4) Zulaikha was drowning in *the river*. (004D)

They was plan to fishing at *the river* near the Amer grandfather's house. (016D)

They rushed to that place and Ali quickly jumped into *the river*. (023D)

Examples of the indefinite article *a* from students' texts are as follows:

Time 1

(5) Last Sunday, my sister and I went to *a river* to collect some flowers. (008A)

Last Saturday, they went to fishing at *a river*. (029A)

Suddenly, when she pulled the flowers, she feld into *a river*. (035A)

Time 2

(6) I went to fishing at *a river*. (023B)

Last holiday, I and my friends decided to went a fishing at *a river*. (030B)

Last weekend, Harry, Potter and I went to *a river* to fishing. (037B)

Time 3

(7) We walk by *a river* to our fishing spot. (003C)

There was *a river* near their house. (011C)

Last Sunday, they went to *a river* behind Ali house. (034C)

Time 4

(8) Near their houses there was *a river*. (011D)

I and my friends were planed to go for fishing at *a river*. (027D)

Aiman and his friends, Azri and Syafiq went for fishing at a river. (048D)

Two types of determiners are observed to occur with the word RIVER, that is, the definite article *the* and the indefinite article *a*. The definite article *the* takes up over 90% of the total usage, while the indefinite article around 10%.

4.1.2.1.2 adj N

When RIVER is used in pattern **adj N**, three grammar books, as indicated in Chapter 2, are applied to differentiate the types of adjectives. Three types of adjectives are observed to occur in the corpus, which are one-word adjectives (e.g., *muddy*, *small*, and

long), phrasal adjectives (e.g., *beautiful and clean* and *long and deep*), and adjective compounds (e.g., *fast flowing* and *fast-flowing*) (Table 4.4).

Table 4.4: Adjectives of the word RIVER

Types	Time 1		Time 2		Time 3		Time 4	
	Freq	%	Freq	%	Freq	%	Freq	%
One-word adjectives	9	100%	7	100%	4	100%	9	53%
Phrasal adjectives	-		-		-		6	35%
Adjective compounds	-		-		-		2	12%

Only one type of adjectives is observed to occur at the first three data sets, whereas three types of adjectives are observed at the last data set. At the first three data sets, one-word adjectives make up 100% of the use of adjectives. When it comes to the last data set, two more types of adjectives (phrasal adjectives and adjective compounds) appear, and take up 47% of the use.

Examples of one-word adjectives from students' texts are shown from (9) to (12), examples of phrasal adjectives at Time 4 are presented in (13), and examples of adjective compounds at Time 4 are shown in (14).

Time 1

(9) I and my friend Najaa go to the *small river* near our house. (020A)

It was a *big river*. (023A)

The girl was getting drown into the *deep river*. (049A)

Time 2

(10) Jasmine has slipped down from the slippery ground to the *muddy river*. (038B)

Sufian and Razak went fishing at a *small river* behind Sufian's house. (051B)

While Minah plucked the flowers, Minah fall into the *deep river*. (052B)

Time 3

(11) Riku and Risa followed their brother and his frineds to a *nearby river*. (010C)

So, they planed to go to a *near river* and fishing there. (041C)

Suddenly, Aminah fall into the *deep river*. (052C)

Time 4

(12) Last weekend, Rahim, Lan and I went for fishing at a *nearby river*. (026D)

There were also a *natural river* that very beautiful. (037D)

We are at the *peaceful river* and now we can see two beautiful princess. (050D)

(13) They walked along the *long and deep river* to find a place for fishing. (028D)

Torres and his friends wanted to go to a *beautiful and clean river*. (041D)

Amin dived in the *cold and cloudy river* to save the drowned girl. (049D)

(14) One of them slipped and felt down into the *fast flowing river*. (025D)

Basyir finished fishing at a *fast-flowing river* near their houses. (038D)

One-word adjectives play a dominant role in the modification of the word in focus, especially in the earlier period of time. As time goes by, two more different types of adjectives (i.e., phrasal adjectives and adjective compounds) are observed to occur. It

suggests the students have started to pay more attention to describing the characters or objects in the story.

4.1.2.1.3 N prep n

When RIVER is used in pattern **N prep n**, five prepositional phrases are observed to occur, as shown in Table 4.5. Two prepositional phrases are observed at Time 1, Time 3 and Time 4, while three are observed to occur at Time 2. It is found that **N near n** is the only pattern, and also the most frequent pattern used across the four data sets.

Table 4.5: Prepositional phrases of the word RIVER

Types	Time 1		Time 2		Time 3		Time 4	
	Freq	%	Freq	%	Freq	%	Freq	%
N near n	7	70%	7	55%	4	45%	8	62%
N in n	2	20%	-		-		-	
N behind n	-		2	15%	3	33%	-	
N at n	-		-		-		2	15%
N nearby n	-		2	15%	-		-	

Two prepositional phrases are identified at Time 1, that is, *near* (15) with a percentage of 70% and *in* (16) with a percentage of 20%.

(15) One day, they went for a fishing at the *river near their house*. (005A)

Hafizah and Farhana are gong to the *river near her house*. (016A)

Adnan and Yusri went to a *river near the Syamsir's house* to catch fish. (051A)

(16) Azriff ... went to fishing at the *river in Kampung Semberong*. (028A)

Last Weekend, Nazmi, Nice and I went to the *river in Taman Molek*. (037A)

Three prepositional phrases are identified at Time 2, which are *near* (17) with a percentage of 55%, *behind* (18) with a percentage of 15% and *nearby* (19) with a percentage of 15%.

(17) One day, they went fishing by a *river near their houses*. (003B)

Saiful and Zainal went to fishing at a *river near their houses*. (004B)

Last week, Ina and Ana went to the *river near the grandfather's house*. (014B)

(18) Last Sunday, they went to a *river behind Ali house*. (034B)

Sufian and Razak went fishing at a small *river behind Sufian's house*. (051B)

(19) On Saturday, Ani and Ina go to the *river nearby Ani home*. (040B)

Abu , Ali and I went to Bethong *river nearby my house* to fishing. (050B)

Two prepositional phrases are identified at Time 3, which are *near* (20) with a percentage of 45% and *behind* (21) with a percentage of 33%.

(20) There was a *river near their house*. (011C)

Ahmad and Abu went for fishing at the *river near their village*. (025C)

They always fishing at the *river near their village*. (049C)

(21) One day, they went for a fishing at the *river behind their house*. (005C)

Aiman and i were went to fishing at *river behind Aiman house*. (024C)

There was a *river behind grandfather's house*. (051C)

Two prepositional phrases are identified at Time 4, which are *near* (22) with a percentage of 62% and *at* (23) with a percentage of 15%.

(22) They were fishing at the *river near their house*. (005D)

Last week, they decided to fish in a *river near their village*. (031D)

Basyir were walking home after fishing at a *river near their houses...*(038D)

(23) ...Ali went fishing by a *river at the end of their village*. (003D)

They walked along the *river at the park*. (035D)

Five types of prepositional phrases are used to post-modify the word RIVER in the data. Two types of prepositional phrases are found to be used at Time 1, Time 3 and Time 4, while three types at Time 2. The fluctuation of the types of prepositional phrases indicates that language development is not necessarily a linear process.

4.1.2.1.4 N clause

When the word RIVER is used in pattern **N clause**, three types of clauses are identified in the corpus (Table 4.6), that is, *which*-clause, *that*-clause, and clause with adjective phrases.

From a long run, the number of clauses as post-modifiers of RIVER increases over time. No clause is observed at Time 1, two clauses observed at Time 2, one clause observed at Time 3 and three clauses observed at Time 4.

Table 4.6: Clauses of the word RIVER

Types	Time 1		Time 2		Time 3		Time 4	
	Freq	%	Freq	%	Freq	%	Freq	%
N <i>which</i>	-		3	60%	-		2	33%
N <i>that</i>	-		2	40%	-		2	33%
N <i>not far from</i>	-		-		2	100%	2	33%

Two types of clauses are identified at Time 2 with the introduction of *which* (24) with a percentage of 60% and *that* (25) with a percentage of 40%.

(24) Sazali invited friends fishing at a *river which is not far from their house*. (029B)

His friend went fishing at the *river which is located near their house*. (049B)

The walked to a *river which near by their house*. (041B)

(25) They went to the *river that is only 1 KM from their house*. (009B)

The *river that they want to go is near from their home*. (018B)

One type of clauses is identified at Time 3 with the node word followed by adjective phrases (26):

(26) Niezam and i will going to fishing at a *river not far from our house*. (023C)

Tajol an exciting to go fishing at the *river not far from their houses*. (034C)

Three types of clauses are identified at Time 4, which are introduced by *which* (27) with a percentage of 33%, by *that* (27) with a percentage of 33%, and followed by adjective phrases (29) with a percentage of 33%.

(27) We like to fishing is at *river which (is)situated far from my house.* (019D)

They decided to fish at the small *river which (is) full with fish.* (049D)

(28) At Kampung Baru had a long and deep *river that called Deed's River.* (028D)

There were also a natural *river that is very beautiful.* (037D)

(29) Ali and his frineds were fishing at a *river not far from their village.* (023D)

Kamal and his FRIENDs went to a *river not far from their house...* (034D)

Three types of clauses are identified in the students' texts, i.e., *which*-clause, *that*-clause, and clause with adjective phrases. No clause was found at Time 1; however, two types of clauses were found at Time 2, one type at Time 3 and three types at Time 4.

It is noticed that two types of determiners, three types of adjectives, five types of prepositional phrases and three types of clause are used to modify the word RIVER in the learner data. The number of types of modifiers increases or decreases as time goes by. We can see that the use of RIVER especially in patterns **adj N** and **N clause** becomes more diverse and complex over time.

4.1.2.2 Emergent lexis with RIVER

The nouns and adjectives in patterns **N n**, **N prep n**, and **adj N** are examined to see how the word RIVER is used with other lexis in learner data. Table 4.7 presents the results of lexis which occur at least twice in different students' texts. Three words (one adjective and two nouns) emerge at Time 1, while seven words (three adjectives and four nouns) are observed to occur at Time 4.

Table 4.7: Emergent lexical items with RIVER

	Time 1	Time 2	Time 3	Time 4
Adjectives	deep	deep	nearby	nearby deep big
Nouns	house bank	house bank side	house village bank	house village bank side

According to the frequency information (Table 4.8), we can see that 2% of the word RIVER is found to be modified at each of the first three data sets, while the percentage rises up to 3% at the last data set, which may suggest that learners' repertoire has expanded across time.

Table 4.8: Frequency information of lexical items with RIVER

	Time 1	Time 2	Time 3	Time 4
Frequency	3 2%	4 2%	4 2%	7 3%

Over time, more lexical items are found to be used with the word RIVER, which may indicate that students' lexical repertoire expands over time.

Through the observation of the use of RIVER, we can see that language development is a complex process. The number of patterns of use of RIVER shows a slight growth over time, and the instances under each pattern show a even more complex picture. The type of determiners used to modify the word RIVER remains the same; the number of prepositional phrases patterned with the word RIVER shows a fluctuation in the data; the type of clauses and adjectives to modify the word RIVER becomes more diverse over time; and the variety of adjectives and nouns patterned with the word RIVER gets increased over time.

4.2 GIRL

A total of 718 instances of GIRL are observed to occur in the learner data: 169 instances (sing.:113/pl.:56) at Time 1, 208 instances (sing.:145/pl.:63) at Time 2, 164 instances (sing.:92/pl.:72) at Time 3, and 177 instances (sing.:122/pl.:55) at Time 4. The raw and normalized frequency of use of GIRL across the data is summarized in Table 4.9.

Table 4.9: Raw and normalized frequency of use of GIRL

	Time 1	Time 2	Time 3	Time 4
Size of sub-corpus (tokens)	8627	11205	11649	12309
Raw frequency	169	208	164	177
Normalized frequency	196	186	141	144

196 instances of GIRL per 10,000 words are found at Time 1, 186 at Time 2, 141 at Time 3, and 144 at Time 4. As can be seen, the normalized frequencies across the data show a constant decrease across time, although the data size in each sub-corpus grows bigger.

A total of 718 concordance lines of GIRL were identified in the data. The normalized frequency of the word in each data set shows a decrease over time. All these concordance lines were examined to explore the patterns of use of GIRL.

4.2.1 Patterns of use of GIRL

A total of six patterns of use of GIRL are observed to occur in the data. They are **N n**, **poss-N n**, **det N**, **adj N**, **N prep n**, and **N clause** (Table 4.10). Pattern **N n** indicates that the noun GIRL is followed by another noun (as in *the girl parents* and *the girl voice*). Pattern **poss-N n** indicates the noun GIRL is used as a possessive (as in *the girl's voice* and *the girl's hand*). Pattern **det N** indicates that the noun GIRL is preceded by determiners (as in *that girl* and *another girl*). Pattern **adj N** indicates that the noun GIRL is preceded by adjectives (as in *the young girl* and *the little girl*). Pattern **N prep n** indicates that the noun GIRL is followed by prepositional phrases (as in *the girl at the shore* and *the girl near the river*). Pattern **N clause** indicates that the noun GIRL is followed by clauses (as in *the girl which was in the river* and *the girl who fell into the pool*).

Table 4.10: Patterns of use of GIRL

Patterns	Time 1		Time 2		Time 3		Time 4	
	Freq	%	Freq	%	Freq	%	Freq	%
det N	138	92%	160	94%	145	94%	146	97%
poss-N n	16	9%	12	6%	5	3%	14	8%
adj N	13	9%	7	4%	7	5%	11	7%
N clause	8	5%	7	4%	3	2%	4	3%
N prep n	4	3%	7	4%	3	2%	-	-
N n	3	2%	3	1%	4	2%	6	3%

The number of patterns of use of GIRL experiences a slight decrease over time. Six patterns (**N n**, **poss-N n**, **det N**, **adj N**, **N clause** and **N prep n**) are identified at first three data sets, while five patterns (**N n**, **det N**, **adj N**, **poss-N n** and **N clause**) at last data set. Pattern **N prep n** is absent from Time 4.

It is noticed that the most frequent pattern is **det N**, with a percentage of over 92% at each data set. The frequency of the rest patterns (**N n**, **poss-N n**, **adj N**, **N prep n** and **N clause**) is far behind with a percentage of less than 10% across the corpus (see Section 4.2.2.1 for examples of each of these patterns). In terms of each pattern, there is no constant increase or decrease in the percentage, except for pattern **det N**. It is thus necessary to zoom in on each pattern, and find the change and difference of use of patterns at different time.

4.2.2 Observation on instances under each pattern

Six patterns of use of GIRL are identified in the data, that is, **N n**, **poss-N n**, **det N**, **adj N**, **N prep n**, and **N clause**. Closer observation is then made on the instances under each pattern of use of the word. Examples of the use of GIRL in the learner production are given. In addition, the type of function words (such as determiners, prepositions and clauses in patterns **det N**, **N prep n**, and **N clause**) is concerned in Section 4.2.2.1. Meanwhile, the variety of content words (like nouns and adjectives in patterns **N n**, **poss-N n**, **N prep n**, and **adj N**) is considered in Section 4.2.2.2.

4.2.2.1 Instances under each pattern of use of GIRL

4.2.2.1.1 det N

When GIRL is used in pattern **det N**, eight types of determiners are applied by the students over time, as can be seen from Table 4.11. The determiners can be categorized into five groups which are articles, numerals, possessive determiners, demonstratives, and semi-determiners. The article, especially the definite article *the*, is the most frequent type of determiners in the data, which accounts for over half of the total usage of the determiners. Numerals come to the second place, with a percentage of over 14%. The rest three categories of determiners (possessive determiners, demonstratives, and semi-determiners) only account for 7% or less in each data set.

More specifically, the use of articles *the* and *a* and the numeral *two* is observed to occur across the corpus, the use of possessives is only observed at Time 2, the use of the

demonstrative *that* is only found at later point at Time, and the use of semi-determiners is present in three data sets, except for Time 3.

Table 4.11: Determiners of the word GIRL

Types	Categories	Time 1		Time 2		Time 3		Time 4	
		Freq	%	Freq	%	Freq	%	Freq	%
the	Articles	93	67%	108	61%	100	69%	100	68%
a		8	6%	13	7%	12	8%	3	2%
one	Numerals	-		5	3%	-		-	
two		22	16%	27	15%	23	16%	20	14%
their	Possessive	-		2	1%	-		-	
that	Demonstrative	-		-		3	2%	9	6%
another	Semi-determiners	2	1%	3	2%	-		-	
other		-		2	1%	-		2	1%

Four determiners are identified at Time 1, which are the definite article *the* (30) with a percentage of 67%, the indefinite article *a* (31) with a percentage of 6%, the numeral *two* (32) with a percentage of 16%, and the semi-determiner *another* (33) with a percentage of 1%.

(30) After a minute, *the girl* open her eyes. (005A)

Ahmad ran as fast as he could and saved *the girl* from drowned. (039A)

The girl has no injury. (052A)

(31) They rushed tere and saw *a girl* was felt in the pond. (026A)

Finally, *a girl* was safe. (028A)

A girl is into the river.(044A)

(32) He saw *two girls* were planting the flower. (028A)

They were *two girls* picking some flower at side of the lake. (033A)

There were *two girls* who were collect some flower at the river side. (048A)

(33) *Another girl* was thanked to Ali because saved her best friend. (032A)

Another girl also has three flowers. (041A)

Six determiners are identified at Time 2, which are the definite article *the* (34) with a percentage of 61%, the indefinite article *a* (35) with a percentage of 7%, the numeral *one* (36) with a percentage of 3%, the numeral *two* (37) with a percentage of 15%, the semi-determiner *another* (38) with a percentage of 2%, and the semi-determiner *other* (39) with a percentage of 2%.

(34) We asked *the girl* and her friend name. (003B)

They sent *the girl* to her home. (005B)

Chong Beng took *the girl* to the hospital ... (006B)

(35) He had saved *a girl* from accident. (004B)

After a few mintes, *a girl* was safe by Feroz. (028B)

Feroz and his friends send *a girl* at her house. (028B)

(36) At the home, Amran tells his family that he safe *one girl* at the river. (036B)

Ali felt so happy because he has safed *one girl*. (048B)

Suddenly, *one girl* fell into the river. (048B)

(37) ...they saw there was *two girls* were pick up some flower. (018B)

There were *two girls* picked the flower at side of the lake. (033B)

At the river, have *two girls* and three boys. (044B)

(38) Then, *other girl* called her friend's father. (006B)

The other girl were screamed for help...(046B)

(39) *Another girl* was crying. (004B)

Nazim was calming down *another girl* who was so frightened. (008B)

Another girl shouting for help. (048B)

Four determiners are identified from students' texts at Time 3, which are the definite article *the* (40) with a percentage of 69%, the indefinite article *a* (41) with a percentage of 8%, the numeral *two* (42) with a percentage of 16% and the demonstrative *that* (43) with a percentage of 2%.

(40) Firdaus without hesitation, jumped into the river to save *the girl*. (004C)

When they saw *the girl* that drowned in the river ...(006C)

The girl was fainted. (025C)

(41) Akaba, Shin and Sena was shocked when *a girl* was shouting for help. (006C)

On his way to Sungai Itik, *a girl* suddenly fell into the river. (007C)

It different for *a girl* because they love flower. (013C)

(42)...*two girls* were plucking some beautiful flowers near the river. (015C)

...*two girls* from Irfan's school were collecting a lot of flower, Ema and Jika. (016C)

At there, *two girls* were plucking some beautiful flowers at the riverside. (024C)

(43) Maybe the sound from *that girl*. (004C)

Then, Firdaus said to *that girl* must take care of herself next time. (004C)

I had self-sacrifice to saved *that girl* from get drown. (030C)

Five determiners are identified from students' texts at Time 4, which are the definite article *the* (44) with a percentage of 68%, the indefinite article *a* (45) with a percentage of 2%, the numeral *two* (46) with a percentage of 14%, the demonstrative *that* (47) with a percentage of 6%, and the semi-determiner *other* (48) with a percentage of 1%.

(44) Ameer and his friend bring *the girl* back to her house. (016D)

He tried to swim to *the girl* and grabed her to the river bank. (025D)

So they ran to *the girl*. (052D)

(45) He told that he heard *a girl* shouting for help. (015D)

We saw *a girl* was float and snk in the water. (026D)

We save *a girl*. (050D)

(46) ...we saw *two girls* had plucked some flower at the bank of the river. (030D)

There were *two girls* at the park but Hisham recognized them... (045D)

...they saw *two girls* were plucking a yellow flower beside the riverbank. (049D)

(47) Then, her's friend shout loudly for a help to save *that girl* from drawn. (030D)

I tried to save *that girl* but I'm failed to do that. (030D)

After a few minutes, Samad had saved *that girl*. (036D)

(48) *The other girl*, Amira shouted for help. (009D)

The other girl was very shocked and scared. (050D)

Two types of innovative language use with the singular and plural forms of GIRL are observed to occur at earlier periods in time. Some the participants did not add the suffix -s to the countable plural noun GIRL. Examples of creative use like *two girl* are found at Time 1 (49) for three times and at Time 2 (50) for nine times, and also examples like *their girl* are found at Time 2 (51) for two times.

Time 1

(49) ...we saw *two girl*, Izzati and Izzah were took a few flowers there. (019A)

We walked to go there. At the road, we saw *two girl* at the side river. (047A)

Suddenly, *two girl* pick some flowers at the river. (047A)

Time 2

(50) There was *two girl* were plucked flower near the river. (005B)

Two girl pick up many types of flower there. (018B)

...they saw *two girl* at the river side plucking some flowers. (048B)

(51) The girl's parents were proud to him because safe *their girl*. (039B)

Minah's parent feel very proud to Abu and Angah because safe *their girl*.(052B)

Table 4.12 shows the percentage of the innovative language use at different periods of time. It can be seen that the use of determiners of GIRL becomes more conventionalized over time, which may reflect that students are more aware of the “traditional” use of determiners.

Table 4.12: The innovative language use of determiners of GIRL

	Time 1	Time 2	Time 3	Time 4
Innovative language use	3 2.2%	11 6.2%	-	-

To sum up, five categories of determiners are used by the students over time, which are articles, numerals, possessive determiners, demonstratives, and semi-determiners. The article, especially the definite article *the*, is the most frequent type of determiners in the data, which accounts for over half of the total usage. Also, two types of innovative language use with the singular and plural forms of GIRL are observed to occur at earlier periods in time, but the use of determiners becomes more conventionalized over time.

4.2.2.1.2 N prep n

When GIRL is used in pattern **N prep n**, two prepositions are identified at students’ texts, which are *at* and *in*, as shown in Table 4.13.

It is surprising to find that the use of prepositional phrases as post-modifiers of the word GIRL is absent from Time 4. It may suggest that language development does not happen once and for all, which is consistent with the study of Larsen-Freeman (2010,

2015). Prepositional phrases introduced by *at* are observed at the first three data sets, and phrases introduced by *in* is only found at Time 2.

Table 4.13: Prepositional phrases of the word GIRL

Types	Time 1		Time 2		Time 3		Time 4	
	Freq	%	Freq	%	Freq	%	Freq	%
N <i>at</i> n	3	75%	4	57%	2	67%	-	-
N <i>in</i> n	-	-	2	29%	-	-	-	-

One type of prepositional phrases is identified at Time 1, which is *at* (52) with a percentage of 75%.

(52) The *girl at the shore* are shouting for help too. (06A)

Rahman help Ali to pulled the *girl at the river bank*. (009A)

At the road, we saw the two *girls at the side river*. (049A)

Two types of the prepositional phrases are identified at Time 2, which are *at* (53) with a percentage of 57% and *in* (54) with a percentage of 29%.

(53) At the home, Amran tells his family that he safe one *girl at the river*. (036B)

At the road, we saw a *girl at the river side*. (047B)

He saw two *girl at the river side* plucking some flowers. (048B)

(54) Ali quickly jumped into the river and helped the *girl in the river*. (006B)

They saw the *girl in the river*. (036B)

One type of prepositional phrases is identified at Time 3, which is *at* (55) with a percentage of 67%.

(55) The *girl at the river* was drowning. (025C)

When they arrived at the river, there were two *girl at the riverbank*. (004C)

It is noticed that two types of prepositional phrases are used to modify the word GIRL at students' texts, which are phrases with *at* and *in*. Surprisingly, no prepositional phrase is found at Time 4.

4.2.2.1.3 N clause

When GIRL is used in pattern **N clause**, three types of clauses are identified: one is introduced by *who*, one by *which* and one by *that*, as given in Table 4.14. Two types of clauses are identified at the first two data sets, while one type identified at the last two data sets. It shows the number of types of clauses decreases as time goes by. Clauses introduced by *who* is the only clause used through all the data sets.

Table 4.14: Clauses of the word GIRL

	Time 1		Time 2		Time 3		Time 4	
Types	Freq	%	Freq	%	Freq	%	Freq	%
N <i>who</i>	6	75%	3	43%	2	67%	2	50%
N <i>which</i>	2	25%	-		-		-	
N <i>that</i>	-		4	57%	-		-	

Two types of clauses are identified at Time 1, which are clauses introduced by *who* (56) with a percentage of 75% and *which* (57) with a percentage of 25%.

(56) Ali was happy because they have saved a *girl who fell into the lake*. (032A)

Kamal jump... to saved the *girl who was drowning* in the lake. (033A)

One of the *girl who was wearing a blackskirt*, has three flowers ... (49A)

(57) Ahmad jumped... to save the little *girl which was in the river*. (004A)

Azman getting down into the river to safe the *girl which is fight with life*. (049A)

Two types of clauses are identified at Time 2, which are clauses introduced by *who* (58) with a percentage of 43% and *that* (59) with a percentage of 57%.

(58) Nazim was calming down the another *girl who was so frightened*. (008B)

A *girl who name Kiara* thanked him for his good deed. (028B)

The *girl who felt into the pool* was cannot swim. (046B)

(59) The *girl that drowned* was safe now. (006B)

The *girl that pick up* some flower was tripped into the lake. (018B)

All the three boys helped the *girl that fell into the river*... (041B)

Clauses introduced by *who* are identified at Time 3 and Time 4, as follows:

(60) ...there were two *girls who are Jasmine* and Nadiah.(037C)

There were two *girls who are plucking some flowers* nearby. (048C)

(61)...two *girls who are plucking some flowers* by the river (003D)

Near the river, there were two *girls who plucked some flowers*. (048D)

Three types of clauses are used to modify the word GIRL: clauses introduced by *who*, by *which* and by *that*. Two types of clauses are identified at the first two data sets, while one type identified at the last two data sets. It shows the number of types of clauses decreases as time goes by.

Five types of determiners, two types of clauses and three types of clause are used to modify the noun GIRL in the data. The number of types of modifiers increases or decreases as time goes by. We can see that the use of GIRL especially in patterns **N prep n** and **N clause** shows a decline in terms of the diversity and complexity as time goes by. It indicates that language development is not a linear process.

4.2.2.2 Emergent lexis with GIRL

Adjectives and nouns in patterns **N n**, **poss-N n**, **adj N** and **N prep n** are examined to see how the word GIRL is used with other lexis in learn data. Table 4.15 presents the results of lexis which occur at least twice in different students' texts. At Time 1, two words emerge, while five words are observed to occur at Time 4.

Table 4.15: Emergent lexical items with GIRL

	Time 1	Time 2	Time 3	Time 4
Adjectives	little	poor	beautiful	beautiful
Nouns	parents	parents river	parents hand	parent parents hand friend

According to the frequency information, we can see from Table 4.16 that 1% of the word GIRL is found to be modified at Time 1 and Time 2, 2% at Time 3 and 3% at Time 4. There is a growth in the frequency of GIRL being modified, which may suggest that learners' lexical repertoire has enlarged over time.

Table 4.16 Frequency information of lexical items with GIRL

	Time 1	Time 2	Time 3	Time 4
Frequency	2 1%	3 1%	3 2%	5 3%

Over time, more lexical words are found to be used with the word GIRL, which may indicate that students' lexical repertoire expands over time.

Through the observation of the use of GIRL, we can see that language development is a complex process. Although the number of patterns of use of GIRL shows a slight decline over time, the instances under each pattern show a more complex picture: there is a fluctuation in terms of the number of determiners patterned with the word GIRL, but the use of determiners becomes more conventionalized over time; the type of

prepositional phrases and clauses becomes less diverse over time; the variety of adjectives and nouns patterned with the word GIRL gets increased over time.

4.3 FLOWER

A total of 381 instances of FLOWER are observed to occur in the learner data: 94 instances (sing.:29/ pl.: 65) at Time 1, 98 instances (sing.: 30/ pl.: 68) at Time 2, 81 instances (sing.: 29/ pl.: 52) at Time 3, and 108 instances (sing.: 34/ pl.: 74) at Time 4. The raw and normalized frequency of use of FLOWER across the data is summarized in Table 4.17.

Table 4.17: Raw and normalized frequency of use of FLOWER

	Time 1	Time 2	Time 3	Time 4
Size of sub-corpus (tokens)	8627	11205	11649	12309
Raw frequency	94	98	81	108
Normalized frequency	108	87	70	88

108 instances of FLOWER per 10,000 words are found at Time 1, 87 at Time 2, 70 at Time 3, and 88 at Time 4. As can be seen, the normalized frequencies across the data generally descend across time, although the data size in each sub-corpus increases.

A total of 381 concordance lines of FLOWER are identified in the data. The normalized frequency of the word in each data set shows a decrease over time. All these concordance lines were examined to explore the patterns of use of FLOWER.

4.3.1 Patterns of use of FLOWER

A total of four patterns of use of FLOWER are observed to occur in the data. They are **det N**, **adj N**, **N prep n**, and **N clause** (Table 4.18). Pattern **det N** indicates that the noun FLOWER is preceded by determiners (as in *some flowers* and *many flowers*). Pattern **adj N** indicates that the noun FLOWER is preceded by adjectives (as in *the wild flowers* and *the beautiful flowers*). Pattern **N prep n** indicates that the noun FLOWER is followed by prepositional phrases (as in *the flowers near the pond* and *the flowers at the river*). Pattern **N clause** indicates that the noun FLOWER is followed by clauses (as in *the flowers that she hold* and *flowers that were very beautiful*).

Table 4.18: Patterns of use of FLOWER

Patterns	Time 1		Time 2		Time 3		Time 4	
	Freq	%	Freq	%	Freq	%	Freq	%
det N	51	54%	61	62%	46	57%	67	62%
adj N	17	18%	19	19%	12	15%	27	25%
N prep n	31	33%	26	27%	19	23%	32	30%
N clause	2	2%	-		-		-	

The number of patterns of use of FLOWER experiences a slight decrease over time: four patterns (**det N**, **adj N**, **N clause** and **N prep n**) at Time 1, while three patterns (**det N**, **adj N**, and **N prep n**) at the other three data sets.

It is noticed that the most frequent pattern is **det N**, with a percentage of at least 54% at each data set. Pattern **N prep n** comes to the second place, with a percentage of over

23% (see Section 4.3.2.1 for examples of each of these patterns). In terms of each pattern, there is no constant increase or decrease in the percentage. It is thus necessary to zoom in on each pattern, and find the change and difference of use of patterns at different times.

4.3.2 Observation on instances under each pattern

As mentioned earlier, four patterns of the use of FLOWER are identified in the data, that is, **det N**, **adj N**, **N prep n**, and **N clause**. Closer analysis is then made on the instances under each pattern of use of the word. Examples of the use of FLOWER in the learner production are given. In addition, the types of function words (such as determiners, prepositions and clauses, in patterns **det N**, **N prep n**, and **N clause**) are concerned in Section 4.3.2.1. Meanwhile, the variety of content words (like nouns and adjectives in patterns **N prep n**, and **adj N**) is considered in Section 4.3.2.2.

4.3.2.1 Instances under each pattern of use of FLOWER

4.3.2.1.1 det N

When FLOWER is used in pattern **det N**, four types of determiners are used by students over time, as can be seen from Table 4.19. They can be further categorized into two groups: articles and quantifiers. More specifically, two types of articles (the definite article *the* and the indefinite article *a*) and two types of quantifiers (*some* and *many*) are identified in the data. The definite article *the* and the quantifier *some* are the most frequent determiners of FLOWER in the learner data.

Table 4.19: Determiners of the word FLOWER

Types	Categories	Time 1		Time 2		Time 3		Time 4	
		Freq	%	Freq	%	Freq	%	Freq	%
the	Articles	23	45%	30	49%	13	28%	31	46%
a		8	16%	7	11%	3	7%	6	9%
some	Quantifiers	13	25%	19	31%	21	46%	20	30%
many		4	9%	4	7%	4	9%	4	6%

Four types of determiners are identified from students' texts at Time 1, which are the definite article *the* (62) with a percentage of 45%, the indefinite article *a* (63) with a percentage of 16%, the quantifier *some* (64) with a percentage of 25%, the quantifier *many* (65) with a percentage of 9%.

(62) My sister picked *the flowers* while I was collecting it. (008A)

They were so excited and picked *the flowers*. (022A)

Sometimes, my frineds and I also take *the flowers*. (024A)

(63) They are see a flower at the river. (027A)

They want to picked a flower there. (036A)

The girls are taking a flower. (044A)

(64) They saw two little girls took *some flowers* which are beside the river. (004A)

There were two girls pick *some flowers* at the river. (005A)

They wanted to get *some flowers* to give her teacher. (034A)

(65) They want to collect the *many flowers*. (002A)

At the river have *many flowers*. (044A)

Four determiners are identified from students' texts at Time 2, which are the definite article *the* (66) with a percentage of 49%, the indefinite article *a* (67) with a percentage of 11%, the quantifier *some* (68) with a percentage of 31%, the quantifier *many* (69) with a percentage of 7%.

(66) *The flowers* was very beautiful and she wanted to keep it. (007B)

They plucked *the flowers* by a lake at the park.(021B)

In their way, they saw two girls picked *the flowers* beside the river. (029B)

(67) They go to the lake to collect a flower. (017B)

Feroz looked the two girls planted a flower beside the lake. (028B)

The girls are walking while taking a flower at the river. (044B)

(68) They saw two little girls were took *some flowers*. (004A)

They wanted to find *some flowers* for their Science folio. (011B)

I and my friends plucked *some flowers* to gave our parents. (030B)

(69) There were *many flowers* there. (011B)

They are *many flowers* at the river. (027B)

Four types of determiners are identified from students' texts at Time 3, which are the definite article *the* (70) with a percentage of 28%, the indefinite article *a* (71) with a percentage of 7%, the quantifier *some* (72) with a percentage of 46%, the quantifier *many* (73) with a percentage of 9%.

(70) Muhaiminah and Basyirah were didn't plucked *the flowers* at the river. (024C)

When they take *the flowers*, three boys...walk behind them. (040C)

The flowers are very colourfull. (049C)

(71) Both of them were plucking *a flower*. (013C)

There, we saw Aminah and Siti plucked *a flower*. (039C)

(72) Fatimah were putting *some flowers* for their science project. (031C)

Salmah and Fatimah go to the river to take *some flowers*.(040C)

They also picked up *some flowers* which is their hobby. (043C)

(73) *Many flowers* at the river.(014C)

There are *many flowers* such as morning glory... (037C)

Four types of determiners are identified from students' texts at Time 4, which are the definite article *the* (74) with a percentage of 46%, the indefinite article *a* (75) with a percentage of 9%, the quantifier *some* (76) with a percentage of 30%, the quantifier *many* (77) with a percentage of 6%.

(74) *The flowers* very beautiful.(014D)

While plucking *the flowers*, her friend, Balqis, fell into the river. (020D)

They saw the two girl was pick up *the flowers* at the river. (044D)

(75) They saw teo girls was picking *a flower* eside the river. (016D)

...the boys were usually gave Dianza *a flower* ... (018D)

(76) There were *many flowers* grew by the river. (010D)

They plucked *many flowers* including Hibiscus... and many more. (023D)

When they walking near the river, they saw *many flowers* at the river. (040D)

(77) Lina saw *some flowers* near the lake. (001D)

Awatif and Rohana took *some flowers* near the river. (037D)

Near the river, there were two girls who plucked *some flowers*.(048D)

Instances of innovative language uses are observed through all the four data sets. Students are found to use the indefinite article *a* to modify the plural-formed noun *flowers* for two times at Time 1 (78), which indicates that students have a language system themselves, and language learning is an ongoing process.

(78) They are taking *a flowers* near a river as a schoolwork. (013A)

They were walking two womens plucking *a flowers*. (045A)

Also, students are found to apply the quantifier *many* to modify the singular-formed noun *flower* for two times at the first three data sets, as follows:

(79) While we walking along the river, we pluck *many colourful flower*. (020A)

Many beautiful and colourful flower at the river. (036A)

(80) ...the pond have *many beautiful flower*. (002B)

Many flower behind river. (016B)

(81) *Many flower* at behind the river. (016C)

Many beautiful flower at there. (036C)

Meanwhile, students are found to use quantifier *some* to modify the singular-formed noun *flower* for three times at Time 1 (82), seven times at Time 2 (83), four times at Time 3 (84), and five times at Time 4 (85).

(82) There also has two girls, was collected *some flower* near the river bank. (030A)

They were two girls picking *some flower* at side of the lake. (033A)

There were also two girls who were collect *some flower* at the river side. (048A)

(83) They were plucking *some beautiful flower* near the lake. (015B)

At the side of the lake, there was *some flower* that is beautiful. (018B)

Near the river, the two of couple see *some flower*. (027B)

(84) Then, the girl gave Firdaus *some flower* and smiled at him. (004C)

I had saw a two girls was plucked *some beautiful flower*. (030C)

They went there take *some flower* around there. (037C)

(85) Aminah took *some beautiful flower* to bring back to home. (017D)

Dianza was fell into the lake while pick *some flower* at river bank. (018D)

we saw two girls had plucked *some flower* at the bank of the river. (030D)

The table below shows the percentage of innovative language use of determiners of FLOWER. We can see that the use of determiners of FLOWER becomes more conventionalized over time.

Table 4.20: The innovative language use of determiners of FLOWER

	Time 1		Time 2		Time 3		Time 4	
Innovative language use	7	14%	9	15%	6	13%	5	7%

Two categories of determiners are used by students over time, which are articles and quantifiers. The definite article *the* and the quantifier *some* are the most frequent determiners of FLOWER in the learner data. Although instances of innovative language uses are observed through all the four data sets, the use of determiners of FLOWER becomes more conventionalized over time.

4.3.2.1.2 N prep n

When the word FLOWER is used in pattern **N prep n**, many prepositions are observed to occur with the word (Table 4.21). Four types of prepositional phrases are identified at Time 1 and Time 4, three at Time 2, and two at Time 3. Prepositions *at* and *near* are observed across the four data sets, with a percentage of over 62%.

Table 4.21: Prepositional phrases of the word FLOWER

Types	Time 1		Time 2		Time 3		Time 4	
	Freq	%	Freq	%	Freq	%	Freq	%
N at n	12	39%	9	35%	7	37%	16	50%
N near n	7	23%	9	35%	9	47%	8	25%
N beside n	6	19%	4	15%	-		2	6%
N by n	2	6%	-		-		-	
N in n	-		-		-		2	6%

Four prepositional phrases are identified from students' texts at Time 1, which are *at* (86) with a percentage of 39%, *near* (87) with a percentage of 23%, *beside* (88) with a percentage of 19%, and *by* (89) with a percentage of 6%.

(86) There were two girls pick some *flowers at the river*. (005A)

They saw two girls was plucked the *flowers at the riverbank*. (009A)

They saw two girls were plucking *flowers at the bank of the river*. (051A)

(87) They are taking a *flowers near a river* as a schoolwork. (013A)

They go there to collected *flowers near the lake*. (017A)

Zaiton was collecting *flowers near the bank of river*. (023A)

(88) They also pluck the *flower beside the lake*. (001A)

The two girls wanted to pick up some *flowers beside the river*. (037A)

...they saw two girls which was plucking some *flowers beside the river*. (049A)

(89) ...two girls picking *flowers by the river bank*. (003A)

Suddenly, while plucking *flowers by the river*, Aina fell into the river. (010A)

Three prepositional phrases are identified from students' texts at Time 2, which are *at* (90) with a percentage of 35%, *near* (91) with a percentage of 35%, and *beside* (92) with a percentage of 15%.

(90) At the river, Ani and Ina take a few of *flowers at the river*. (040B)

They saw two girls plucked *flowers at the bank of the river*. (051B)

Minah and Amirah is plucking the beautiful *flowers at the riverside*. (052B)

(91) At there, they saw two girls were plucking *flowers near the lake*.(024B)

They were plucking the *flowers near a pond*. (026B)

We saw two girls were plucking some *flowers near the river*. (037B)

(92) ...we saw Aida and Sarah were plucking *flowers beside the river*.(019B)

In their way, they saw two girls picked the *flowers beside the river*. (029B)

We saw two girls were plucked some *flowers beside the pool*. (046B)

Two prepositional phrases are identified from students' texts at Time 3, which are *at* (93) with a percentage of 37% and *near* (94) with a percentage of 47%.

(93) Nita decided to pluck some *flowers near a lake*. (021C)

Afie was plucking *flowers near the river*. (044C)

We saw two girls were plucked some *flowers near the river*. (046C)

(94) Muhaiminah and Basyirah were didn't plucked the *flowers at the river*. (024C)

Ahmad saw two beautiful girls picked *flowers at the river bank*. (025C)

They were plucking *flowers at the pond*.(045C)

Four prepositional phrases are identified from students' texts at Time 4, which are *at* (95) with a percentage of 50%, *near* (96) with a percentage of 25%, *beside* (97) with a percentage of 6%, and *in* (98) with a percentage of 6%.

(95) Lina saw some *flowers near the lake*. (001D)

Johana and Irma were plucked *flowers near the river*. (005D)

While walked, they saw two girls plucked *flowers near a riverbank*. (009D)

(96) Liyana was attracted to some *flowers at the river bank*. (011D)

When they walking near the river, they saw many *flowers at the river*. (040D)

They saw the two girl was pick up the *flowers at the river*. (044D)

(97) Farah want to pluck *flower beside the pond*... (002D)

They picked the *flowers beside the rievrr*. (014D)

...they saw two girls were plucking a yellow *flower beside the riverbank*. (049D)

(98) Actually, Ika wanted to pluck a beautiful *flower in the water*. (015D)

B-mok and Zack were taking the beautiful *flower in the side of river*. (047D)

Five types of prepositional phrases are used to modify the word FLOWER in the data.

Four types of prepositional phrases are identified at Time 1 and Time 4, three at Time 2, and two at Time 3.

4.3.2.1.3 N clause

When FLOWER is used in pattern **N clause**, one type of clauses are observed to occur only at Time 1 (99), as shown in Table 4.22.

Table 4.22: Clauses of the word FLOWER

	Time 1		Time 2		Time 3		Time 4	
Types	Freq	%	Freq	%	Freq	%	Freq	%
N <i>that</i>	2	100%	-		-		-	

(99) The *flowers that she hold* also fell into the river. (015A)

There were many types of *flowers that were very beautiful* at the lake. (018A)

It is noticed that only one type of clauses are observed to occur at learner data, which is **N *that***. Such use only emerges at Time 1.

Four types of determiners, five types of prepositional phrases and one type of clauses are used to modify the word FLOWER in the learner data. We can see that the use of FLOWER in pattern **N prep n** shows a slight fluctuation in terms of the diversity and complexity as time goes by.

4.3.2.2 Emergent lexis with FLOWER

The nouns and adjectives in patterns **adj N** and **N prep n** are examined to see how the word FLOWER is used with other lexis in learner data. Table 4.23 presents the results of nouns and adjectives which occur at least twice in different students' texts.

At Time 1, four words (two adjectives and two nouns) emerge, while five words (two adjectives and three nouns) are observed to occur at Time 4.

Table 4.23: Emergent lexical items with FLOWER

	Time 1	Time 2	Time 3	Time 4
Adjectives	beautiful	beautiful	beautiful	beautiful
	colourful	colourful		colourful
Nouns	river	river	river	river
	lake	lake	river bank	lake riverbank

According to the frequency information from Table 4.24, we can see that 4% of the word FLOWER is found to be modified at the first three data sets, while the percentage rises up to 5% at the last data set.

Table 4.24: Frequency information of lexical items with FLOWER

	Time 1	Time 2	Time 3	Time 4
Frequency	4 4%	4 4%	3 4%	5 5%

Over time, more lexical words are found to be used with the word FLOWER, which may indicate that students' lexical repertoire expands over time.

Through the observation of the use of FLOWER, we can see that language development is a complex process. Although the number of patterns of use of FLOWER shows a slight decrease over time, the instances under each pattern show a more complex picture: the type of determiners used with the word FLOWER remains the same; the number of

prepositional phrases of FLOWER experiences a fluctuation; the number of clauses declines over time; and the variety of nouns and adjectives has expanded over time.

4.4 LAKE

A total of 231 instances of lake are observed to occur in the learner data: 63 instances (sing.) at Time 1, 56 instances (sing.) at Time 2, 64 instances (sing.) at Time 3, and 48 instances (sing.) at Time 4. The raw and normalized frequency of use of lake across the data is summarized in Table 4.25.

Table 4.25: Raw and normalized frequency of use of *lake*

	Time 1	Time 2	Time 3	Time 4
Size of sub-corpus (tokens)	8627	11205	11649	12309
Raw frequency	63	56	64	48
Normalized frequency	73	50	55	39

71 instances of RIVER per 10,000 words are found at Time 1, 50 at Time 2, 55 at Time 3, and 39 at Time 4. As can be seen, the normalized frequencies across the data descend across time, although the sizes of the sub-corpora get swollen.

Generally, a total of 231 concordance lines of *lake* are identified in the data. The normalized frequency of the word in each data set shows a decrease over time. All these concordance lines were examined to study the patterns of use of *lake*.

4.4.1 Patterns of use of *lake*

A total of four patterns of use of *lake* are observed to occur in the data (Table 4.26). They are **N n**, **det N**, **adj N**, and **N prep n**. Pattern **N n** indicates that the noun *lake* was followed by another noun (as in *the lake water* and *the lake name*). Pattern **det N** indicates that the noun *lake* is preceded by determiners (as in *the lake* and *a lake*). Pattern **adj N** indicates that the noun *lake* is preceded by adjectives (as in *the beautiful lake* and *the huge and beautiful lake*). Pattern **N prep n** indicates that the noun *lake* was followed by prepositional phrases (as in *the lake near their house* and *the lake near their school*).

Table 4.26: Patterns of use of *lake*

	Time 1		Time 2		Time 3		Time 4	
Patterns	Freq	%	Freq	%	Freq	%	Freq	%
N n	2	3%	-	-	-	-	-	-
det N	59	97%	53	98%	62	97%	44	92%
adj N	-	-	2	4%	-	-	-	-
N prep n	3	5%	2	4%	5	8%	4	8%

The number of patterns of use of *lake* experiences a slight decrease over time: three patterns (**N n**, **det N**, **adj N**, and **N prep n**) at first two data sets respectively, while two patterns (**det N** and **N prep n**) at the last two.

It is noticed that the most frequent pattern is **det N**, with a percentage of at least 92% at each data set. The frequency of the rest two patterns (**adj N**, and **N prep n**) is far behind with a percentage of less than 8% across the corpus (see Section 4.4.2.1 for examples of each of these patterns). In terms of each pattern, there is no constant increase or

decrease in the percentage. It is thus necessary to zoom in on each pattern, and find the change and difference of use of patterns at different time.

4.4.2 Observation on instances under each pattern

As mentioned earlier, four patterns of use of *lake* are identified, that is, **N n**, **det N**, **adj N**, and **N prep n**. Closer observation is then made on the instances under each pattern of use of the word. Examples of the use of *lake* in the learner production are given. In addition, the type of function words (such as determiners and prepositions in patterns **det N** and **N prep n**) is concerned in Section 4.4.2.1. Meanwhile, the variety of content words (like nouns and adjectives in patterns **N n**, **N prep n**, and **adj N**) is considered in Section 4.4.2.2.

4.4.2.1 Instances under each pattern of use of *lake*

4.4.2.1.1 det N

When *lake* is used in pattern **det N**, two types of determiners are identified, that is, the definite article *the* and the indefinite article *a* (Table 4.27). The definite article *the* is the most frequent determiner patterned with the word *lake*.

Table 4.27: Determiners of the word *lake*

Types	Categories	Time 1		Time 2		Time 3		Time 4	
		Freq	%	Freq	%	Freq	%	Freq	%
the	Articles	56	95%	47	92%	57	92%	41	93%
a		3	5%	4	8%	5	8%	3	7%

Examples of use of the definite article *the* from students' texts across time are given below:

Time 1

(100) They also pluck the flower beside *the lake*. (001A)

Ali was happy because they have saved a girl who fell into *the lake*. (032A)

One of them jumped into *the lake* and took Adila to the shore. (043A)

Time 2

(101) Shiera was drowned in *the lake* and asked for help. (001B)

One evening, Rashid with his friend, Rizal and Chong went to *the lake*. (015B)

There were three boys walked around *the lake*. (017B)

Time 3

(102) At *the lake*, there were so many people came to fishing, jog and cycled. (018C)

Fatimah's father came to *the lake* and brought her to the nearby hospitals. (031C)

Lan Chau who was a good swimmer immediately jumped into *the lake*. (043C)

Time 4

(103) Fahmy jumped into *the lake* and saved Lina. (001D)

Then, we together went to *the lake* happily. (032D)

It was Aminah, she was felt into *the lake* because the slippery surface. (039D)

Examples of use of the indefinite article *a* from students' texts across time are as follows:

Time 1

(104) They see two little girls are taking flowers at the shore of *a lake*. (006A)

Last Saturday, Ali and his friends wanted to fishing in *a lake*. (032A)

Last Sunday, Maniam and Adila went to *a lake*. (043A)

Time 2

(105) One day, Siti and Aminah go to *a lake*. (017B)

They plucked the flowers by *a lake* at the park. (021B)

Last Saturday, Fareez, Amirul and Rahman went to *a lake*. (024B)

Time 3

(106) Samad went to *a lake* near their home like just throw away stone. (017C)

Nita decided to pluck some flowers near *a lake*. (021C)

Most of them are birds who want to get drink from water in *a lake*. (037C)

Time 4

(107) Shima walked together beside *a lake* in the Botanical Garden. (001D)

Saorah asked Lucy out for a walk to *a lake* near their neighbourhood. (022D)

He always invite me and Amir to *a lake* at weekends. (032D)

Two types of determiners are found to modify the word *lake* in the learner data, that is, the definite article *the* and the indefinite article *a*. The definite article *the* is the most frequent determiner.

4.4.2.1.2 N prep n

When *lake* is used in pattern **N prep n**, only one preposition is observed to emerge in the data, that is, *near* (Table 4.28) with a percentage of over 50%, as shown in (108) to (111).

Table 4.28: Prepositional phrases of the word *lake*

Types	Time 1		Time 2		Time 3		Time 4	
	Freq	%	Freq	%	Freq	%	Freq	%
N <i>near</i> n	3	100%	2	50%	5	100%	4	75%

Time 1

(108) Ali, Ramli and Rahman went for fishing at the *lake near their house*. (009A)

Ahamd, Zakri and Ramu planned to go to the *lake near their house*. (015A)

Sarah and Siti went for a walk at the *lake near their neighbourhood*. (022A)

Time 2

(109) Feroz invite his friends to go fishing at the *lake near their village*. (028B)

Kamal and his friends were going fishing at the *lake near their house*. (033B)

Time 3

(110) Samad went to a *lake near their home* like just throw away stone. (017C)

Sarah and Sabrina went for a walk at the *lake near their neighbourhood*. (022C)

One day, they decided to fish in the *lake near their school*. (031C)

Time 4

(111) Riza and Farouq planned to go fishing at the *lake near their village*. (015D)

They walked along the *lake near their house* like throw a stone. (017D)

Saorah asked Lucy out for a walk to a *lake near their neighbourhood*. (022D)

It is noticed that only one type of prepositions is observed to emerge across the sub-corpora, that is, *near*.

Two types of determiners and one type of prepositions are found to modify the word *lake* in the learner data, but patterns **det N** and **N prep n** have not shown much difference in terms of the diversity and complexity as time goes by.

4.4.2.2 Emergent lexis with *lake*

The nouns and adjectives in patterns **N n**, **adj N** and **N prep n** are examined to see how the word *lake* is used with other lexis in learn data. Table 4.29 presents the results of nouns and adjectives which occur at least twice in different students' texts. At Time 1, two words (two nouns) emerge, while one word (one noun) are observed to occur at Time 4.

Table 4.29: Emergent lexical items with *lake*

	Time 1	Time 2	Time 3	Time 4
Adjectives	-	beautiful	-	-
Nouns	house name	-	house	house

According to the frequency information (Table 4.30), we can see that 3% of the word *lake* is found to be modified at the first data set, while 2% is found at the last three data sets.

Table 4.30: Frequency information of lexical items with *lake*

	Time 1	Time 2	Time 3	Time 4
Frequency	2 3%	1 2%	1 2%	1 2%

Through the observation of the use of *lake*, we have not see much language development over time. The type of determiners and prepositional phrases remains the same, and the number of the nouns and adjectives also remains the same.

4.5 FRIEND

A total of 501 instances of FRIEND are observed to occur in the learner data: 118 instances (sing.: 60/ pl.:58) at Time 1, 147 instances (sing.: 71/ pl.:76) at Time 2, 103 instances (sing.: 57/ pl.:46) at Time 3, and 133 instances (sing.: 62/ pl.:71) at Time 4. The frequency information of FRIEND is summarized in Table 4.31.

Table 4.31: Raw and normalized frequency of use of FRIEND

	Time 1	Time 2	Time 3	Time 4
Size of sub-corpus (tokens)	8627	11205	11649	12309
Raw frequency	118	147	103	133
Normalized frequency	137	131	88	92

137 instances of FRIEND per 10,000 words are found at Time 1, 131 at Time 2, 88 at Time 3, and 92 at Time 4. As can be seen, the normalized frequencies across the data show a decrease across time, although the data size in each sub-corpus grows bigger.

A total of 501 concordance lines of FRIEND were identified in the data. The normalized frequency of the word in each data set shows a decrease over time. All these concordance lines were examined to explore the patterns of use of FRIEND.

4.5.1 Patterns of use of FRIEND

A total of three patterns of use of FRIEND are observed to occur in the data (Table 4.32). They are **det N**, **adj N**, and **N clause**. Pattern **det N** indicates that the noun FRIEND is preceded by determiners (as in *her friend* and *my friend*). Pattern **adj N** indicates that the noun FRIEND is preceded by adjectives (as in *the good friend* and *the best friend*). Pattern **N clause** indicates that the noun FRIEND is followed by clauses (as in *my best friend who live in Taman Sentosa*).

Table 4.32: Patterns of use of FRIEND

	Time 1		Time 2		Time 3		Time 4	
Patterns	Freq	%	Freq	%	Freq	%	Freq	%
det N	99	84%	112	76%	77	75%	109	82%
adj N	10	8%	21	14%	16	16%	7	5%
N clause	-		2	1%	-		-	

The number of patterns of use of FRIEND experiences a fluctuation over time: two patterns (**det N** and **adj N**) identified at Time 1, Time 3 and Time 4, whereas three patterns (**det N**, **adj N**, and **N clause**) identified at Time 2.

It is noticed that the most frequent pattern is **det N**, with a percentage of over 75% at each data set. The frequency of the rest two patterns (**adj N** and **N clause**) is far behind with a percentage of less than 16% across the corpus (see Section 4.5.2.1 for examples of each of these patterns). In terms of each pattern, there is no constant increase or decrease in the percentage. It is thus necessary to zoom in on each pattern, and find the change and difference of use of patterns at different time.

4.5.2 Observation on instances under each pattern

As mentioned earlier, three patterns of use of FRIEND, i.e., **det N**, **adj N** and **N clause** are observed to occur in the data. Closer observation is then made on the instances under each pattern of use of the word. Examples of the use of FRIEND in the learner production are given. In addition, the types of function words (such as determiners and clauses in patterns **det N** and **N clause**) are concerned in Section 4.5.2.1. Meanwhile, the variety of content words (like adjectives in pattern **adj N**) is considered in Section 4.5.2.2.

4.5.2.1 Instances under each pattern of use of FRIEND

4.5.2.1.1 det N

When FRIEND is used in pattern **det N**, six types of determiners are used by students over time, as can be seen from Table 4.33. They can be further categorized into three groups: articles, numerals and possessive determiners. The possessive is the most frequent determiner in the data, which accounts for at least 81% of the total usage of the determiners.

The use of the indefinite article *a* is observed to occur at earlier periods in time. Possessive determiners (*my*, *his* and *her*) emerge across the corpus during different periods of time. The possessive *their* is found to be used to determine the plural-formed noun friends at Time 4. The use of numerals is observed to occur at later periods of time.

Table 4.33: Determiners of the word FRIEND

Types	Categories	Time 1		Time 2		Time 3		Time 4	
		Freq	%	Freq	%	Freq	%	Freq	%
a	Article	2	2%	5	4%	2	3%	-	-
my	Possessives	29	29%	15	8%	11	14%	16	15%
his		39	39%	66	44%	24	31%	50	46%
her		25	25%	19	29%	33	43%	32	29%
their		-	-	-	-	-	-	2	2%
three	Numeral	-	-	2	2%	3	4%	2	2%

Four types of determiners are identified at Time 1, which are the indefinite article *a* (112) with a percentage of 2%, the possessive *my* (113) with a percentage of 29%, the

possessive *his* (114) with a percentage of 39%, and the possessive *her* (115) with a percentage of 25%.

(112) Siti, who is a naughty girl has *a friend*, Ah Mei. (021A)

The girl is scared she call the boys to help *a friend*. (044A)

Ali, Ramli and Rahman are *a best friend*. (009A)

(113) Her friend said, "Thank you, because you had saved my friend." (004A)

She always picking flowers with *her friend*, Hidayah. (007A)

The another girl was thanked to Ali because saved *her best friend*. (032A)

(114) *His friend* shouted for help. (009A)

Kamal and *his friend* brought the girl to the girl's house. (033A)

Ahmad was like "Superman" when he together gather with *his friend*. (039A)

(115) *My friend* was in the river. (004A)

She is *my friend*. (018A)

My friend and I went back home happily but a lot of thirsty. (038A)

Five determiners are identified at Time 2, which are the indefinite article *a* (116) with a percentage of 4%, the possessive *my* (117) with a percentage of 8%, the possessive *his* (118) with a percentage of 44%, the possessive *her* (119) with a percentage of 29%, and the numeral *three* (120) with a percentage of 2%.

(116) *A friend* in need is a friend indeed (001B)

They also become *a friend* and always played happily together. (006B)

A boy jump on the river to helping *a friend* and the girl. (044B)

(117) Fatin was very happy because *her best friend* was safe now. (001B)

Her friend shouted for a help. (005B)

Her friend was trying to get help by shouted for help. (006B)

(118) Ali and *his friend* hear the shout and caget for a while. (025B)

Grateful, Rashid and *his friend* heard the shouted. (015B)

Su thanked to Ali and *his friend* because save her life. (034B)

(119) Please help *my friend!* (015B)

At 10.00a.m., *my friend* came to my house. (030B)

My friend and I went to KCF and ate the tasty meal used the money. (050B)

(120) In the village, live *three best friends* name Zaki, Farham and Izuddin. (005B)

...the *three good friends* talked to each other. (006B)

Four type of determiners are identified at Time 3, which are the indefinite article *a* (121) with a percentage of 3%, the possessive *my* (122) with a percentage of 14%, the possessive *his* (123) with a percentage of 31%, and the possessive *her* (124) with a percentage of 43%.

(121) ...*a friend* in need is a friend indeed. (001C)

Nina and Intan was *a best friend*. (011C)

After that Chua and Aminah became *a good friend*. (039C)

(122) Ema want Irfan and his friend helped *her friend* Jika. (016C)

Her friend was very shocking and did not knew how to help her friend. (041C)

What *her friend* can do only hope that somebody came to save her friend. (050C)

(123) Ali and *his friend* heard somebody in the emergency. (034C)

Raju and *his friend* heard the shouted. (036C)

The girl's parent was very thankfull to Udin and *his friend* (049C)

(124) Please help *my friend!* (015C)

My friend, my brother, Hao and me Asakura Yoh went to a park...(037C)

She was very thankful to me and *my friend* because saving her. (044C)

Five types of determiners are identified at Time 4, which are the possessive *my* (125) with a percentage of 15%, the possessive *his* (126) with a percentage of 46%, the possessive *her* (127) with a percentage of 29%, the possessive *their* (128) with a percentage of 2%, and the numeral *three* (129) with a percentage of 2%.

(125) She is *my friend*. (018D)

Then *My friend* were laughed. (050D)

My friend were very shoked went we heard people shouted. (050D)

(126) Dianza walked around the lake while *her friend* pick flower as usual. (018D)

Her friend had pushed she than felt into river. (019D)

Her friend was shoched and very scared... (052D)

(127) *His friend* was decided to go fishing at the river near their house. (036D)

They invited Ali and *his friend* to go to a party at Yati's house. (043D)

Her parents were thanked to Amin and *his friend* to save their daughter. (049D)

(128) Farouq get a meaningful story to tell *their friends* when go to school. (015D)

...they were very happy because they had helped *their friends*. (020D)

(129) Then, the *three best friends* were very shocked. (005D)

The *three best friends* and Zaiton were also brought to the hospital. (006D)

It is noticed that students were aware of using the numeral *three* to quantify the plural-formed noun *friends* at Time 2 and Time 4. However, students at Time 3 are found to use the numeral *three* to determine the singular-formed noun *friend* for three times, as in (130).

(130) There were *three friend* named Ali, Ahmad and Abu went for fishing. (025C)

The *three friend* walked to the river because can save someone life. (025C)

In a village, was found *three good friend* which is Ali, Abu and Alan. (029C)

Three categories of determiners are used by students over time, which are articles, numerals and possessive determiners. The possessive is the most frequent determiner in the data. One instance of innovative language use is found at Time 3.

4.5.2.1.2 N clause

When FRIEND is used in pattern **N clause**, the clause type is rather limited (Table 4.34). Only one type of clauses is observed to occur, that is, clauses introduced by *who*. Also, this type of use is only present at Time 2 (131).

Table 4.34: Clauses of the word FRIEND

Types	Time 1		Time 2		Time 3		Time 4	
	Freq	%	Freq	%	Freq	%	Freq	%
N <i>who</i>	-		2	100%	-		-	

(131) Shiera and Fatin are my *friend who lives next to my house*. (001B)

Ali is my best *friend who live in Taman Sentosa*. (032B)

It is notice that only one type of clauses is observed to occur, that is, clauses introduced by *who*. Also, this type of use is only present at Time 2.

Three categories of determiners and one type of clause are used to modify the word FRIEND in the learner data. We can see that the use of FRIEND in patterns **det N** and **N clause** shows a slight difference as time goes by.

4.5.2.2 Emergent lexis with FRIEND

The adjectives in pattern **adj N** are examined to see how the word FRIEND is used with other lexis in learner data. Table 4.35 presents the results of the adjectives which occur

at least twice in different students' texts. Two exactly same words (*best* and *good*) are observed to occur across four data sets.

Table 4.35: Emergent lexical items with FRIEND

	Time 1	Time 2	Time 3	Time 4
Adjectives	best	best	best	best
	good	good	good	good

According to the frequency information (Table 4.36), it can be seen that the percentage of the word FRIEND being modified across the data sets roughly remains the same.

Table 4.36: Frequency information of lexical items with FRIEND

	Time 1	Time 2	Time 3	Time 4
Frequency	2 2%	2 1%	2 2%	2 2%

Through the observation of the use of FRIEND, it can be seen that language development is a complex process. Although the number of patterns of use of FRIEND fluctuates over time, the instances under each pattern show a more complex picture: the number of determiners has increased over time; and the type of clauses and the number of adjectives in pattern **adj N** fluctuate over time.

From the observation of these selected words, we can see that a total of four patterns, **det N**, **adj N**, **N prep n** and **N clause**, are identified. The number of patterns of each noun increases or decreases, as time goes by. The observation on each pattern shows the

complex use of determiners, adjectives, prepositional phrases, and clauses. The emergent lexis of the selected words has increased, which may indicate that students' lexical repertoire has expanded over time.

4.6 Conclusion

In this chapter, I have investigated how the use of the selected nouns in the data changes over a period of 24 months. A total of six patterns of use of the nouns are identified, that is, **det N**, **adj N**, **N prep n**, **N clause**, **N n** and **poss-N n**. The findings show that the number of patterns of each noun varies, as time goes by. A closer observation is made of each pattern, and instances under patterns are examined. The findings show a rather complex picture of the use of determiners, adjectives, prepositional phrases, clauses and emergent lexis. The next chapter will discuss the findings, provide the implications from this study, and make recommendations for future research.

CHAPTER 5

DISCUSSION AND CONCLUSION

This research examined five high frequency nouns, *river*, *girl*, *flowers*, *lake* and *friend*, together with their respective singular or plural forms, in the data to explore L2 development. The patterns of use of these nouns were identified, quantified and observed. This chapter will address the two research questions raised in Chapter 1, illustrate the implications of language learning and teaching, and make recommendations for future research.

5.1 The research questions

Research question 1: What are the recurrent patterns of use of the selected nouns used by Malaysian secondary school students at four different points in time in the subset of LoCLaNT?

To answer the first question, a total of 2676 concordance lines were analysed, and four types of patterns (**det N**, **adj N**, **N prep n**, **N clause**, **N n** and **poss-N n**) were identified in the data. The results show that the number of patterns of use of each noun varies at different periods in time, as can be seen from Table 5.1.

Table 5.1 Patterns of use of the selected nouns across time

	Time 1	Time 2	Time 3	Time 4	
	Patterns				Examples
RIVER	det N				<i>the/a river</i>
	N n				<i>river bank/side</i>
	adj N				<i>deep/big river</i>
	N prep n				<i>the river near her house</i> <i>the river behind their house</i> <i>the river at the park</i>
	-	N clause			<i>a river which is not far from their house</i> <i>the river which is located near house</i> <i>The river that they want to go</i>
GIRL	det N				<i>one/that/another girl</i>
	poss-N n				<i>girl's parents/hand/friend</i>
	adj N				<i>little/poor/beautiful girl</i>
	N clause				<i>a girl who fell into the lake</i> <i>a girl who felt into the pool</i> <i>The girl that drowned</i>
	N prep n			-	<i>the girl at the shore</i> <i>the girl in the river</i> <i>the girl at the river bank</i>
	N n				<i>girl friend/parents</i>
FLOWER	det N				<i>some/many flowers</i>
	adj N				<i>beautiful/colourful flowers</i>
	N prep n				<i>flowers near the lake</i> <i>flowers at the riverbank</i> <i>flowers beside the pool</i>
	N clause	-	-	-	<i>flowers that she hold</i> <i>flowers that were beautiful</i>
LAKE	det N				<i>a lake/the lake</i>
	N prep n				<i>the lake near their house</i>
	N n	-	-	-	<i>the lake name</i>

	-	adj N	-	-	<i>the beautiful lake</i>
FRIEND	det N				<i>my/his/her friend</i>
	adj N				<i>good/best friend</i>
	-	N clause	-	-	<i>my best friend who live in Taman</i> <i>my friend who lives next to my house</i>

There is a slight increase in the number of patterns of use of the word RIVER over time. Four patterns of use of the word RIVER are identified at Time 1, while five patterns identified at the following three data sets. There is a decline in the number of patterns related to the words GIRL, FLOWER and LAKE. Six patterns of use of the word GIRL are identified at the first three data sets, while five patterns identified at Time 4; four patterns of use of the word FLOWER are identified at Time 1, and three patterns identified at the last three data sets; and three patterns of use of the word LAKE are identified at the first two data sets, while two patterns identified at last two data sets. There is a fluctuation in the number of patterns concerning the word FRIEND. Two patterns of use of the word FRIEND are identified at the first and last two data sets, while three patterns identified at Time 2 and Time 3.

It is also noticed that all these four patterns can fall into three categories: 1) determiner-related patterns and 2) modifier-related patterns, and 3) word-formation patterns. The first group involves **det N**; the second **adj N**, **N prep n**, and **N clause**; and the third **N n** and **poss-N n**. Among all the patterns, **det N** is the most frequent. It makes sense that English nouns, especially countable nouns like RIVER, GIRL, FLOWER, LAKE and FRIEND, tend to be preceded by determiners.

Research question 2: How do the patterns of use of these nouns change over time in the subset of LoCLaNT?

To address the second research question, the instances under each pattern with regard to the function words and lexical words were further studied. The findings show a rather complex picture of the use of determiners, adjectives, prepositional phrases, clauses and emergent lexis.

The use of determiners becomes more conventionalized over time, as shown in Table 5.2. It shows that about 2% usages of determiners at Time 1 are what is called innovative language use, while the percentage drops to 0.9% at Time 4. It is known that determiners are often restricted to the number and countability of the noun with which they can occur (Maros, Tan, & Salehuddin, 2007). It is found in learners' writing that they have confusion with the application of the suffix -s to a noun. Two cases happened: the suffix -s was applied to a singular noun, or the suffix -s was not applied to a plural noun. A possible reason is that the plural markers are used for nouns in English which do not exist in Malay (Darus & Subramaniam, 2009).

Table 5.2: The innovative language use of determiners

	Time 1	Time 2	Time 3	Time 4
Innovative language use	10 2.0%	20 3.5%	9 1.7%	5 0.9%

The use of adjectives, especially of RIVER, becomes more complex over time. Apart from adjectives like *deep* and *small* used to modify the word, phrasal adjectives (as in

the deep and big river/ a beautiful and clean river) and adjective compounds (as in *a fast-flowing river/ the fast flowing river*) are observed to occur at a later point in time.

The use of post-modifiers (the prepositional phrases and clauses) indicates a complex picture. A total of seven types of prepositional phrases are observed to occur with the selected words in the data (Figure 5.1). Five types of prepositional phrases are identified at Time 1, six types of prepositional phrases identified at Time 2, three types of prepositional phrases identified at Time 3, and four types of prepositional phrases identified at Time 4. There is a decline in the types of prepositional phrases used with the selected nouns over time, which may suggest that language development is a non-linear process.

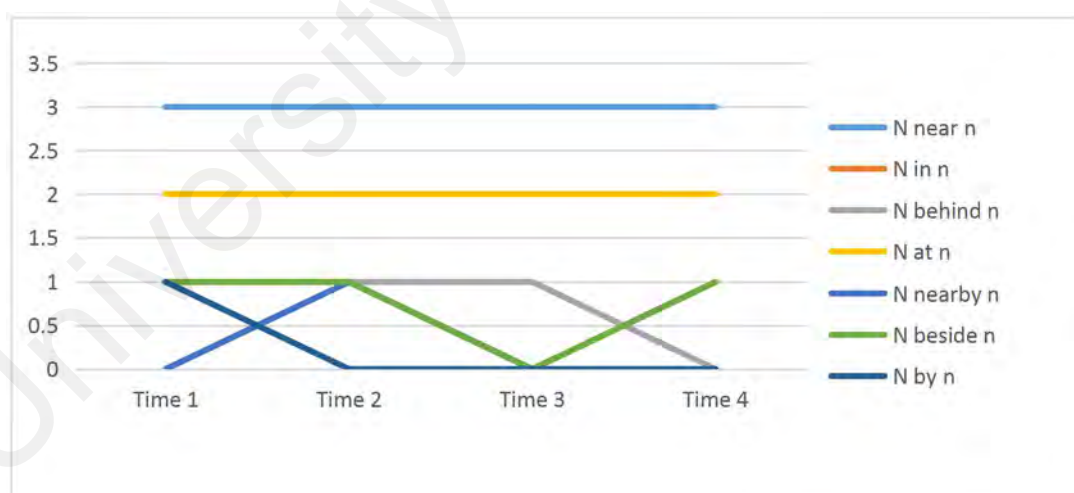


Figure 5.1: The number of prepositional phrases of the selected nouns

A total of four types of clauses are identified with the selected words in the learner data (Figure 5.2). Three types of clauses are found at the first two data sets, two types of clauses found at Time 3, and four types of clauses found at Time 4. There is a growth in

terms of the types of clauses used with the selected nouns over time, which may suggest that learners' lexical repertoire has expanded in time.

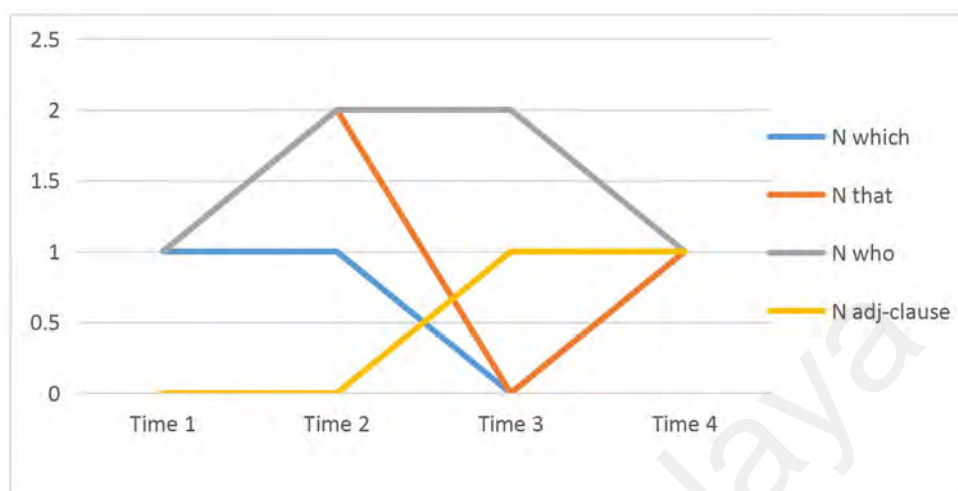


Figure 5.2: The types of clauses of the selected nouns

The list of emergent lexis of the selected nouns is presented in Table 5.3. As can be seen, 13 words are found to be used to modify the selected words at Time 1, while 20 words are identified at Time 4. In other words, 2.0% of use of these nouns is found to be modified at Time 1, while 2.9% to be modified at Time 4. It may indicate students are more aware of using lexical items to modify the key words, and also students' lexical repertoire gets expanded across time.

Table 5.3: Emergent lexis used with the selected nouns

	Time 1	Time 2	Time 3	Time 4
RIVER	deep house bank	deep house bank side	nearby house village bank	nearby deep big house village bank side

GIRL	little parents	poor parents river	drowned beautiful parents hand	beautiful parents parent hand friend
FLOWER	beautiful colourful river lake	beautiful colourful river lake	beautiful river river bank	beautiful colourful river lake riverbank
LAKE	house name	beautiful	house	house
FRIEND	good best	good best	good best	good best

Through the observation on the use of these selected words, it is realized that language development is a non-linear, complex process (e.g., Larsen-Freeman, 2006, 2009; Ortega, 2009; Ellis, 2011; Chau, 2015). Chau (2015) explains the process of L2 development as both a constructive process and a reductionist process. The former shows the increment in language resources, while the latter is the decline in those resources. That is,

Language development is not just about accumulating and storing an ever larger pool of language resources over time; it also involves embracing reduction of some linguistic forms, expressions and patterns as a means of progression (Chau, 2015: 82)

Evidence from this research supports the findings of a constructive process and a reductionist process operating at the same time. The findings here indicate the complex process of language learning process. L2 development suggests two types of process: construction and reduction. The linguistic resources are in a state of flux (Ushioda, 1996; Ellis & Larsen-Freeman, 2006; Chau, 2015).

Recently, scholars like Larsen-Freeman (2015) suggest a new designation of the discipline, namely, Second Language Development. The word 'acquisition' is derived from the Latin word 'acquistionem', and it means the act of obtaining or getting, which implies that language is a commodity to be acquired, and to acquire a language is to get something once and for all (Larsen-Freeman, 2010, 2015). On the contrary, linguists and researchers gradually realize that language keeps changing, even the term target language seems misleading since the target is continually moving (Larsen-Freeman 1997). It is natural to regard language as a complex dynamic system rather than an object. Instead, language should be seen as an ever-developing resource, rather than a commodity to be obtained.

5.2 Implications of the study

Two main implications can be drawn from the present research, that is, 1) the difference between error and no-error analysis of learner data, and 2) the relationship between learners' linguistic system and learner corpus.

This study aims to describe how L2 development takes place over time, through the observation of the selected nouns (RIVER, GIRL, FRIEND, FLOWER and LAKE in both plural and singular forms) in the data, with the help of corpus linguistics.

In the field of SLA, scholars and practitioners look at language learner and learner language through different lenses. The lens this research has adopted is to respect language learners, and testify that learner language has a system in its own right. This research does not take the native speakers' language productions as a norm or a standard to measure the L2 development. It is hoped that this study can not only provide evidence of how learner language develops over time, but also help researchers look at learner language from a different perspective. Findings from this research provide a fuller picture of how a L2 develops, compared with those from the researches with methods like EA which focuses only on errors made by learners. In the practice of language teaching, meanwhile, teachers may realize that making 'errors' is a process of language learning, and it is suggested for teachers to create a comfortable environment for students to learn and use English, and downplay the role of 'errors' and 'mistakes' in language learning and teaching.

As mentioned many times in this research, learner language is a system in its own right (Selinker,1972). Corpus linguistics, a method based on the observation of naturally-occurring language, is an ideal way to describe the system.

5.3 Recommendations for future study

There are several recommendations for future study. First of all, different learner corpora are suggested to be used to investigate language development over time. The current research is set in a Malaysian context with secondary school students being the participants. Corpora collected in other countries with English as a L2 or as a foreign language can be used to explore how learner language develops across time. Meanwhile,

corpora with learners from different age groups are recommended to study how learners at different age groups develop their L2 competence in time.

Secondly, different word classes, such as verbs, adjectives, prepositions and adverbs, are recommended to be examined. This research takes nouns as a starting point to investigate language development, and certain types of noun patterns are observed to occur in the learner data. It will be interesting to see how different patterns would occur with different word classes, as time goes by.

Further, the data was analyzed by the research alone. Further study is recommended to involve a second rater/coder.

Last but not least, word forms are highly recommended to be used as units of analysis for future research. The current study focuses on the noun lemma, and certain types of patterns are yielded in the data. It is possible to see different results coming out of the different focuses .

5.4 Conclusion

This research explores how the use of the selected nouns changes over a period of 24 months in the texts contributed by Malaysian secondary school students. The finding shows that a total of six patterns of use of the nouns are identified, which are **det N**, **adj N**, **N prep n**, **N clause**, **N n** and **poss-N n**. The number of patterns of each noun varies

at different periods in time, which indicates a rather complex picture of the use of determiners, adjectives, prepositional phrases, and clauses. It is also found that students' lexical repertoire is expanding over time. The results show that both a constructive process and a reductionist process (Chau, 2015) operate at the same time, during language development.

Through the observation on the use of these selected words, it is realized that language development is a non-linear, complex process (e.g., Larsen-Freeman, 2006, 2009; Ortega, 2009; Ellis, 2011; Chau, 2015).

University of Malaya

REFERENCES

- Alsagoff, L. (2016). Interpreting error patterns in a longitudinal primary school corpus of writing. *The Asian Journal of Applied Linguistics*, 3(1), 114-124.
- Barlow, T. (2005). Computer-based analyses of learner language. In R. Ellis, & G. Barkhuizen, *Analysing learner language* (pp. 335-357). Oxford: Oxford University Press.
- Bley-Vroman, R. (1983). The comparative fallacy in interlanguage studies: the case of systematicity. *Language Learning*, 33(1), 1-17.
- Biber, D., Gray, B., & Poonpon, K. (2011). Should we use characteristics of conversation to measure grammatical complexity in L2 writing development? *TESOL Quarterly*, 45(1), 5-35.
- Biber, D., Johansson, S., Leech, G., Conrad, S., Finegan, E., & Quirk, R. (1999). *Longman grammar of spoken and written English*. London: Longman.
- Biber, D., & Reppen, R. (1998). Comparing native and learner perspectives on English grammar: A study of complement clauses. In S. Granger (ed.), *Learner English on computer* (pp. 145-158). London: Longman.
- Brians, P. (2008). *Common errors in English usage* (2nd ed.). Sherwood: William James & Co.
- Chau, M. H. (2015). *From language learners to dynamic meaning makers: A longitudinal investigation of Malaysian secondary school students' development of English from text and corpus perspectives*. Unpublished doctoral thesis, University of Birmingham.
- Cook, G. (1998). The uses of reality: A reply to Ronald Carter. *ELT Journal*, 52(1), 57-63.
- Cook, V. (1992). Evidence for multicompetence. *Language Learning*, 44(4), 557-591.
- Cook, V. (2010). The relationship between first and second language learning revisited. In E. Macaro (ed.), *The continuum companion to second language acquisition* (pp. 137-157). London: Bloomsbury.
- Cook, V. (2013). Multicompetence. In C. A. Chapelle (ed.), *The encyclopedia of applied linguistics*. Oxford: Wiley-Blackwell.
- Corder, S. P. (1967). The significance of learner's errors. *International Review of Applied Linguistics in Language Teaching*, 5(4), 161-170.
- Corder, S. P. (1974). Error analysis. In J. P. B. Allen, & S. P. Corder (eds.). *Techniques in applied linguistics*. London: Oxford University Press.
- Darus, S., & Subramaniam, K. (2009). Error analysis of the written English essays of secondary school students in Malaysia: A case study. *European Journal of Social Sciences*, 8(3), 483-495.

- Dash, N. S. (2005). *Corpus linguistics and language technology: With reference to Indian languages*. New Delhi: Mittal Publications.
- Edge, J. (1989). *Mistakes and correction*. London: Longman.
- Ellis, N. C. (2011). The emergence of language as a complex adaptive system. In J. Simpson (ed.), *Handbook of applied linguistics* (pp. 666-679). London: Routledge.
- Ellis, N. C., & Larsen-Freeman, D. (2006). Language emergence: Implications for applied linguistics-Introduction to the special issue. *Applied Linguistics*, 27(4), 558-589.
- Ellis, R. (2008). *The study of second language acquisition* (2nd edn.). Oxford: Oxford University Press.
- Ellis, R., & Barkhuizen, G. (2005). *Analysing learner language*. Oxford: Oxford University Press.
- Fitikides, T. J. (1989). *Common mistakes in English*. London: Longman.
- Francis, G., Hunston, S., & Manning, E. (1998). *Collins Cobuild Grammar Patterns 2: nouns and adjectives*. London: HarperCollins.
- Gass, S. M., & Selinker, L. (2008). *Second language acquisition: An introductory course* (3rd ed.). New York: Routledge.
- Gast, V. (2013). Contrastive analysis. In M. Byram & A. Hu (eds.), *The routledge encyclopedia of language teaching and learning* (pp. 153-158). London: Routledge.
- Granger, S. (1998). *Learner English on computer*. London: Addison Wesley Longman Limited.
- Granger, S. (2004). Computer learner corpus research: current status and future prospects. In U. Connor, & Th. A. Upton (eds.), *Applied corpus linguistics: A multidimensional perspective* (pp. 123-145). Amsterdam: Rodopi.
- Granger, S., & Paquot, M. (2009). Lexical verbs in academic discourse: A corpus-driven study of learner use. In M. Charles, S. Hunston, & D. Pecorari (eds.), *Academic writing: At the interface of corpus and discourse* (pp. 193-214). London: Continuum.
- Hardie, A. (2016). Corpus linguistics. In K. Allan (ed.), *The routledge handbook of linguistics* (pp. 502-515). New York: Routledge.
- Hendrickson, J. M. (1978). Error correction in foreign language teaching: Recent theory, research, and practice. *The Modern Language Journal*, 62(8), 387-398.
- Hornby, A. S. (1954). *A guide to patterns and usage in English*. London: OUP.
- Hunston, S. (2002). *Corpora in applied linguistics*. Cambridge: Cambridge University

Press.

- Hunston, S., & Francis, G. (1999). *Pattern Grammar: A corpus-driven approach to the lexical grammar of English*. Amsterdam: John Benjamins.
- Hmouma, M. A. A. (2014). Interlanguage syntax of the noun phrase in English: An error analysis of third year high school students of English in Zawiya, Libya. *IOSR Journal Of Humanities And Social Science*, 19 (5), 36-43.
- Kennedy, G. (2003). *Structure and meaning in English: A guide for teachers*. Harlow: Pearson Education.
- Kalmos, B. (2011). Language learner errors and dyslexia: a tale of unlikely parallels. *Working Papers in Language Pedagogy*, 5, 1-15.
- Lan, G. (2014). Investigating the relationship between second language writing proficiency and noun modifications. *US-China Foreign Language*, 13(5), 363-373.
- Larsen-Freeman, D. (1997). Chaos/complexity science and second language acquisition. *Applied Linguistics*, 18(2), 141-165.
- Larsen-Freeman, D. (2006). The emergence of complexity, fluency, and accuracy in the oral and written production of five Chinese learners of English. *Applied Linguistic*, 27(4), 590-619.
- Larsen-Freeman, D. (2009). Adjusting expectations: The study of complexity, accuracy, and fluency in second language acquisition. *Applied Linguistics*, 30(4), 579-589.
- Larsen-Freeman, D. (2010). Having and doing: Learning from a complexity theory perspective. In P. Seedhouse, S. Walsh, & C. Jenks (eds.), *Conceptualising learning in applied linguistics* (pp. 52–68). Basingstoke: Palgrave Macmillan.
- Larsen-Freeman, D. (2011). The emancipation of language learner. *Studies in Second Language Learning and Teaching*. 2(3), 297-309.
- Larsen-Freeman, D. (2015). Saying what we mean: Making a case for ‘language acquisition’ to become ‘language development’. *Language Teaching*, 48(04), 491-505.
- Levitzky-Aviad, T. (2012). Lexical richness and variation in the writing of school-age EFL learners at different learning stages and different educational systems. In Y. Tono, Y. Kawaguchi, & M. Minegishi (eds.), *Developmental and crosslinguistic perspectives in learner corpus research* (pp.159-168). Amsterdam: John Benjamins.
- Li, S. (2015). *A corpus-based study of the high frequency nouns ‘time’ and ‘thing’: Investigating the role of phraseology in the construction of meaning in discourse*. Unpublished doctoral dissertation, University of Birmingham.
- Lozano, C., & Mendikoetxea, A. (2013). Learner corpora and Second Language Acquisition: The design and collection of CEDEL2. In A. Díaz-Negrillo, N.

- Ballier & P. Thompson (eds.), *Automatic treatment and analysis of learner corpus data* (pp. 65-100). Amsterdam: John Benjamins.
- Mackey, A. (1999). Input, interaction and second language development: An empirical study of question formation in ESL. *Studies in Second Language Acquisition*, 21(4), 557-587.
- Mahlberg, M. (2005). *English general nouns*. Amsterdam: John Benjamins.
- Mamelina, L (2013). The process of second language acquisition and its implications for teaching. *Scientific Contributions*, 17, 149-153.
- Maros, M., Tan, K. H., & Salehuddin, K. (2007). Interference in learning English: Grammatical errors in English essay writing among rural Malay secondary school students in Malaysia. *BANGI: Jurnal Sains Sosial dan Kemanusiaan*, 2(2), 15.
- Martin, N. J. L. (2009). *A corpus-based study of the phraseological behaviour of abstract nouns in medical English*. Paper presented at the Corpus Linguistics Conference, Liverpool, UK.
- McEnery, T., & Hardie, A. (2011). *Corpus linguistics: method, theory and practice*. Cambridge: Cambridge University Press.
- Mitchell, R., & Myles, F. (2004). *Second language learning theories* (2nd ed.). Oxford: Oxford university.
- Mitchell, R., Myles, F., & Marsden, E. (2013). *Second language learning theories*. New York: Routledge.
- Mousavi, A., & Rauof Moini, M. (2014). A corpus study of shell nouns in published research articles of education. *Social and Behavioral Sciences*, 98, 1282-1289.
- Myles, F. (2005). Interlanguage corpora and second language acquisition research. *Second Language Research*, 21, 373-391.
- Myles, F. (2015). Second language acquisition theory and learner corpus research. In S. Granger, G. Gilquin, & F. Meunier, (eds.), *The Cambridge handbook of learner corpus research* (pp. 307-308). Cambridge: Cambridge University Press.
- Nation, I. S. (2001). *Learning vocabulary in another language*. Cambridge: Cambridge University Press.
- Ortega, L., & Byrnes, S. (2008). *The longitudinal study of advanced L2 capacities*. New York: Routledge.
- Ortega, L. (2009). *Understanding Second Language Acquisition*. London: Hodder Arnold.
- Ortega, L. (2013). SLA for the 21st century: Disciplinary progress, transdisciplinary relevance, and the bi/multilingual turn. *Language Learning*, 63(1), 1-24.

- Parkinson, J., & Musgrave, J. (2014). Development of noun phrase complexity in the writing of English for academic purposes students. *Journal of English for Academic Purposes, 14*(2), 48-59.
- Pica, T. (1983). Methods of morpheme quantification: Their effect on the interpretation of second language data. *Studies in Second Language Acquisition, 6*(1), 69-78.
- Pienemann, M., & Johnston, M. (1987). Factors influencing the development of language proficiency. In D. Nunan (ed.), *Applying second language acquisition research* (pp. 45-141). Adelaide: National Curriculum Resource Centre.
- Quirk, R., Greenbaum, S., Leech, G., Svartvik, J., & Crystal, D. (1985). *A comprehensive grammar of the English language*. London: Longman.
- Rif'ah, F. N., & Sulisty, G. H. (2004). *An obligatory occasion analysis of the use of past regular and past irregular forms in recount texts of VIII graders of SMP1 Bogonegoro*. Master's thesis, State University of Malang.
- Rustipa, K. (2011). Contrastive analysis, error analysis, interlanguage and the implication to language teaching. *Ragam Jurnal Pengembangan Humaniora, 11*(1), 16-22.
- Saville-Troike, M. (2012). *Introducing second language acquisition*. Cambridge: Cambridge University Press.
- Selinker, L. (1972). Interlanguage. *International Review of Applied Linguistics, 10*(4), 209-232.
- Snedeker, J., Geren, J., & Shafto, C. L. (2012). Disentangling the effects of cognitive development and linguistic expertise: A longitudinal study of the acquisition of English in internationally-adopted infants and preschoolers. *Cognitive Psychology, 65*(1), 39-76.
- Sinclair, J. M. (1991). *Corpus, concordance, and collocation*. Oxford: Oxford University Press.
- Sinclair, J. M. (1998). Lexical item. In E. Weigand (ed.), *Contrastive lexical semantics* (pp. 1-24). Amsterdam: John Benjamins.
- Stefansson, E. G. (2013). *Second language acquisition: The effect of age and motivation*. Master's thesis, University of Iceland.
- Ushioda, E. (1996). Developing a dynamic concept of motivation. In T. Hickey, & J. Williams (eds.), *Language education and society in a changing world* (pp. 239-245). Dublin: Multilingual Matters.
- Yunus, M. M., Sulaiman, A., Kamarulzaman, M. H., & Ishak, N. M. (2013). Language learning difficulties among Malaysian gifted students. *Asian Social Science, 9*(15), 130.
- Zhao, S. (2015). An error analysis of English attributive clauses by non-English major college students. *Advances in Social Science, Education and Humanities*

Research, 15, 1005-1010.

University of Malaya

Appendix A

Coding of the patterns

Codes are adopted from Hunston and Francis (1999: 45) and additional codes used in the current study are from Francis, Hunston, and Manning (1996, 1998)

v:	verb group
n:	noun group
adj:	adjective group
adv:	adverb group
that	clause introduced by <i>that</i> (realized or not)
which	clause introduced by <i>which</i> (realized or not)
-ing	clause introduced by an "-ing" form
to-inf:	clause introduced by to-infinitive form
wh:	clause introduced by a wh-word (including <i>how</i>)
with quote:	used with direct speech
det:	a determiner
prep:	a prepositional phrase

Appendix B

Noun patterns from Cobuild Collins grammar patterns 2

Nouns with determiners and modifiers		Noun patterns with <i>be</i>	
<i>a</i> N	n N	<i>the</i> N <i>be</i> that	<i>it</i> v-link poss N <i>that</i>
<i>the</i> N	num N	poss N <i>be</i> that	<i>it</i> v-link det N to-inf
poss N	ord N	<i>the</i> N <i>be</i> to-inf	<i>it</i> v-link N to-inf
adj N		poss N <i>be</i> to-inf	<i>it</i> v-link poss N to-inf
Nouns followed by a clause		<i>the</i> N <i>be</i> -ing	<i>it</i> v N to-inf
N <i>that</i>	N to-inf	Noun patterns with <i>there</i>	
Nouns followed by a preposition		<i>there be</i> det N about n	<i>there be</i> det N to n/-ing
N <i>about</i> n	N <i>in</i> n	<i>there be</i> det N <i>in</i> n/-ing	
N <i>against</i> n	N <i>in favour of</i> n	Nouns following a preposition	
N <i>among</i> pl-n	N <i>into</i> n	<i>at</i> N	<i>out of</i> N
N <i>as</i> n	N <i>of</i> n	<i>by</i> N	<i>to</i> N
N <i>as to</i> n	N <i>on</i> n	<i>from</i> N	<i>under</i> N
N <i>behind</i> n	N <i>over</i> n	<i>in</i> N	<i>with</i> N

Appendix C

The Prompt Used in Data Collection

