

The Effectiveness of Focused Instruction of Formulaic Sequences in Augmenting L2 Learners'
Academic Writing Skills: A Quantitative Research Study

by

Lina AlHassan

A dissertation submitted to the Faculty of Graduate and Postdoctoral Affairs
in partial fulfillment of the requirements for the degree of

Master

in

Applied Linguistics and Discourse Studies

Carleton University

Ottawa, Ontario

© 2014

Lina AlHassan

Abstract

Research has increasingly focused on the effectiveness of formulaic sequences in augmenting second language (L2) learners' academic writing skills. These formulaic sequences, which may constitute as much as 52.3% of written discourse (Erman & Warren, 2000), play a vital role in improving L2 learners' writing proficiency and enhancing their performance in academic contexts (Jones & Haywood, 2004; Lewis, 1997). However, experimental testing of such a role is largely undeveloped in research. This quantitative research study is an attempt to investigate the effects of explicitly teaching formulaic sequences on twelve L2 learners' academic writing skills. The study results suggest that an explicit instructional approach to formulaic sequences can enhance their subsequent acquisition. Moreover, formulaic sequences increase L2 learners' writing proficiency because they function as frames to which L2 learners resort when approaching a writing task to compose an academic piece of writing.

Keywords: Formulaic sequences, focused instruction, L2 writing proficiency

Acknowledgements

I would like to express my deepest gratitude to the dissertation supervisor, Professor David Wood. Without his solid guidance, invaluable suggestions, and constructive feedback, this dissertation would not have been completed. I am also grateful to Professor Janna Fox for her continuous support and valuable advice in statistics. As well, I would like to thank Professor Ellen Cray for her valuable assistance with the use and interpretation of the Corpus of Contemporary American English.

I am also grateful to Estelle Lalande, Don Myles, and Farzad Hashemi for their willingness to evaluate the data for this study despite the busy schedules to which they were committed. Special thanks are to Joan Grant and Connie Wall for their assistance in securing a place for conducting my research study. I would also like to thank my participants for volunteering in this study.

I owe my deepest appreciation to my mother, Isaaf Mokayed, and my siblings for their selfless support and continuous encouragement despite the atrocities of war they have been experiencing on a daily basis for three years in my beloved hometown, Aleppo. I am also indebted to my husband, Hassan, for his unflagging support and encouragement.

Toward the end, I dedicate this thesis to the spirit of my father, Fathi AlHassan, who imparted an endless love and quest for knowledge and learning to me and has always been the boundless source of encouragement, persistence, and inspiration throughout my life.

Table of Contents

Abstract	ii
Acknowledgements	iii
List of Tables	vi
List of Appendices	ix
Chapter One: Introduction	1
Chapter 2: Literature Review	6
Formulaicity of Spoken and Written Language: An Overview	7
The Nature, Types, and Functions of Formulaic Sequences	13
The holistic nature of formulaic sequences	13
Types of formulaic sequences.....	16
Functions of formulaic sequences.....	24
Formulaic Sequences and Academic Writing Skills.....	27
Formulaic Sequences in the Classroom	34
The lexical approach to language teaching.....	34
Incorporating formulaic sequences in classroom materials	35
Focused instruction of formulaic sequences	36
Classroom activities	39
Enriching L2 learners’ linguistic repertoire with formulaic sequences	39
Augmenting L2 learners’ academic writing skills	42
Summary of Chapter Two.....	44
Chapter 3: Motivation for the Study	46
Chapter 4: Methodology	50
Participants.....	51
Instruments.....	54
Instruments used for data collection	54
Instruments used during the training period	56

Selection of Formulaic Sequences	61
Procedure	68
Data Analysis	71
Variables	78
Chapter Five: Findings	84
Quantitative Content Analysis of Data	84
Descriptive Statistics.....	88
Cronbach's Alpha	94
Paired-Samples T-Tests	95
Correlation Coefficients.....	104
Chapter Six: Discussions of the Quantitative Analysis Results	108
Focused Instruction of Formulaic Sequences	108
Focused Instruction of Formulaic Sequences and L2 Writing Proficiency	113
Formulaic Sequences and L2 Learners' Better Achievement.....	116
The Use of Formulaic Sequences as Zones of Safety as well as Links	118
Chapter Seven: Conclusion	123
Quantitative Results	123
Formulaic Sequences and Language Teaching.....	123
Limitations	126
Conclusion and Suggestions for Future Research	128
References	131

List of Tables

Table 1: A Predefined Categories List.....	72
Table 2: A Predefined Categories List.....	73
Table 3: The Frequency of Formulaic Sequences.....	85
Table 4: The Occurrence of Formulaic Sequences.....	87
Table 5: Descriptive Statistics: The Frequency of Formulaic Sequences	89
Table 6: Descriptive Statistics: The Occurrence of Formulaic Sequences.....	90
Table 7: Descriptive Statistics: The Evaluation of the First Judge.....	91
Table 8: Descriptive Statistics: The Evaluation of the Second Judge	92
Table 9: Descriptive Statistics: The Evaluation of the Third Judge	93
Table 10: Reliability Statistics.....	94
Table 11: Item-Total Statistics.....	94
Table 12: Paired Samples Statistics: The Frequency of Formulaic Sequences	95
Table 13: Paired Samples Test: The Frequency of Formulaic Sequences.....	96
Table 14: Paired Samples Statistics: The Occurrence of Formulaic Sequences.....	97
Table 15: Paired Samples Test: The Occurrence of Formulaic Sequences	97
Table 16: Paired Samples Statistics: The First Judge’s Evaluation.....	99
Table 17: Paired Samples Test: The First Judge’s Evaluation	99
Table 18: Paired Samples Statistics: The Second Judge’s Evaluation	100
Table 19: Paired Samples Test: The Second Judge’s Evaluation.....	101
Table 20: Paired Samples Statistics: The Third Judge’s Evaluation	102
Table 21: Paired Samples Test: Third Judge’s Evaluation.....	102
Table 22: A Correlation Matrix: Formulaic Sequences and the First Judge’s Evaluation	104
Table 23: A Correlation Matrix: Formulaic Sequences and the Second Judge’s Evaluation.....	105
Table 24: A Correlation Matrix: Formulaic Sequences and the Third Judge’s Evaluation.....	106
Table A1: The First Judge’s Evaluation	146
Table A2: The Second Judge’s Evaluation.....	147
Table A3: The Third Judge’s Evaluation.....	148
Table C1: The Formulaic Sequences Extracted from the COCA.....	169
Table C2: The Frequency of the Stem Collocations in the COCA.....	170
Table C3: The Collocations Extracted from the COCA.....	173
Table D1: Descriptive Statistics: The Mean Difference of Frequency.....	177

Table D2: Tests of Normality: The Mean Difference of Frequency	178
Table D3: Descriptive Statistics: The Mean Difference of Occurrence	180
Table D4: Tests of Normality: The Mean Difference Occurrence	181
Table D5: Descriptive Statistics: The Mean Difference of the First Judge's Evaluation.....	182
Table D6: Tests of Normality: The Mean Difference of the First Judge's Evaluation.....	184
Table D7: Descriptive Statistics: The Mean Difference of the Second Judge's Evaluation	185
Table D8: Tests of Normality: The Mean Difference of the Second Judge's Evaluation	186
Table D9: Descriptive Statistics: The Mean Difference of the Third Judge's Evaluation	187
Table D10: Tests of Normality: The Mean Difference of the Third Judge's Evaluation.....	189
Table D11: Descriptive Statistics: The First Judge's Evaluation	190
Table D12: Descriptive Statistics: The Second Judge's Evaluation.....	191
Table D13: Descriptive Statistics: The Third Judge's Evaluation.....	192
Table E1: Manual Coding: Salem's Pretest	194
Table E2: Manual Coding: Salem's Posttest	195
Table E3: Manual Coding: Salem's Delayed Posttest	196
Table E4: Manual Coding: Ellie's Pretest	197
Table E5: Manual Coding: Ellie's Posttest.....	198
Table E6: Manual Coding: Ellie's Delayed Posttest.....	199
Table E7: Manual Coding: David's Pretest	201
Table E8: Manual Coding: David's Posttest.....	202
Table E9: Manual Coding: David's Delayed Posttest	203
Table E10: Manual Coding: Mary's Pretest	205
Table E11: Manual Coding: Mary's Posttest.....	206
Table E12: Manual Coding: Mary's Delayed Posttest	207
Table E13: Manual Coding: Eva's Pretest.....	209
Table E14: Manual Coding: Eva's Posttest	210
Table E15: Manual Coding: Eva's Delayed Posttest.....	212
Table E16: Manual Coding: Ali's Pretest.....	213
Table E17: Manual Coding: Ali's Posttest	214
Table E18: Manual Coding: Ali's Delayed Posttest.....	216
Table E19: Manual Coding: MATA's Pretest	218
Table E20: Manual Coding: MATA's Posttest.....	219
Table E21: Manual Coding: MATA's Delayed Posttest	221

Table E22: Manual Coding: Jacob's Pretest.....	223
Table E23: Manual Coding: Jacob's Posttest	224
Table E24: Manual Coding: Jacob's Delayed Posttest.....	227
Table E25: Manual Coding: Moo's Pretest	229
Table E26: Manual Coding: Moo's Posttest.....	230
Table E27: Manual Coding: Moo's Delayed Posttest.....	231
Table E28: Manual Coding: Gabriella's Pretest	232
Table E29: Manual Coding: Gabriella's Posttest.....	233
Table E30: Manual Coding: Gabriella's Delayed Posttest	235
Table E31: Manual Coding: Shireen's Pretest.....	236
Table E32: Manual Coding: Shireen's Posttest	237
Table E33: Manual Coding: Shireen's Delayed Posttest.....	238
Table E34: Manual Coding: Abedeer's Pretest	240
Table E35: Manual Coding: Abedeer's Posttest.....	241
Table E36: Manual Coding: Abedeer's Delayed Posttest.....	243

List of Appendices

Appendix A: The Three Judges' Evaluation	146
Appendix B: Instruments	149
Prompt Used for Data Collection.....	149
Worksheet 1: Topic Sentences.....	150
Worksheet 2: Concluding Sentences	152
Worksheet 3: List of Collocations	155
Worksheet 4: Collocation Practice.....	156
Worksheet 5: Formulaic Sequences.....	157
Worksheet 6: Practice	160
Worksheet 7: Formulaic Sequences.....	162
Worksheet 8: Practice	164
Worksheet 9: A Table of Formulaic Sequences	165
Worksheet 10: Practice Sheet	166
Worksheet 11: A Cloze Activity.....	167
Worksheet 12: Full Text for Dictogloss.....	168
Appendix C: Formulaic Sequences Extracted from the COCA	169
Appendix D: Tests of Normality	177
Tests of Normality for Paired-Samples T-Tests	177
Tests of Normality for Correlation Coefficients.....	190
Appendix E: Quantitative Content Analysis of the Elicited Texts	194
Salem's Textual Data.....	194
Ellie's Textual Data	197
David's Textual Data	201
Mary's Textual Data	205
Eva's Textual Data.....	209
Ali's Textual Data.....	213
MATA's Textual Data	218
Jacob's Textual Data.....	223

Moo's Textual Data	229
Gabriella's Textual Data.....	232
Shireen's Textual Data.....	236
Abedeer's Textual Data	240

Chapter One

Introduction

As early as the 1980s, a number of linguists were arguing against the Chomskyan approach, which hypothesizes that what comprises any natural language is a set of infinite utterances that are generated based on syntactic rules (Pawley & Syder, 1983). To Barlow (2000), Liu and Huo (2011), and Wray (2000), language users tend to repeatedly employ particular utterances in their language production and disregard all other possible grammatical expressions that can be as useful to express the same concept. This growing interest in what is known to be the formulaicity of language production is linked to a burgeoning of new linguistic theories which focus on performance rather than on competence per se and highlight the significant role of formulaic sequences in language production.

Moreover, the emergence of usage-based theory and corpus linguistics have added further evidence in support of the importance of formulaicity in language production (Barlow, 2000; Beckner et al., 2009; Cortes, 2004; Weinert, 1995). The formulaic nature of language production has been observed by several scholars with reference to native speakers' preference for certain utterances over others when they tend to convey their messages in everyday communication (Ellis, 1996; Howarth, 1998; Lewis, 2000a; Wray, 1999). These utterances are referred to as formulaic sequences, which can be defined as prefabricated chunks that are stored in and retrieved from the memory as wholes (Chen, 2009; Ding, 2007; Wood, 2006; Wray & Fitzpatrick, 2008). Furthermore, the corpus-based analysis of language use has demonstrated that formulaic sequences make up a large portion of discourse, particularly academic discourse, and, hence, they have been viewed as essential constituents of proficient academic skills due to their

various functions in discourse (Conklin & Schmitt, 2008; Cortes, 2004; Coxhead & Byrd, 2007; Lewis, 1997).

The ubiquity of formulaic sequences in L1 discourse requires that formulaic sequences be considered a cornerstone of any full account of second language acquisition and production (Boers & Lindstromberg, 2009; Jones & Haywood, 2004; Schmitt & Underwood, 2004). Experts in the field of second language acquisition have argued that formulaic sequences constitute a major part of L2 learners' linguistic repertoire, and, thus, they are believed to play a vital role in L2 learners' productive skills, mainly speech fluency and pragmatic competence, among other aspects of proficiency (Fitzpatrick, 2005; Osborne, 2008; Wood, 2009a; Wray, 2002; Wray & Perkins, 2000).

In addition, formulaic sequences have been viewed as essential building blocks of proficient academic writing, a complex process for many L2 learners whose expressive skills are underdeveloped (Hyland, 2003, 2006; Leki, 2006; Reppen, 2002; Silva, 1993). The importance of formulaic sequences in promoting L2 writing skills can be attributed to the fact that they represent the frequent expressions with which L2 writers have to deal in the academic register in order to be defined as proficient writers (Coxhead & Byrd, 2007; Martinez & Schmitt, 2012). As Lewis (1997) puts it, academic writers demonstrate a tendency to frequently use expressions that are peculiar to academic prose.

This view of academic prose, as replete with formulaic sequences, implies that academic writing is a complex process that merges the mastery of grammar and lexicon with the concise knowledge of the academic genres as well as the vocabulary typical of the field in question (Coxhead & Byrd, 2007). More importantly, it has been suggested that the failure to employ formulaic sequences in academic writing may not only affect the writing quality, but also the

overall evaluation of L2 students' performance in academic contexts (Cortes, 2004; Lewis, 1997).

Due to their vital role in proficient language production, formulaic sequences have been identified by several scholars as necessitating a major component of classroom materials (Li & Schmitt, 2009; Nation & Newton, 1997; Nattinger & DeCarrico, 1992; Sinclair, 1991). It has been also suggested that teachers should be selective in their choice of the target formulaic sequences (Boers & Lindstromberg, 2009; Lewis, 1997). In other words, language teachers should choose the formulaic sequences that best account for their students' needs in the target context. They also advocate an explicit teaching approach to this language phenomenon which, although very frequent in discourse, is only partially transparent or totally opaque to many L2 learners (Erman & Warren, 2000; Jones & Haywood, 2004).

It is interesting to note that although this role has been intensively investigated in regards to speech fluency and pragmatic competence, the tendency to view formulaic sequences as a keystone of L2 learners' proficient academic writing has not yet been empirically examined to any great extent.

In this respect, to highlight the effectiveness of formulaic sequences in augmenting L2 learners' academic writing skills, the present study aims to explore how the mastery of different types of formulaic sequences can help L2 learners manipulate academic discourse in academic reports. The study was conducted over a ten-week period with twelve participants from four different L1 backgrounds. It involved the participants in focused instruction on formulaic sequences for ninety minutes a week. Every week, a number of formulaic sequences were presented to the participants, and the presentation stage was followed by practice and production stages. Three timed written paragraphs were elicited from each participant by using the same

prompt, a line graph that depicts the changes in commodity prices over one year. The textual data were collected at three different points in time: at the beginning of the study before any instruction (a pretest); at the end of the study (a posttest); and fifteen to twenty days after the end of the study (a delayed posttest).

The elicited texts were evaluated blindly by three judges, and the textual data were analyzed based on quantitative content analysis. The raw scores, i.e., the variables derived based on quantitative content analysis along with the three judges' evaluation, were analyzed statistically by computing paired-samples t-tests to highlight any significant differences in the frequency and occurrence of the target formulaic sequences and the judges' evaluation before and after the training period. In addition to paired-samples t-tests, Cronbach's alpha was computed to investigate the inter-rater reliability of the three judges' evaluations. Pearson product-moment correlation coefficients were also run to investigate any possible associations between the frequency and occurrence of the target formulaic sequences in the elicited texts and the evaluation of the three judges. Other statistical tests were conducted, including measures of central tendency, measures of variability, and tests of normality, to summarize the raw scores and investigate the distribution of the dependent variables.

This study is divided into seven chapters. Each of these chapters, except Chapter One, opens with an introduction that introduces the reader to its ultimate aim and ends with a summary that sums up its contents. The first chapter, which is the introduction to this thesis, encapsulates the content of this dissertation by briefly highlighting the gap in literature that has motivated this study and providing insights into its design. The second chapter is entirely devoted to examining the literature related to the overarching theme of this research study: investigating the effectiveness of the focused instruction of formulaic sequences in augmenting L2 learners'

academic writing skills. Thus, in order to introduce the reader to the various theories, approaches, and methods presented by experts in the field of second language acquisition and discourse analysis, the second chapter falls into five major sections which tackle the formulaicity of language production, the nature, types, and functions of formulaic sequences, the relationship between formulaic sequences and writing proficiency, and the various approaches to teaching formulaic sequences.

The third chapter defines the gap in literature and includes the research questions that have guided this study. The fourth chapter, in turn, provides the reader with a detailed description of the empirical study designed to answer the research questions. Thus, it is divided into six sections that offer a description of the participants investigated in this study, the instruments used during the training period, the methods related to the selection of the target formulaic sequences, the procedure followed in this study, the coding method applied to the textual data, the statistical tests computed on the raw scores that were derived based on manual coding and elicitation instruments, and the variables tested in this study.

Chapter Five presents the findings of the manual quantitative coding method of the textual data and the results of several statistical tests, namely, measures of central tendency and variability, Cronbach's alpha, paired-samples t-tests, and Pearson product-moment correlation coefficients. The sixth chapter offers an in-depth interpretation of the results presented in Chapter Five and aims at tracing the connection between the findings of this study and the theories and discussions presented in Chapter two. The last chapter concludes this thesis by reiterating the major findings of this study and highlighting their implications. It also highlights the limitations of the study and provides some directions for future research.

Chapter 2

Literature Review

The present chapter offers a synthesis of research on formulaic sequences to cast light on their significant role in second language acquisition and production in various contexts, particularly academic contexts. With this goal as a guide, this theoretical background is divided into five major sections: Formulaicity of Spoken and Written Language; The Nature, Types, and Functions of Formulaic Sequences; Formulaic Sequences and Academic Writing Skills; Formulaic Sequences in the Classroom; Summary of Chapter Two.

The first section in this chapter investigates the recent growth of the theoretical interest in the formulaicity and/or idiomaticity of language production from the perspectives of both first and second language acquisition. It then moves on to highlight the way in which this language phenomenon, which has been long overlooked in favor of the rule-governed, creative nature of language, has been brought to the forefront with the emergence of usage-based theory and corpus linguistics. The second section explores the holistic nature of formulaic sequences, which is thought to be particularly important for enhancing L2 learners' speech fluency and proficiency. It also outlines the phraseological as well as distributional classifications of formulaic sequences and their functions in both spoken and written discourse as proposed by prominent figures in the field of second language acquisition.

The third section, in turn, presents an overview of writing skills in general, and academic writing in particular. Moreover, it provides an intensive account of all relevant literature on the significant role of formulaic sequences in upgrading L2 learners' expressive skills. It also explores the major challenges that result from L2 learners' difficulty in using the formulaic sequences peculiar to academic discourse. The fourth section presents insights into the lexical approach to language teaching and the importance of making formulaic sequences salient to L2

learners through focused instruction. It concludes with some sample activities designed to help the successful acquisition and internalization of formulaic sequences into L2 learners' linguistic repertoire. The chapter concludes, in the fifth section, with a brief summary of the key discussions, which function as an incentive and a profound stimulus for conducting this exploratory study. It also paves the way for the research gap and questions that will be identified in the third chapter.

Formulaicity of Spoken and Written Language: An Overview

It has been generally accepted that language is not entirely creative, rule-governed, or composed afresh each time it is produced (Barlow, 2000; Boers, Eyckmans, Kappel, Stengers, & Demecheleer, 2006; Conklin & Schmitt, 2008; Ellis & Simpson-Vlach, 2009; Hyland, 2011; Miller & Weinert, 1998; Williams, 2008; Wood, 2001; Wray, 2002). That is, despite the creative nature of language that theoretically enables them to produce and interpret an infinite set of utterances, language users tend to employ particular utterances in their language production and disregard all other possible grammatical expressions that would seem to be equally useful to express the same concept (Ellis, 1996; Lewis, 1997, 2000a; Liu & Huo, 2011; Oakey, 2002; VanPatten & Williams, 2007; Wray, 2000). In this respect, with the burgeoning of new linguistic theories, including usage-based theory, which have mainly focused on performance rather than on competence per se, the concept of creativity of language has been situated over time and has been augmented, at least partially, by what is known to be the formulaicity and/or idiomaticity of language production (Boers et al., 2006; Lewis, 1997; Pawley & Syder, 1983; Weinert, 1995).

As early as the 1980s, a considerable number of linguists began to question the Chomskyan approach, which hypothesizes that what comprises any natural language is a set of infinite utterances that are generated based on syntactic rules. Among those scholars are Pawley and

Syder (1983) who provide insights into the formulaic nature of language in its widest sense. In their analysis of the language used by English speakers, Pawley and Syder (1983) claim that English native speakers do not exploit the creative power of syntax in their everyday communication, but they rather tend to employ utterances that are natively like in form and, hence, disregard all other utterances that might be considered 'odd' despite their grammatical accuracy:

Native speakers do not exercise the creative potential of syntactic rules to anything like their full extent, and that, indeed, if they did do so they would not be accepted as exhibiting natively like control of the language. The fact is that only a small proportion of the total set of grammatical sentences are natively like in form— in the sense of being readily acceptable to native informants as ordinary, natural forms of expression, in contrast to expressions that are grammatical but are judged to be 'unidiomatic', 'odd' or 'foreignisms.' (p. 193)

The formulaic nature of language use has also been a focus as a result of the growing interest in the usage-based theory of language which focuses on performance, what language users actually do with the language, rather than on competence, what they are capable of doing (Beckner et al., 2009; Weinert, 1995). According to this theory, corpus-based analyses of language use “verify that communication largely consists of prefabricated sequences, rather than an ‘open choice’ among all available words” (Beckner et al., 2009, p. 6).

This hypothesis has been empirically demonstrated with reference to language users' choice of utterances, speech fluency, articulatory patterns in speech production, and historical changes in language use. The assumption that language production is at least partially formulaic rather than being entirely constructed from scratch by virtue of grammar rules is manifested in language users' obvious tendency to bypass all generative rules available in language and repeatedly deploy the utterances common in their speech community (Beckner et al., 2009;

Weinert, 1995). To Beckner et al. (2009), “speakers do not choose randomly from among all conceivable combinatorial possibilities when producing utterances. Rather there are conventional ways of expressing certain ideas” (p. 6). In addition, speech fluency reflects the formulaic nature of language production. Taking into consideration the limited storage capacity and the constraints of the short-term memory, “a plausible explanation for the fluency is that speakers rely on ready-made memorized sequences in language use rather than novel generation from rules” (Weinert, 1995, p. 85).

The degree of phonological reduction, the tendency to reduce frequently co-occurring words, in speech production also supports the formulaic nature of language production (Beckner et al., 2009; Gregory, Raymond, Bell, Fosler-Lussier, & Jurafsky, 1999). The fact that language users produce highly frequent co-occurring words with a degree of phonological reduction suggests that frequent word combinations are stored in and retrieved from the long-term memory as chunks rather than being creatively constructed every time they are produced (Beckner et al., 2009; Gregory et al., 1999). Last but not least, the historical changes in language can be also viewed in relation to the formulaic nature of language. These changes are represented in English contractions which were developed from the amalgamation of frequently recurring words and modal auxiliaries that have become fused with their most frequent collocates (Beckner et al., 2009). In this respect, usage-based theory has given equal weight to both productive rules and unanalyzed chunks as opposed to the long prevailing views of language as being entirely creative (Beckner et al., 2009; Weinert, 1995).

Further evidence has been added to the formulaic nature of language use as a result of the remarkable emergence of corpus linguistics (Barlow, 2000; Cortes, 2004; Coxhead & Byrd, 2007), a landmark in applied linguistics which

... aims to explore the extent to which certain features of language use are associated with contextual factors ... to uncover characteristic patterns of language use and to generalize from the collected texts to other texts of a similar type, or to the language as a whole. (Hall, Smith, & Wicaksono, 2011, p. 79)

Corpus linguistics provides tools and techniques that analyze language use in order to identify the meaning and frequency of its constituents based on their actual use in natural contexts (Barlow, 2000; Biber, Conrad, & Repper, 1998). At this point, corpus-based analysis of language production has demonstrated that “much of language in use is not creative in the Chomskyan sense, but it is based mainly on the use of prefabricated or semi-fabricated chunks” (Barlow, 2000, p. 318). Hence, instead of creating utterances from scratch, language users deploy grammatical rules in order to amalgamate prefabricated chunks and transform them into meaningful utterances which they can repeatedly utilize in their language production (Barlow, 2000; Lewis, 1997).

The aforementioned arguments have been validated by several scholars with reference to native speakers' preference of certain utterances over others when they tend to convey their messages in everyday communication. For example, despite the fact that marriage proposals can be fulfilled using such utterances like ‘I wish to be wedded to you’ or ‘My becoming your spouse is what I want,’ English speakers usually express their wishes with the familiar utterance ‘I want to marry you’ (Pawley & Syder, 1983). Not only do language users deploy familiar utterances when proposing to their partners, but they also resort to such utterances in most of their spontaneous communication that is part and parcel of their everyday lives (Conklin & Schmitt, 2008; Ellis, 1996; Howarth, 1998; Lewis, 1997, 2000c; Pawley & Syder, 1983; Wray, 1999, 2002).

In addition to their high frequency in spoken language, formulaic sequences, as demonstrated in corpus-based analysis of language use, are also widespread in written discourse, particularly academic writing (Biber, Conrad, & Cortes, 2004; Conklin & Schmitt, 2008; Ellis, Simpson-Vlach, & Maynard, 2008; Howarth, 1998; Li & Schmitt, 2009). Paquot (2008), for example, maintains that corpus-based analysis of language use has revealed the nature of academic discourse as marked by the existence of prefabricated sequences which have essential syntactic and semantic functions, especially in writing. Coxhead (2008), in turn, states that academic discourse is characterized by a high frequency of particular multiword sequences which might not be as frequent in other registers. Moreover, Li and Schmitt (2009) note that “it has been generally agreed that formulaic sequences like ‘*as a result*’ and ‘*it should be noted that*’ are central to the creation of academic texts” (p. 86). Likewise, Simpson-Vlach and Ellis (2010) state that formulaic sequences are salient entities which have significant functions in written academic language due to the fact that “the language necessary for proficiency in academic contexts is quite different from that required for basic interpersonal communicative skills” (p. 487).

Lewis (1997), in turn, presents an extreme view on the formulaic nature of language production and argues that language is no longer seen as divided into grammar and vocabulary since such a view of language “has been challenged and shown to be seriously misguided from both strictly linguistic and pedagogical points of view” (p. 255). To him, what constitutes language, whether spoken or written, are four types of lexical items that range from ‘absolutely fixed’ to ‘very generative’ chunks. Such a view of language can be demonstrated with reference to L2 learners’ generated utterances which, although they do not breach the linguistic rules of generative grammar, are deemed unnatural and sometimes odd by English native speakers due to the fact that “not all possible sentences of English – i.e., those that are well formed according to

the patterns of the language— are actual or even probable utterances of the language” (Lewis, 1997, p. 258).

As such, prefabricated sequences are believed to constitute about half of written English discourse and have different functional patterns as well as constructions that relate to different academic fields and registers (Biber et al., 2004; Boers & Lindstromberg, 2009; Ellis & Simpson-Vlach, 2009; Ellis et al., 2008; Erman & Warren, 2000; Hyland, 2008; Nattinger & DeCarrico, 1992; Schmitt & Carter, 2004; Simpson-Vlach & Ellis, 2010).

The indisputable existence of formulaic sequences in L1 discourse means that these formulas are to be considered a cornerstone of any full account of second language acquisition (Ellis, 1996; Ellis et al., 2008; Jones & Haywood, 2004; Oakey, 2002; Schmitt & Underwood, 2004; Wray & Fitzpatrick, 2008, 2010). In this regard, experts in the field of second language acquisition have argued that formulaic sequences constitute a major part of L2 learners’ linguistic repertoire, and, hence, they are believed to play a vital role in L2 learners’ receptive and productive skills (Boers & Lindstromberg, 2009; Fitzpatrick, 2005; Osborne, 2008; Schmitt, Dornyei, Adolphs, & Durow, 2004; Stengers, Boers, Housen, & Eyckmans, 2011; Wood, 2006, 2009a, 2009b, 2010a; Wray & Perkins, 2000).

At this point, it has been demonstrated that mastering formulaic sequences can facilitate the role of both the L2 speaker and the L2 listener in the sense that they reduce the speaker’s processing time required for speech production and maximize comprehension on the part of the listener (Wray, 2000). Moreover, formulaic sequences are believed to help L2 learners approach native-like proficiency since their mastery will make speech sound natural, idiomatic, and native-like (Boers et al., 2006; Liu & Huo, 2011; Qi & Ding, 2011; Wray, 2000). According to Wray (2000), gaining a full command of a new language “requires the learner to become sensitive to

the native speakers' preferences for certain sequences of words over others that might appear just as possible" (p. 463).

Based on the above-mentioned arguments, it can be noted that formulaic sequences constitute a major component of both L1 and L2 language users' linguistic repertoires, and, hence, they play a vital role in their language production, whether spoken or written.

The Nature, Types, and Functions of Formulaic Sequences

The holistic nature of formulaic sequences. With the growing interest in the formulaicity of language production, the body of research concerning formulaic sequences has attempted to determine the characteristics that identify this language phenomenon. Several researchers in the field of second language acquisition have defined formulaic sequences as prefabricated chunks that are stored in and retrieved from the memory as wholes (Arnaud, Ferragne, Lewis, & Maniez, 2008; Chen, 2009; Conklin & Schmitt, 2008; Pawley & Syder, 1983; Sinclair, 1991; Tremblay & Baayen, 2010; Wood, 2001, 2006). Among those researchers is Weinert (1995) who points out that formulaic sequences

refer to multi-word (*How do you do?*) or multi-form strings (*rain-ed, can-'t*) which are produced or recalled as a whole chunk, much like an individual lexical item, rather than being generated from individual lexical items/forms with linguistic rules. (p. 182)

The holistic nature of formulaic sequences has been likewise posited in Wray's (2002) definition of a formula as

a sequence, continuous or discontinuous, of words or other meaning elements, which is, or appears to be, prefabricated: that is, stored and retrieved whole from memory at the time of use, rather than being subject to generation or analysis by the language grammar. (p. 9)

As well, Wood (2010a) defines formulaic sequences as “multiword or polymorphemic units of language, stored in memory as if they are single lexical units, and recalled and produced as wholes” (p. 42).

The assumption that formulaic sequences are stored and processed holistically in long-term memory has been validated with reference to mental processing. It has been suggested that the holistic nature of formulaic sequences is clearly shown in the processing advantage they offer to both native speakers and proficient L2 users: they are processed more quickly and potentially differently than entirely creative expressions (Conklin & Schmitt, 2008, 2012; Gibbs & Gonzales, 1985; Tabossi, Fanari, & Wolf, 2009). Such claims have been tackled from different angles.

First and foremost, the holistic nature of formulaic sequences manifests itself in native speakers’ ability to produce utterances which exceed the largest unit of language that can be processed in the short-term memory, 8-10 words, without hesitation (Conklin & Schmitt, 2012). This ability can be attributed to the holistic nature of formulaic sequences, which are stored in and retrieved from the long-term memory at the time of language production to compensate for the limited capacity of the short-term memory (Conklin & Schmitt, 2012). Accordingly, the processing advantage of formulaic sequences, which is extended to different types of formulaic sequences as idioms, lexical bundles, collocations, and so on, can be seen as evidence in support of the assumption that formulaic sequences are stored and retrieved as if they were single entities (Conklin & Schmitt, 2012).

This processing advantage has also been empirically demonstrated in research studies on idiom processing. Gibbs and Gonzales (1985) conducted an empirical study to highlight any possible differences in native speakers’ processing of a number of idioms and their novel control

phrases. The findings of their studies, in which thirty-four undergraduate native speakers of English were asked to identify the links between idioms and their literal expressions, adduced further evidence to the holistic nature of formulaic sequences. In this study, the participants processed idiomatic expressions more quickly than novel control phrases, and, hence, Gibbs and Gonzales (1985) conclude that in the case of language comprehension and production “retrieving idioms from the lexicon operates in a similar manner to normal lexical access processes” (p. 257).

Similarly, Tabossi et al. (2009) investigate the holistic nature of formulaic sequences in an empirical study that includes a wider range of formulaic sequences, namely, decomposable idioms, non-decomposable idioms, and clichés. Tabossi et al.’s (2009) empirical study, in which the participants were required to identify the meaning of both formulaic sequences and their matched literal expressions, demonstrates that formulaic sequences, regardless of their type, are processed faster due to their holistic nature.

The fact that formulaic sequences are processed holistically has also been validated in a participant-paced line-by-line reading procedure. Conklin and Schmitt (2008) conducted a study in which nineteen native speakers of English and twelve L2 learners of English were assigned a line-by-line reading task, i.e., they read one line at a time and pushed a button to bring up a new line. The results of this study demonstrate that in the case of both native speakers and proficient L2 users, the lines replete with formulaic sequences were processed faster than the ones that were composed of novel expressions. Accordingly, Conklin and Schmitt (2008) posit that “formulaic sequences offer processing efficiency because single memorized units, even if made up of a sequence of words, are processed more quickly and easily than words which are generated creatively” (p. 75).

Accordingly, the assumption that formulaic sequences are acquired and recalled as single lexical items rather than being constructed by virtue of individual items and grammar rules implies that formulaic sequences, whether referred to in literature as formulaic sequences, multiword sequences, lexical bundles, prefabricated chunks, formulas, ready-made language, fixed expressions, collocations, prefabs, set phrases, prefabricated routines, etc., represent the same language phenomenon that is unitary in nature and, hence, their acquisition resembles that of vocabulary (Biber, 2006; Boers et al., 2006; Coxhead & Byrd, 2007; Jones & Haywood, 2004; Schmitt & Carter, 2004; Simpson-Vlach & Ellis, 2010; Weinert, 1995).

More importantly, formulaic sequences, being acquired and internalized as single entities, are believed to constitute a major component of L2 learners' linguistic repertoire and, thus, positively affect their language production, in both speaking and writing (Biber et al., 2004; Boers et al., 2006; Ellis, 1996; Liu & Huo, 2011; Nattinger & DeCarrico, 1992; Wood, 2010b; Wray, 2000, 2012).

Types of formulaic sequences. Conscious of their importance in language production, different scholars have attempted to classify formulaic sequences on the basis of their discourse or pragmatic functions, their frequency, and their syntactic structures in both spoken and written discourse. It is worth mentioning that despite the slight differences in the terms used to list formulaic sequences in discourse, the following types can be pinpointed in literature.

In an attempt to develop a taxonomy that represents the major discourse functions of prefabricated chunks, Biber et al. (2004) classify lexical bundles, the most frequent recurring multi-word sequences which have both structural and functional characteristics in a register, into three main categories and each of which serves different functions in discourse. To them,

although lexical bundles may have various functions in different registers, they can be classified into three major categories based on their most common use.

The first category includes ‘stance bundles,’ which are particularly used to express attitudes and help interpret subsequent propositions, and can be divided into epistemic and attitudinal/modality stance bundles (Biber et al., 2004). As for epistemic stance bundles, they are for the most part personal and are mainly utilized to express uncertainty rather than certainty (e.g. *I don't know if...*); however, they can also be impersonal stance bundles that convey, for the most part, a degree of certainty as in ‘*are more likely to be...*’ (Biber et al., 2004). Attitudinal stance bundles, in turn, are usually used to evince the speaker’s feeling towards an upcoming proposition, and, therefore, they may express a desire, ‘*I don't want to...*,’ an obligation, ‘*you have to do...*,’ an intention, ‘*what we're going to...*,’ or a certain ability ‘*to come up with*’ (Biber et al., 2004).

The second category of lexical bundles encompasses discourse organizing bundles that are employed by language users to introduce or elaborate on a topic and are, thus, subcategorized into topic introduction bundles and topic elaboration bundles (Biber et al., 2004). Topic introduction bundles are, as Biber et al. (2004) suggest, essentially recurrent in classroom contexts as they help teachers introduce new topics and reveal their intentions as in ‘*I want to talk about ...*.’ Topic elaboration bundles, in turn, are discourse markers which the speaker tends to deploy when he/she believes that further clarification on a particular topic is needed (e.g. *you know I mean...*); these bundles are also highly frequent in textbooks as they help compare and contrast information as in ‘*as well as the...*’ and ‘*on the other hand*’ (Biber et al., 2004).

Referential bundles constitute the third category of lexical bundles and are often used either to signal an entity as of paramount importance or to refer to physical and abstract entities (Biber

et al., 2004). These bundles are divided into four subcategories, namely identification/focus bundles, imprecision bundles, bundles specifying attributes, time/place/text-deixis bundles (Biber et al., 2004). Identification/focus bundles, as Biber et al. (2004) note, have various functions among which is focusing on the following noun phrase (e.g. *those of you who...*), organizing and summarizing lengthy discourse (e.g. *that's one of the...*), and introducing a topic and initiating a discussion (e.g. *one of the things...*). The second subcategory of referential bundles, i.e., imprecision bundles, is employed to highlight an imprecise reference or to show that a specified reference can be expanded by other references of the same type as in '*and things like that...*' (Biber et al., 2004). Bundles specifying attributes, as their name may indicate, are usually manipulated to specify the identifying attributes including the quantity, the size, or the abstract characteristics of the subsequent head noun, e.g. '*a little bit of....*,' '*the size of the....*,' and '*the nature of the....*' respectively (Biber et al., 2004). Time/place/text-deixis bundles, which make up the last subcategory of referential bundles, are employed to refer to time, place or location; moreover, they are context-specific as their meaning is determined by the context in which they are used as in '*at the end of the.../ as shown in figure 4.4...*' (Biber et al., 2004).

In addition to Biber et al.'s (2004) discourse based classification of lexical bundles, Hyland (2008) categorizes four-word bundles according to their functions in academic discourse, mainly research articles, Master's theses, and doctoral dissertations. To Hyland (2008), lexical bundles that are composed of four words can be divided into three broad categories based on their functions in academic discourse. To start with, research-oriented bundles, which help writers frame their activities and experiences, constitute the first category of lexical bundles; they are usually used to refer to locations that include time or place (*at the same time.../ in the present study...*), introduce a procedure (*the role of the...*), list qualifications (*the magnitude of the...*),

provide description (*the structure of the...*), or introduce a topic (*the currency board system*) in academic discourse (Hyland, 2008).

The second category is referred to as text-oriented bundles which play a major role in producing coherent and cohesive texts (Hyland, 2008). Among those bundles are transition signals, e.g. '*in addition to the ...*,' resultative signals as in '*these results suggest that...*,' structuring signals such as '*in the next section...*,' and framing signals, e.g. '*with respect to the...*' (Hyland, 2008). Participant-oriented bundles, which are employed to express the writer's attitude, e.g. '*are likely to be...*,' and directly address the readers as in '*it should be noted...*,' are the last category in Hyland's (2008) list.

Simpson-Vlach and Ellis (2010), in turn, present The Academic Formulas List (AFL), a list of the formulaic sequences that are recurrent in both academic spoken and written language as well as the ones that are typical of academic writing alone or academic speaking alone. The AFL, which is comparable with Coxhead's (2000) Academic Word List (AWL), categorizes formulaic sequences based on their pragmatic functions in different registers. As such, Simpson-Vlach and Ellis's (2010) list classifies formulaic sequences into three major groups, namely "referential expressions," "stance expressions," and "discourse organizing expressions," under which other subcategories are listed.

First, referential expressions constitute the largest portion of the three groups and are divided into five subcategories (Simpson-Vlach & Ellis, 2010). The first subcategory is referred to as "specification of attributes" and can, in turn, be divided into three types that serve various functions in academic discourse, i.e., intangible framing attributes, tangible framing attributes, and quantity specifications (Simpson-Vlach & Ellis, 2010). As for intangible framing attributes, Simpson-Vlach and Ellis (2010) explain that the vast majority of these formulaic sequences are

composed of ‘*a/the N of*’ (e.g. *the notion of...*) and can be sometimes preceded by a preposition (e.g. *on the basis of...*). These attributes, as Simpson-Vlach and Ellis (2010) suggest, may frame concrete entities (e.g. *based on the total volume...*) or abstract concepts (e.g. *even with the notion of eminent...*). More importantly, they may construct an attribute of a subsequent phrase, formulate a whole clause, or connect a verb with a following clause (Simpson-Vlach & Ellis, 2010). Tangible framing attributes, in turn, specify the physical or measurable attributes of the subsequent noun as in ‘*the level of shade...*’ (Simpson-Vlach & Ellis, 2010). The last type of specification of attributes, quantity specification, is very similar to tangible framing attributes and includes expressions that specify the quantity of the following noun phrase as in ‘*there are three...*’; they can also be anaphoric and refer to the preceding noun ‘*the combination of these two...*’ (Simpson-Vlach & Ellis, 2010).

In addition to specification of attributes, identification and focus is the second subcategory and involves explicatory phrases such as ‘*as an example*’ and clause stems like ‘*this would be...*’ (Simpson-Vlach & Ellis, 2010). To Simpson-Vlach and Ellis (2010), identification and focus attributes are of special importance in academic discourse since listing examples and identifications are major functions in academic contexts. The third subcategory is contrast and comparison and includes all comparative and contrastive clusters such as ‘*as opposed to...*’ (Simpson-Vlach & Ellis, 2010). Deictics and locatives are, as Simpson-Vlach and Ellis (2010) explain, the fourth subcategory and involve expressions that include proper nouns ‘*the United Kingdom*’ or refer to “physical locations in the environment (e.g. *the real world*) or to temporal or spatial reference points in the discourse (e.g. *a and b, at this point*)” (p. 505). The fifth and last subcategory, vagueness markers, is a very limited one in discourse in that it includes only four phrases (*and so on, and so forth, and so on and so, and blah blah blah*), three of which are

typical of spoken discourse (Simpson-Vlach & Ellis, 2010). Notwithstanding, the importance of these phrases in academic discourse can be attributed to their high frequency (Simpson-Vlach & Ellis, 2010).

Stance expressions represent the second category of formulaic sequences and are divided into six subcategories, namely, hedges, epistemic stance, obligation and directive, ability and possibility, evaluation, and intention/volition (Simpson-Vlach & Ellis, 2010). To start with, hedges comprise all the formulaic sequences that serve hedging functions in discourse such as ‘*there may be...*’ (Simpson-Vlach & Ellis, 2010). Epistemic stance, in turn, includes the formulaic sequences that help present beliefs, claims, thoughts, etc., for example, ‘*... assume that it...*’ (Simpson-Vlach & Ellis, 2010). Moreover, while obligation and directive formulas, the third subcategory, serve the function of giving directions as in ‘*it should be noted...*’, ability and possibility formulas constitute the fourth subcategory and are mainly used in speaking as they help form actions or propositions as in ‘*you can see...*’ (Simpson-Vlach & Ellis, 2010). Furthermore, evaluation formulas constitute the fifth subcategory and include such sequences as ‘*it is obvious that...*’ or ‘*important role in...*’ (Simpson-Vlach & Ellis, 2010). The last subcategory, which frequently occurs in spoken discourse, is intention/volition and includes sequences that help convey the speaker’s intentions as in ‘*so let me just...*’ (Simpson-Vlach & Ellis, 2010).

Third, the last category of formulaic sequences is referred to as discourse organizing expressions and is subdivided into four subcategories (Simpson-Vlach & Ellis, 2010). The first subcategory of this type, metadiscourse and textual reference, is genre specific and includes formulaic sequences that help introduce topics, e.g. *in the next section...* (Simpson-Vlach & Ellis, 2010). Topic introduction and focus formulas constitute the second subcategory and help frame

an entire phrase ahead as in ‘*take a look at...*’ (Simpson-Vlach & Ellis, 2010). Topic elaboration, the fourth subcategory of this group, is divided into non-causal topic elaboration, where a phrase like ‘*it turns out that...*’ helps elaborate on a topic without any explicit causal relations whatsoever, and cause and effect formulaic sequences, whose main function is to introduce results, e.g. *as a result...* (Simpson-Vlach & Ellis, 2010). According to Simpson-Vlach and Ellis (2010), the latter is particularly important in EAP contexts. Finally, discourse markers can connect either sentence constituents or clauses together as in ‘*...as well as...*’ or ‘*in other words...*’; they may express agreement, disagreement, gratitude, or surprise such as ‘*no no no*’ (Simpson-Vlach & Ellis, 2010).

In addition to the aforementioned classifications, other scholars have attempted to broadly classify formulaic sequences according to their syntactic structure. It is worth noting that the following classification is mainly based on Coxhead and Byrd’s (2007) list which clearly overlaps with other categorizations proposed by prominent scholars in the field of discourse analysis. The first type includes “multi-word combinations that are structural or semantic units” (Coxhead & Byrd, 2007, P. 135). This type comprises phrasal verbs that consist of a verb and its particle as one set, e.g. ‘*look up/ look after*’ or prepositional verbs such as ‘*agree to, deal with, etc.*’ (Coxhead & Byrd, 2007; Howarth, 1998; Nattinger & DeCarrico, 1992). The same class has been referred to as grammatical collocations by Lewis (2000a). Idioms are the second type of formulaic sequences and include immutable sequences whose meaning is opaque as it cannot be interpreted from the individual words that constitute it (Arnaud & Savignon, 1997; Coxhead & Byrd, 2007; Handl, 2008; Omazic, 2008; Simpson-Vlach & Ellis, 2010). Examples of idioms are ‘*shoot the breeze,*’ ‘*blow the gaff,*’ ‘*back to square one,*’ ‘*pig in a poke,*’ ‘*beat around the*

bush,’ etc. (Boers et al., 2006; Cortes, 2004; Coxhead & Byrd, 2007; Howarth, 1998; Lewis, 2000a; Wray, 1999, 2002).

The third type encompasses words that are likely to appear and/or collocate together in discourse more frequently than what is expected by chance, i.e., collocations such as ‘*tell a story*,’ ‘*commit suicide*,’ ‘*under attack*,’ etc. (Boers et al., 2006; Coxhead & Byrd, 2007; Howarth, 1998; Kennedy, 2008; Lewis, 2000a; Martin, 2008; Nattinger & DeCarrico, 1992; Philip, 2008; Woolard, 2000). They can be either lexical collocations which are constructed of two lexical words (e.g. *suggest an alternative*) or grammatical collocations which are composed of a lexical word and grammatical word as in ‘*aware of*’ (Lewis, 2000a). Collocations, as Martin (2008) suggests, are more likely to be viewed as lexical units in which “two elements show a degree of binding/fixation or restriction to each other, thus forming a unit that fits somewhere in between idioms and free combinations” (p. 56). However, collocations may also have variable lexical slots that can be filled with semantically-related lexicon and, hence, demonstrate the strong semantic and syntactic ties that relate these sequences (Philip, 2008). Such a relationship is manifested in collocations consisting of a fixed part and a variable one as in ‘*an accident of birth*’ where the noun ‘*birth*’ can be replaced with ‘*war, history, fate, etc.*’ or in ‘semi-packaged phrases’ such as ‘*the idea*’ where the slot can be filled with semantically related lexis among which are ‘*faintest, slightest, foggiest, etc.*’ (Philip, 2008).

Verb features and valency patterns constitute the fourth type in this broad list and, as Coxhead and Byrd’s (2007) explain, have an important role in different registers since they embody the strong relationship between grammar and lexicon, i.e. the characteristic features of verbs restrict the lexicogrammatical features of the following constituents in an utterance as in the case of verbs that require that-clause, infinitives, etc. as complements. Lexical bundles, in

turn, are the fifth type and refer to such fixed sequences as ‘*in order to,*’ ‘*to sum up,*’ ‘*on the other hand,*’ etc. that occur frequently in different registers of academic discourse (Biber et al., 2004; Cortes, 2004; Coxhead & Byrd, 2007; Hyland, 2008; Wray, 1999). The sixth type includes semi-fixed sequences that amalgamate semantic requirements with lexicon and syntax— language users’ choice of lexicon is determined by the semantic boundaries of certain lexical sets: modals necessitate a certain sets of words as in ‘*may well be...*’ (Coxhead & Byrd, 2007).

Sentence heads and frames are the last type of formulaic sequences and include short sequences like ‘*and finally...*,’ ‘*the ---- of the----*,’ etc. or longer ones as ‘*There is a growing body of evidence that ...*’ or ‘*it has been asserted ...*’ (Boers et al., 2006; Coxhead & Byrd, 2007; Lewis, 1997). It should be noted that sentence frames are also referred to as “sentence builders” because they help frame long sentences and structure written texts appropriately (Lewis, 1997; Oakey, 2002). Accordingly, these sentence frames and heads are believed to be important for augmenting L2 learners’ academic writing skills (Coxhead & Byrd, 2007; Lewis, 1997; Oakey, 2002).

Functions of formulaic sequences. As mentioned earlier, the importance of formulaic sequences can be ascribed to “their frequencies of use and obvious discourse functions” (Biber et al., 2004, p. 400). As such, due to their holistic nature and frequent occurrence in discourse, formulaic sequences are believed to be a keystone of L2 learners’ fluent and proficient language production (Boers & Lindstromberg, 2009; Ellis, 2007; Ellis et al., 2008; Granger, 1998; Kormos & Dénes, 2004; Lewis, 2000b; Weinert, 1995; Wood, 2006, 2009a, 2009b, Wray, 2000). That is, the mastery of formulaic sequences, which function as links or frames to express language users’ communicative messages, reduces the amount of time required for language processing and,

hence, facilitates “efficient and effective communication, and particularly fluent speech” (Wood, 2009a, p. 40).

Building a rich repertoire of formulaic sequences and integrating them into speech production may also identify L2 learners as proficient language users whose speech is replete with complex language that is sensitive to native speakers’ choice of utterances (Boers et al., 2006; Boers & Lindstromberg, 2009; Qi & Ding, 2011; Wray, 2000, 2002). According to Boers and Lindstromberg (2009),

... the broader that repertoire is (in terms of chunks as well as single words), the greater the likelihood that the student will be able to retrieve from it the lexis that appropriately and precisely conveys the message waiting to be conveyed. In addition, the use of chunks can help students to be perceived as idiomatic language users ..., disposing of a relatively impressive lexical richness and syntactic complexity. (p. 37)

In addition to speech fluency, L2 pragmatic competence has also been attributed, at least partially, to the mastery of formulaic sequences. Littlewood (2004), for instance, speculates that the clearest evidence of the existence of formulaic sequences in L2 learners’ linguistic repertoire is provided by set phrases which L2 learners often produce as a means of coping with common or important situations that are beyond their linguistic or pragmatic competence. Accordingly, when, in several cases, L2 learners’ other output shows no evidence that they have mastered the structures necessary for producing some complex utterances, their advanced linguistic and pragmatic competence should be attributed to the use of formulaic sequences which they have successfully acquired and internalized into their long-term memory (Chowdhury & Tarannum, 2010; Granger, 1998; Littlewood, 2004; Weinert, 1995; Wood, 2010a).

Formulaic sequences can also contribute to L2 learners' linguistic accuracy in the sense that these multiword strings, being acquired and internalized in the long-term memory as wholes, are likely to mitigate L2 learners' possible grammatical errors when communicating in the target language (Boers et al., 2006; Wray & Fitzpatrick, 2008). As Boers et al. (2006) hypothesize, "these pre-fabricated chunks constitute 'zones of safety' and appropriate use of them may thus confine the risk of 'erring' to the spaces in between the formulaic sequences in one's discourse" (p. 247). It is worth noting that the role of formulaic sequences in increasing L2 learners' linguistic accuracy has not been extensively empirically investigated in writing or in speech.

Wray and Fitzpatrick (2008), in turn, after an in-depth study in which six participants were required to memorize and practice formulaic sequences, conclude that integrating formulaic sequences in the participants' speech production provided them with the "opportunity to sound like native speakers, promoted their fluency, reduced the panic of on-line production in stressful encounters, ... and provided materials that could be used in other contexts too" (p. 143).

Formulaic sequences can decrease the loads imposed by language processing since they function as frames that might help language users express their communicative messages fluently, allowing the speaker some time for planning the next utterance, generating specific lexicon, and processing novel pieces syntactically (Wood, 2009b). Formulaic sequences may also foster L2 learners' expressive skills as a result of successfully acquiring various chunks that can be used creatively in communication (Schmitt, 2000). That is, after acquiring formulaic sequences as chunks, L2 learners may realize that these sequences can be used creatively as in "How are you today?" where the adverbial phrase of time can be replaced with "this evening," "this fine morning," etc. (Schmitt, 2000).

Formulaic sequences have also been viewed as building blocks which are essential for proficient writing skills since they may “become a part of the structural material used by advanced writers, making the students’ task easier because they work with ready-made sets of words rather than having to create each sentence word by word” (Coxhead & Byrd, 2007, p. 135). Moreover, the successful implementation of these prefabricated chunks helps define L2 learners as proficient writers due to their utilization of prefabricated chunks that are typical of their academic registers (Coxhead & Byrd, 2007; Granger, 1998; Hyland, 2007; Nattinger & DeCarrico, 1992; Simpson-Vlach & Ellis, 2010). More importantly, these frequently occurring sequences are associated with versatile communicative functions providing a frame for expressing different stances in discourse, and, hence, they are “crucially important for the construction of discourse in all university registers” (Biber, 2006, p. 174). In a nutshell, formulaic sequences, as Wood (2005) puts it,

... are not just common clichés or very frequent collocations, but they function in all registers and in highly specialized content domains. They may have quite low general frequency of occurrence but be frequent within certain cultural groups or specific genres. (p. 67)

Formulaic Sequences and Academic Writing Skills

Writing is one of the four language skills which is viewed, by many scholars, as a unique tool for expressing ideas in written forms. It is, as Kress (1994), Llach (2011), Nattinger and DeCarrico (1992), and Paré (2009) suggest, a powerful means manipulated by language users to communicate knowledge and share factual information in readable texts. Similarly, in her account for the importance of writing as a productive skill, Raimes (2002) explains that

Writing, unlike speaking, provides us with a way not only to generate ideas before presenting them, but also to scrutinize the ideas and language we produce; this re-vision, this

seeing again, lets us receive feedback from ourselves and from others and, learning as we go, make changes and corrections. (p. 309)

Writing is also considered a powerful device manipulated by language users to accomplish certain tasks because “we don’t just write, we write something to achieve some purpose: it is a way of getting something done” (Hyland, 2003, p. 18). In other words, writing is viewed as a ‘mediating activity’ in which a writing process takes place to accomplish various tasks and activities including communication and sharing information (Gentil, 2006). It is worth mentioning that this productive language skill has become particularly important with advances in technology as it helps language users convey their ideas in a world that is being more and more dominated by texts and numerical data (Hyland, 2003).

In addition to its expressive function, writing is also viewed as a fundamental assessment medium, especially in academic contexts where most, if not all, evaluations evolve around students’ ability to generate, organize, and translate ideas into academic texts including essays, summaries, etc. (Connor, 2003; Leki, 2006; Llach, 2011; Nation, 1990; Paltridge, 2004; Zhu, 2006). In this regard, university students’ written samples help test their in-depth knowledge and thorough understanding of their academic disciplines:

Written responses to exam questions and writing assignments constituted a more effective method for evaluating students real understanding of course materials and for assessing students’ ability to analyze and synthesize information and to apply their knowledge to solving real world problems. (Zhu, 2006, p. 136)

Writing is believed to promote language acquisition in general and lexical competence in particular (Llach, 2011). When L2 learners are assigned a writing task, they get the chance to practice the various aspects of the target language including its syntax and lexical entities; thus, it

is viewed as a powerful tool that enlarges L2 learners' linguistic repertoire (Llach, 2011). Writing skills have also been viewed as "an indicator of language proficiency especially in academic contexts" (Llach, 2011, p. 51). Writing comprises a major part of most university placement tests and language proficiency tests, and the evaluation of L2 learners' writing is mainly based on their ability to use accurate grammar and rich lexis to express their ideas academically in written texts (Llach, 2011).

However, in contrast to other language skills, writing, especially in academic contexts, is considered a complex and painstaking process, mainly for L2 learners who may feel daunted at the prospect of being underestimated due to their underdeveloped expressive writing skills (Bacha, 2002; Cook & Bassetti, 2005; Coxhead & Byrd, 2007; Ferris, 2002; Hyland, 2003, 2006; Leki, 2006; Raimes, 2002; Reppen, 2002; Richards & Renandya, 2002; Silva, 1993). Such a complexity has been predominantly ascribed to the nature of academic prose which, unlike creative writing, e.g. poetry, in which the writer's success resides in his/her ability to juxtapose ideas in entirely novel, unfamiliar, and fresh utterances, is characterized by the high frequency of common lexical sequences (Biber et al., 2004; Conklin & Schmitt, 2008; Coxhead, 1998; Jones & Haywood, 2004; Wray, 2002).

As Lewis (2000c) suggests, "in academic writing, where the focus is almost exclusively on accurate communication of information, among colleagues with a shared background in a particular topic, standard words, phrases, collocations and other chunks are an essential prerequisite for effective communication" (p. 189). This view of academic discourse implies that academic writing skills surpass the mastery of lexicon and syntax to encompass the successful implementation of multiword sequences that are viewed as the building blocks of academic

discourse (Biber, 2006; Corson, 1995; Coxhead & Byrd, 2007; Ellis et al., 2008; Hinkel, 2002; Li & Schmitt, 2009; Martinez & Schmitt, 2012).

In this respect, formulaic sequences can be viewed as a cornerstone of successful academic writing skills because they represent the frequent expressions by which L2 writers have to abide in that register (Ellis & Simpson-Vlach, 2009; Lewis, 1997; Martinez & Schmitt, 2012).

For example, it has been suggested that, in several occasions, although L2 learners' generated sentences are grammatically correct, they are judged to be unnatural or odd because they do not adhere to the characteristics that are conventional of a well-formed academic piece of writing. This common problem is highlighted in Lewis' (1997) account of the importance of sentence frames and heads in academic writing:

The frustration of reading a student's essay and thinking 'I know what you mean, but that's not the way to say (= write) it', is most frequently caused by the student's failure to use this type of lexical item. Some are comparatively short and easy (sequences such as *secondly*, ... *and finally*); some are sentence heads serving similar pragmatic purposes (*We come now to a number of important reservations*). (p. 259)

Not only does such deviation affect the quality of a piece of writing, but it may also negatively influence the overall evaluation of L2 students in academic contexts (Lewis, 1997).

Jones and Haywood (2004) point out that EAP programs should be designed in a manner that provides L2 students with the opportunity to acquire, master, and internalize a considerable number of formulaic sequences, since L2 students' acceptable academic performance is contingent on their successful use of lexicon which represents academic discourse. To demonstrate their viewpoint, they refer to Cowie (1992) who hypothesizes that "it is impossible

to perform at a level acceptable to native users, in writing or in speech, without controlling an appropriate range of multiword units” (as cited in Jones & Haywood, 2004, p. 269).

Coxhead and Byrd (2007), in their account of the various components of academic writing, illustrate that “writing for academic purposes involves specialized knowledge of academic genres” (p. 133), which is characterized by particular recurrent vocabulary that is specific to academic discourse. Such vocabulary includes single words as well as formulaic sequences which are highly important since they “come in relatively fixed sequences and ... are likely to be stored in memory as sets rather than created word-by-word” (Coxhead & Byrd, 2007 p. 134). Their holistic nature can, thus, contribute to accurate academic writing. As such, academic writing is a complex process that merges the mastery of grammar and lexicon with the concise knowledge of the academic genres as well as the vocabulary typical of the field in question (Coxhead & Byrd, 2007).

Hyland (2007), in turn, maintains that writing is a complex process in that it “varies with context and cannot be distilled down to a set of abstract cognitive or technical abilities” (p. 150). In this respect, L2 writers’ effective and efficient writing resides not only in their mastery of grammar and lexicon but also in their ability to successfully comply with the intended readers’ expectations by implementing the formulaic sequences typical of the discourse in question. Moreover, Hyland (2008) enunciates that among the difficulties that might face L2 writers in academic contexts is their poor knowledge of this language phenomenon. While the implementation of formulaic sequences in academic writing, as Hyland (2008) suggests, helps shape meaningful and coherent academic texts, their absence can easily signal L2 learners as nonnative writers.

Ellis et al. (2008) note that research on EAP has revealed that each academic discipline has a high frequency of lexical sets and learning to write effectively in a particular discipline is germane to mastering these recurrent sets. To elaborate on this notion, they explain that L2 writers may have advanced knowledge of grammar and lexicon; however, the sentences they generate, although grammatically correct, may sound unnatural and foreign. The difficulty in attaining natively like academic style can be attributed to L2 writers' unawareness of the collocations and formulaic sequences that are representative of their fields. In this regard, Ellis et al. (2008) hypothesize that in addition to their mastery of grammar rules and lexicon, L2 students "have to know the common collocations and lexical bundles, too, not only to increase their reading speed and comprehension ..., but also to be able to write in a natively like fashion" (p. 277).

As well, in their account of the importance of formulaic sequences in academic writing, Li and Schmitt (2009) state that prefabricated chunks are extremely widespread in academic discourse and have various pragmatic functions. By considering their importance in academic discourse, one can notice that

formulaic sequences like *as a result* and *it should be noted that* are central to the creation of academic texts ... As important building blocks of the characteristic features of academic texts, the absence of such formulaic sequences may indicate the lack of mastery of a novice writer in a specific disciplinary community ..., given that to be a successful academic writer, an L2 learner is required to be competent at using these conventional sequences which characterize the learner's discipline. (Li & Schmitt, 2009, p. 86)

Accordingly, learning to write effectively necessitates being able to successfully integrate formulaic sequences into texts since failure to do so may result in producing nonnative pieces of

writing, which, although grammatical, might be judged as foreign or even odd (Li & Schmitt, 2009).

Based on the aforementioned arguments, it can be posited that formulaic sequences are considered a centerpiece of academic writing. However, not only do L2 learners need to utilize formulaic sequences that are peculiar to academic discourse, but they are also required to master their complex and diverse structures and functions in order to construct a coherent and well-developed piece of writing (Cortes, 2004; Granger, 1998; Hyland, 2007, 2008; Li & Schmitt, 2009; Nation, 1990). That is, L2 learners should be capable of accurately using prefabricated sequences since errors in their use can have negative effects akin to those of avoiding them in discourse, i.e., defining language users as outsiders (Handl, 2008).

In addition, L2 learners need to use a wide range of formulaic sequences because overusing certain prefabricated sequences and disregarding others may also result in producing poor pieces of writing due to the high repetition of a limited number of multiword strings; as Granger (1998) puts it, “while the foreign-soundingness of learners’ productions has generally been related to the lack of prefabs, it can also be due to the excessive use of them” (p. 155). Similarly, Hyland (2006) explains that L2 learners’ written products that lacked “good models of target academic genres... seemed to fall back on a limited number of prefabricated ‘lexical bundles’ to avoid grammatical errors, leading them to a repetitive style of writing” (p. 60). Likewise, Paquot (2008) maintains that the ‘non-nativeness’ or ‘unconventionality’ of L2 learners’ writing can be mainly ascribed to their tendency “to overuse a limited number of frequent English collocations and prefabs but to underuse a whole set of native-like phraseological units, especially typical EAP units” (p. 102).

Formulaic Sequences in the Classroom

The lexical approach to language teaching. Since formulaic sequences are so pervasive an element of both spoken and written language, they are key components of competent language production, and, hence, they represent ideal units for second language instruction, especially in academic contexts (Arnaud & Savignon, 1997; Boers & Lindstromberg, 2009; Cortes, 2004; Coxhead, 2008; Hyland, 2007; Li & Schmitt, 2009; Martinez & Schmitt, 2012; Moon, 1997; Nation & Newton, 1997; Nattinger & DeCarrico, 1992; Sinclair, 1991; Wood, 2010a). Such a claim has been clearly propounded in what is known to be a lexical approach to language teaching, in which formulaic sequences are believed to be the building blocks of successful language learning since they serve as what Boers and Lindstromberg (2009) call ‘islands of reliability’ upon which L2 learners depend when communicating in the target language (Boers & Lindstromberg, 2009; Hill, 2000; Lewis, 2000b; Richards & Rodgers, 2001).

The lexical approach to language teaching has emerged as a counterpart to the long prevailing belief in the effectiveness of teaching syntactic structures based on the assumption that each time a sentence is constructed, it is done by applying the relatively laborious syntactic rules to generate combinations from individual words (Hill, 2000; Lewis, 1997; Zimmerman, 1997). While such a notion of language teaching has placed too much focus on grammar by viewing it as the backbone of language, corpus linguistics has revolutionized scholars’ view of the structural syllabus (Hill, 2000; Lewis, 2000b). Thus, the doctrine of structural-based approach to language teaching has faded over time and has been replaced, at least partially, by the assumption that a great deal of language is constructed through a chunking process (Hill, 2000; Lewis, 1997, 2000b).

The lexical approach to language teaching can also be seen as a counteraction to the prevailing belief in the incidental learning of vocabulary, i.e., learning vocabulary as the by-

product of being engaged in other activities such as reading and listening (Coady, 1997; Richards & Renandya, 2002). This view of vocabulary learning has long marginalized its role as a core component of classroom materials. In this respect, the emergence of the lexical approach to language teaching has placed more focus on vocabulary in L2 classrooms based on the assumption that gaining a rich repertoire of lexical entities can enhance L2 learners' fluent and accurate language production (Coxhead, 2000; Hill, 2000; Lewis, 1997, 2000a, 2000b; Richards & Renandya, 2002). As Ellis (1996) puts it, "lexical phrases are as basic to SLA as they are to the L1 ..., and so instruction relies as much on teaching useful stock phrases as it does on teaching vocabulary and grammar" (p. 97).

Incorporating formulaic sequences in classroom materials. The undisputed occurrence of formulaic sequences in discourse implies that these prefabricated chunks should be incorporated into L2 classroom practices (Ellis, 1996; Ellis et al., 2008; Hill, 2000; Martinez & Schmitt, 2012; Nattinger & DeCarrico, 1992; Schmitt, 2000; Wood, 2010a). This claim can be mainly attributed to their widespread use in both spoken and written discourse, and, thus, they should "have a prominent place in language teaching textbooks and materials, as well as tests of language achievement and proficiency" (Martinez & Schmitt, 2012, p. 301).

The frequent use of formulaic sequences entails that introducing L2 learners to multiword sequences is likely to upgrade their academic skills by providing them with chunks that help the production of effective pieces of writing which approach the specific standards of academic genres (Hill, 2000; Nation, 2001). Accordingly, academic lexis can be seen as "a language bar or barrier that students need to transcend in order to move successfully from everyday ways of expressing meaning to the specialized, 'high-status' academic language" (Coxhead & Byrd, 2007, p. 132).

In addition to complying with the norms of academic discourse, raising L2 learners' awareness of formulaic sequences might have other pedagogical implications. According to Schmitt (2000), successful acquisition of L2 lexical entities can promote L2 learners' accuracy since lexical errors have been increasingly viewed as more serious than grammatical ones, especially in the case of writing where such errors, unlike those of grammar, may disrupt or hinder comprehension altogether.

Formulaic sequences can also help mitigate L2 learners' syntactic errors, which can be mainly ascribed to the lack of collocational competence (Hill, 2000). Hill (2000) posits that, in several occasions, L2 learners' undeveloped collocational competence may increase the number of their grammatical errors in both writing and speech "because they do not know the collocations which express precisely what they want to say" (p. 49). Integrating multiword sequences into classroom materials may as well reduce the burden of learning individual words on the part of L2 learners (Nation, 2001).

Introducing formulaic sequences in the classroom is likely to encourage L2 learners to implement them in their academic writing. To Cortes (2004), "it is possible that students were never explicitly instructed to use these target bundles when writing for their academic disciplines, and that is the reason why the use of these expressions is so rare in the student corpora" (p. 421).

Focused instruction of formulaic sequences. A considerable number of researchers advocate an explicit approach to lexical teaching for various reasons. First and foremost, the successful acquisition of L2 lexicon, whether individual words or formulaic sequences that are processed as single entities, is neither instantaneous; nor can it occur by chance (Boers &

Lindstromberg, 2009; Cortes, 2004; Coxhead & Byrd, 2007; Howarth, 1998; Nation, 2002; O’Keeffe, 2012; Schmitt, 2000).

Schmitt (2000), for example, explains that developing both receptive and productive knowledge of lexical items— understanding the meaning of a lexical item and being able to use it productively in both speaking and writing— necessitates a robust support that provides L2 learners with the multiple exposures needed for successful acquisition. To him, “in the case of vocabulary, the more one engages with a word (deep processing), the more likely the word will be remembered for later use” (Schmitt, 2000, p. 121). Moreover, raising L2 learners’ awareness of lexical entities, can, in several ways, prime the noticing of these items in discourse and, hence, positively impact both implicit and explicit knowledge of the target language (Nation, 2002).

Other scholars take their arguments to the other extreme by suggesting that neither multiple exposure nor implicit learning can fully familiarize L2 learners with the various functions of prefabricated sequences, which are complex in nature (Coxhead, 2000; Coxhead & Byrd, 2007; Ellis et al., 2008; Hunston, 2002; Hyland, 2011; Kennedy, 2008; Lewis, 2000b; Li & Schmitt, 2009). Despite their frequency in various contexts, formulaic sequences seem to be complex and sometimes opaque, at least partially, to many L2 learners (Boers & Lindstromberg, 2009; Cortes, 2004; Jones & Haywood, 2004).

According to Boers and Lindstromberg (2009), unlike the case of L1 acquisition where native speakers build up a rich repertoire of formulas as a result of sheer exposure, L2 learners may fail to successfully incorporate formulaic sequences into their linguistic repertoire, even after being engaged in appropriate awareness-raising or noticing activities due to the complex nature of formulaic sequences. These formulas may be either entirely transparent and, hence, pass unnoticed by L2 learners, or be highly opaque, and thus evoke a kind of mental burden on

the part of L2 learners who will, in turn, try to disregard them in context (Boers & Lindstromberg, 2009). As such, Boers and Lindstromberg (2009) conclude that “it is part of a teacher’s role in the classroom to help students not just to notice particular chunks in the course materials but also to commit these chunks to memory” (p. 68).

Jones and Haywood (2004) explain that formulaic sequences which are major components of academic prose, should be explicitly taught since they might not be perceptible to L2 learners who may, therefore, misuse them in academic contexts. Ellis et al. (2008) propose that L2 learners usually encounter difficulties when trying to employ formulaic sequences in their language production, especially writing. Such a difficulty can be ascribed to both the complex nature of formulaic sequences and the negative transfer that might occur from L1 to L2. In this respect, Ellis et al. (2008) note that “our ESL learners clearly need support in learning the formulaic sequences which have high utility in the specialist discourse of EAP” (p. 391). More importantly, through focused instruction, L2 learners will probably get the chance to rehearse and internalize this language phenomenon and ultimately utilize it in their language production (Wood, 2009a).

Based on the discussions proposed by some prominent figures in the field of second language acquisition, it can be posited that focused instruction is likely to elevate the saliency of formulaic sequences, raise L2 learners’ awareness of the different discourse functions that these chunks execute in various contexts, promote L2 learners’ ability to accurately implement them in their language production, and augment L2 learners’ implicit knowledge of the target language (Boers & Lindstromberg, 2009; Cortes, 2004; Ellis et al., 2008; Granger, 1998; Hyland, 2007; Nation, 2002; O’Keeffe, 2012; Wood, 2009a).

Classroom activities. Conscious of their significant role in language production, several scholars have provided insights into integrating formulaic sequences into the language classroom. The suggested activities predominantly aim at promoting the acquisition of formulaic sequences in order to expand L2 learners' linguistic repertoire and upgrade their academic writing skills.

Enriching L2 learners' linguistic repertoire with formulaic sequences. Teachers can implement several activities to enhance the internalization of formulaic sequences into L2 learners' long-term memory. To start with, memorizing formulaic sequences at the early stages of language learning can augment their acquisition (Nattinger & DeCarrico, 1992). Teachers can expand L2 learners' linguistic repertoire by having them, through drills, memorize flexible formulaic sequences that can be used creatively later on. For instance, instead of presenting such a fixed phrase as *'I'd like to express my sympathy (about X),'* teachers can have their students practice flexible formulaic sequences with 'open slots' as in *'I'm (very) sorry (to hear (about) X),'* where all the phrases between brackets can be creatively replaced with other phrases as L2 learners become more competent in the target language (Nattinger & DeCarrico, 1992). The mastery of 'flexible lexical phrases' is believed to help learners produce, in writing or in speech, more complex sentences that are constructed through a chunking process rather than being composed from scratch (Nattinger & DeCarrico, 1992).

In addition to drills, memorization can be promoted through such techniques as writing formulaic sequences with their translations on flash cards, having L2 learners repeat them with appropriate intervals to enhance mental processing, instructing L2 learners to put formulaic sequences in meaningful sentences, etc. (Nation, 2001, 2002). Moreover, associating formulaic sequences with visual images and creating lists that classify them thematically may enhance deep

processing which can, in turn, increase the chances of acquiring and committing formulaic sequences to the long-term memory (Boers & Lindstromberg, 2009; Celce-Murcia & Olshtain, 2000; Hunt & Beglar, 2002).

Dividing up texts, seeing the patterns in chunks, and deliberately learning new chunks are other techniques proposed to help the acquisition of formulaic sequences through focused instruction (Nation, 2001). The first technique, dividing up texts, can be achieved through four main activities. The first two activities, dictation and repetition, exhort L2 learners to retain spoken formulaic sequences in their working memory before producing them chunk by chunk in a written form, as in the case of dictation, or orally as a delayed repetition activity (Nation, 2001). In addition to dictation and delayed repetition, L2 learners can practice dividing up texts through ‘read-and-look-up’ activities in which they are required to read a written text, ideally a dialogue, and practice it in pairs; such an activity is believed to urge L2 learners to depend on their memory so that they can remember as many chunks as possible when practicing a dialogue, for example, with their peers (Nation, 2001). Dividing up texts can also be achieved through delayed copying which involves L2 learners in a text copying process; L2 learners can be assigned a text and instructed to rewrite it through phrase for phrase rather than word for word copying (Nation, 2001).

As far as collocations are concerned, L2 learners’ collocational knowledge can be buttressed by a typical collocation activity in which L2 learners are provided with two lists and instructed to match up the items presented in both lists together to form collocations (Nation & Newton, 1997). Another collocation activity involves discussing the meaning of different collocations that share in common the same stem, i.e., the fixed part of the collocation, and then identify their common

underlying meaning; in so doing, L2 learners can perceive these lexical items as collocations rather than individual words (Nation & Newton, 1997).

The internalization of formulaic sequences into L2 learners' linguistic repertoire may also be supplemented by the strategies suggested by Wood (2010a). Among those strategies is engaging L2 learners in such communicative activities as the mingle jigsaw, a highly cooperative activity in which parts of a text are assigned to L2 learners who, in turn, orally deliver information to their peers to construct the whole text (Wood, 2010a). According to Wood (2010a), this 'information-sharing technique' provides L2 learners with the repetition required for acquiring and internalizing formulaic sequences into their linguistic repertoire. In addition to the mingle jigsaw, teachers can provide L2 learners with lists of formulas which can be manipulated to suit certain contexts and tasks and have them practice these lists in such activities as role plays (Wood, 2010a). L2 learners may also be directed to make notes on the ways in which formulaic sequences occur in various contexts. When learners are not sure of the meaning and function of a particular chunk, they should be encouraged to "combine a general phonetic image of the sequence with a sense of its possible function based on context clues" (Wood, 2010a, p. 196).

Teachers can instruct L2 learners to list the formulaic sequences they encounter outside the classroom, in their communities or in the media, for example, and categorize them based on their functions (Wood, 2010a). L2 learners might as well "use phonetic, context, or component analysis to determine the meaning, function, or structure of a sequence encountered in the community or in the media" (Wood, 2010a, p. 197). L2 learners can also benefit from inspecting the cultural aspects and metaphors that underlie the use and structure of some formulaic sequences. In addition to being sensitive to both learner and context variables, the above-listed

strategies are likely to enhance the acquisition and automatization of formulaic sequences which foster L2 learners' expressive skills in various contexts (Wood, 2010a).

Augmenting L2 learners' academic writing skills. Some experts in the field of second language acquisition have suggested a variety of formulaic sequences activities to help improve L2 learners' academic writing skills. Nattinger and DeCarrico (1992), for example, highlight the usefulness of having L2 learners practice the creative use of some formulaic sequences, which function as sentence builders or frames peculiar to writing tasks. As such, they can use these frames as a backbone for a coherent and well-structured academic piece of writing. Among those frames are topic nominating phrases (e.g. *It has been asserted/ believed/ noted that...*), topic nomination (for example, *the (basic) emphasis/ proposal/ goal of this paper/ article is to ...*), and topic organization as in '*this paper will show/compare/contrast/describe/demonstrate that...*' (Nattinger & DeCarrico, 1992). L2 learners can also be required to outline some academic texts, a process that can make the forms and functions of formulaic sequences in discourse more salient (Nattinger & DeCarrico, 1992).

L2 learners may be required to form sentences by using sentence heads and endings provided in lists (Lewis, 1997). They may also be assigned a task in which they complete sentences or texts using either the chunks presented by their teachers or their own language (Lewis, 1997). Writing activities in which L2 learners use particular formulaic sequences repeatedly can promote the subsequent acquisition of chunks and upgrade L2 learners' writing skills (Conzett, 2000). Furthermore, teachers can prepare a list of formulaic sequences that relate to a specific topic and present it as a prewriting activity to familiarize L2 learners with the formulaic sequences that are typical of certain contexts (Conzett, 2000). Other activities include having L2

learners memorize common formulaic sequences that are typically used to connect ideas and/or sentences together (Conzett, 2000).

Studying the concordances of formulaic sequences has also been suggested as a technique to direct L2 learners' attention to their occurrence in various contexts; for instance, language teachers might raise L2 learners' awareness of the probability of producing several alternative phrases to '*very different from*' by replacing '*very*' with other possible adverbs including '*so, little, fundamentally, etc.*' (Nation, 2001). Similarly, dictogloss, a task-based activity that requires L2 learners to produce 'semantic approximations' to a dictated text (Wajnryb, 1990), can prime L2 learners' noticing of the frequent occurrence of formulaic sequences in written texts (Wood, 2009b). In this activity, L2 learners listen to a text twice, jot down content words, and then reconstruct the whole text (Wajnryb, 1990; Wood, 2009b). Not only does dictogloss give L2 learners the chance to perceive formulaic sequences in texts, but "it can also help them to retain the sequences by having them focus on their constituent parts and see how they fit into the flow of discourse" (Wood, 2009b, p. 11).

Last but not least, chain dictations and student dictations can facilitate the automatization of formulaic sequences. As for chain dictations, students can be divided into groups and each time a group of learners goes to the teacher, listens to the dictation, and returns to their group to repair and complete the text. To Wood (2010a), this type of dictation is specially useful for noticing formulaic language because L2 learners, due to the length of texts, "need to attend to the chunks or formulas it contains in order to minimize processing load" (p. 202).

Student dictation, in turn, is another beneficial activity to deal with "formulaic sequences and effective communication" (Wood, 2010a, p. 202). In this activity, where learners are assigned half a text to be dictated to a partner who has the other half, L2 learners' attention is directed to

the formulaic sequences that are embedded in the text. In other words, they “notice them and retain them in working memory in order to complete the gap-filling mutual dictation” (Wood, 2010a, p. 203). More importantly, L2 learners need to “negotiate word-by-word in order to complete a formulaic sequence, which forces a focus on its structure and component parts” (Wood, 2010a, p. 203).

The above-listed classroom activities and techniques can help develop a linguistic repertoire replete with formulaic sequences and augment L2 learners’ academic writing skills. However, given the vast number of formulaic sequences that are worth teaching and the time constraints of classroom-based language learning, it has been suggested that language teachers’ choice of chunks should depend on L2 learners’ objectives and needs to maximize the efficiency and usefulness of these activities (Boers & Lindstromberg, 2009; Lewis, 1997). In so doing, language teachers will provide their students with formulaic sequences that are prerequisites for effective communication in particular genres and contexts (Boers & Lindstromberg, 2009; Lewis, 2000c).

Summary of Chapter Two

This chapter has highlighted the growing interest in what is known to be the formulaicity of language production, whether spoken or written. It has been demonstrated, with reference to usage-based theory and corpus linguistics, that both first and second language users have a strong inclination to repeatedly deploy particular utterances in their language production despite the infinite sets of utterances that they can, in principle, construct based on the rules of syntax.

This chapter has provided an intensive account of the nature, types, and functions of formulaic sequences. It has been explained that, due to their holistic nature and various syntactic and pragmatic functions in discourse, these formulaic sequences can play a vital role in language production, particularly that of L2 learners. Their significant role includes, but is not limited to,

upgrading L2 learners' fluency, proficiency and pragmatic competence as well as augmenting L2 writing skills.

This chapter has displayed the complex nature of academic writing and the main obstacles that L2 learners encounter in academic contexts. These problems have been predominantly ascribed to the nature of academic prose, which is characterized by the frequent recurrence of the formulaic sequences typical of the field in question, and to L2 learners' failure to successfully incorporate formulaic sequences in their writing. Such failure might negatively affect the overall quality of L2 writing products, which may lead to negative or unfair evaluations at the academic level.

The importance of explicit and focused instruction of formulaic sequences has been demonstrated with reference to discussions proposed by a number of prominent figures in the field of second language acquisition. At this point, several scholars have advocated an explicit approach to teaching this language phenomenon due to its complex nature. They have as well suggested some activities that are believed to help the acquisition and the retention of formulaic sequences in the long-term memory. In sum, this review of literature has established a basis for the objectives and the design of this exploratory research study.

Chapter 3

Motivation for the Study

The third chapter sheds light on the research gap which has motivated this study. It also functions as a rationale for the research questions that are theoretically grounded in the detailed literature review explored above.

To start with, due to the complex nature of formulaic sequences, which are either partially transparent or entirely opaque to L2 learners, several scholars advocate an explicit and a focused teaching approach that aims at increasing the number of formulaic sequences used in L2 learners' academic writing (Boers & Lindstromberg, 2009; Cortes, 2004; Ellis et al., 2008; Jones & Haywood, 2004). Not only does focused instruction prime L2 learners' noticing of formulaic sequences in context and enhance their internalization into their linguistic repertoire, but it also provides L2 learners with the opportunity to practice the use of formulaic sequences in different tasks and, hence, promotes their implementation in writing (Boers & Lindstromberg, 2009; Cortes, 2004; Ellis et al., 2008; Granger, 1998; Hyland, 2007; Nation, 2002; O'Keeffe, 2012). However, such an emphasis on the effectiveness of focused instruction in promoting L2 learners' tendency to integrate formulaic sequences in their writing has not been empirically investigated yet in paragraph and/or essay writing.

It has also been posited that while integrating formulaic sequences into L2 learners' writing can foster their expressive skills and help them operate at a level acceptable in academic contexts, failure to do so may affect the overall quality and evaluation of their academic writing (Biber, 2006; Coxhead & Byrd, 2007; Granger, 1998; Hyland, 2007; Lewis, 1997; Nattinger & DeCarrico, 1992; Simpson-Vlach & Ellis, 2010). To Nattinger and DeCarrico (1992), teaching L2 learners the creative use of some sentence builders or frames can help them produce a

coherent and well-structured academic piece of writing. However, these claims need to be buttressed with research evidence.

It has been suggested that a mastery of formulaic sequences can upgrade L2 learners' overall accuracy. Schmitt (2000) and Hill (2000), for example, hypothesize that incorporating multiword sequences into classroom materials can reduce the number of lexical and syntactic errors in both speech and writing on the part of L2 learners. Similarly, Boers et al. (2006) postulate that formulaic sequences, when successfully stored in memory, may improve L2 learners' linguistic accuracy because they "constitute 'zones of safety' and appropriate use of them may thus confine the risk of 'erring' ... in one's discourse" (p. 247). These robust hypotheses have been mainly tested in speech production. However, it is worth mentioning that although Boers et al.'s (2006) study demonstrates an obvious increase in the number of formulaic sequences used by their participants, they conclude that formulaic sequences have no direct effects on L2 learners' perceived accuracy.

Boers et al.'s (2006) findings might be ascribed either to the nature of speech, which is, for the most part, instantaneous in that L2 learners have very little, if any, time to elaborate on the utterances produced when communicating in the target language, or to the kind of support that L2 learners have received during the study, i.e., chunk-noticing. Boers et al. (2006) conducted their research study on 32 university students majoring in English. The participants, whose English proficiency levels were either upper-intermediate or advanced, were divided into two groups, taught by the same instructor, and exposed to the same authentic materials. However, the only difference between both groups related to the emphasis that was placed on formulaic expressions in the experimental group. In other words, the experimental group's attention was directed to the target formulaic sequences and the participants "were encouraged to appreciate

the syntagmatic dimension of vocabulary” (Boers et al., 2006, p. 249). According to Schmitt (2000), L2 learners should be provided with activities that intensify deep processing, a prerequisite for developing both receptive and productive knowledge of lexical items. Thus, the participants, in the above-mentioned study, might have developed one type of knowledge, i.e., receptive, but not the productive one.

The importance of engaging L2 learners in activities that enhance deep processing, rather than awareness raising or noticing, to promote the successful acquisition and automatization of formulaic sequences has been mainly ascribed to the complex nature of this language phenomenon which is either transparent and, hence, passes unnoticed by many L2 learners, or entirely opaque and, thus, imposes a processing burden on L2 learners who will eventually disregard them in context (Boers & Lindstromberg, 2009; Ellis et al., 2008; Lewis, 2000b). To Wood (2009b), providing L2 learners with activities that prime deep processing, which force them to focus on the structures and functions of formulaic sequences, can help the successful acquisition and internalization of this language phenomenon into their linguistic repertoire. At this point, one can suggest that teaching L2 learners formulaic sequences explicitly and providing them with the appropriate practice will probably enhance L2 learners’ overall linguistic accuracy.

To investigate the effectiveness of the focused instruction of formulaic sequences in augmenting L2 learners’ academic writing skills and to bridge the above-highlighted gaps in literature, the following research questions have guided this study:

1. Will the participants integrate formulaic sequences in their writing as a result of focused instruction? If yes,
2. Will the participants receive better evaluation from the judges for the texts produced

- after the training period?
3. Will the increase, if any, in the use of the target formulaic sequences be associated with better evaluation?
 4. Will these formulaic sequences be treated as ‘zones of safety’ and links that lead to more proficient writing as far as content and language are concerned?

Chapter 4

Methodology

The present chapter offers a detailed description of the exploratory study designed to answer the research questions posited in the previous chapter.

The study was conducted over a ten-week period. Twelve participants volunteered in this study with two main motivations: to upgrade their academic writing skills or to get the required mark in the academic IELTS test for university admission. While nine participants were native Arabic speakers, the other three participants were from different L1 backgrounds, i.e., Chinese, Spanish, and Turkish.

The study took the form of focused instruction on formulaic sequences. Throughout the training period, the researcher met with the participants on a weekly basis for 90 minutes during which she explicitly taught the participants a number of formulaic sequences that were extracted from lists suggested by scholars or from the Corpus of Contemporary American English (COCA) due to their high frequency and direct relation to the genre in question, Economics. The focused instruction was followed by production and practice stages using both contextualized activities and decontextualized worksheets.

Each participant wrote three texts which were produced at three different points in time: prior to the training period (a pretest), at the end of the training period (a posttest), and fifteen to twenty days after the training period (a delayed posttest). The elicited texts were analyzed by the researcher following quantitative content analysis, and they were evaluated by three blind judges who are experienced in teaching English for Academic Purposes (EAP). Several statistical tests were also carried out to investigate any statistically significant differences in the use of the target formulaic sequences as well as the participants' writing proficiency.

Participants

The pool of participants comprised twelve L2 learners of English. During the training period, eleven participants were enrolled in different language programs: the English for Academic Purposes program (EAP), Carleton University, Ottawa, Canada and the Pre-English for Academic Purposes program (Pre-EAP), Algonquin College, Ottawa, Canada. Only one participant was not enrolled in any ESL programs but was preparing for the academic IELTS test.

The participants, seven males and five females, came from four different L1 backgrounds: nine Arabic speakers, one Spanish speaker, one Turkish speaker, and one Chinese speaker. It is worth noting that all the names used in this study are pseudonyms chosen by the participants to remain anonymous.

Salem was a thirty-six-year-old man from Yemen. He graduated from Cairo University, Egypt with a PhD degree in International Law. At the time the study was conducted, he had been in Canada for one year. Salem had taken two ESL courses (level four and level five) for six months at Algonquin College, and, during the training period, he was enrolled in ESLA 1300, Carleton University, Ottawa, Canada. He was taking English courses because he wanted to get a post-graduate diploma in Law from a Canadian University.

Mata was a twenty-one-year-old man from Yemen. He completed secondary education in his home country. Mata had arrived in Canada eight months prior to the training period to pursue his undergraduate degree in Civil Engineering. He had finished two ESL courses at Algonquin College, and he was taking ESLA1300 at Carleton University to satisfy the language proficiency requirement before starting a Bachelor's degree in engineering.

Ali was a twenty-nine-year-old man from Yemen. He earned a Bachelor's degree in Computer Science from Hadramout University, Yemen. Ali had been in Canada for eight months

during which he took two ESL courses at Carleton University, Ottawa, Canada. He had finished ESLA 1300 and was attending ESLA 1500 during the training period. Ali had to take these courses to satisfy the language proficiency requirement before applying for a Master's degree in Public Administration.

David was a twenty-year-old man from Yemen. He completed secondary education in his home country. David had been in Canada for eight months and had finished one ESL course at Algonquin College. During the training period, he was enrolled in ESLA 1500. He was taking English courses, concurrently with his undergraduate degree in engineering at Carleton University, to satisfy the language proficiency requirement.

Abedeer was a Jordanian woman in her forties. She had been in Canada for almost four years. She had completed her undergraduate degree in accounting in Jordan. Abedeer was taking ESLA 1500 at Carleton University to satisfy the language proficiency requirement for a Master's degree in Economics.

Mary was a thirty-one-year-old woman from Venezuela. After getting her Bachelor's degree in International Affairs from Venezuela, Mary completed her Master's in international law in Spain. She had been in Canada for eight months. She was taking ESLA 1900 at Carleton University to satisfy the language proficiency requirement before pursuing a Master's degree in Political Science.

Shireen was a Jordanian female in her thirties. She received her Bachelor's degree in Hospital and Health Care Administration from a Jordanian university. Shireen had been in Canada for ten years. During the training period, she was enrolled in ESLA 1900 at Carleton University to satisfy the language proficiency requirement for her Master's degree in Social Work.

Gabriella was a nineteen-year-old student from Turkey. She had completed her secondary education and moved to Canada three months prior to the training period to pursue a bachelor's degree in Economics at Carleton University. During the training period, Gabriella was enrolled in ESLA 1900, concurrently with three undergraduate courses at Carleton University, to satisfy the language proficiency requirement.

Moo was a twenty-year-old man from the Kingdom of Saudi Arabia. He had completed secondary education in his home country and had been in Canada for one year. He had taken three pre-EAP courses at Algonquin College and was enrolled in pre-EAP 3A during the training period. Moo was taking these courses to get the required mark in the academic IELTS test before starting his Bachelor's degree in Engineering.

Jacob was a nineteen-year-old student from Yemen. He had completed secondary education in his home country and moved to Canada to study Engineering. After arriving in Canada, three months before the study was conducted, Jacob started studying English at Algonquin College. He had finished pre-EAP 2B course and was enrolled in pre-EAP 3A during the training period. He was taking English courses at Algonquin College to sit the academic IELTS exam before applying to Carleton University for a Bachelor's degree in Engineering.

Ellie was a twenty-year-old man from Yemen. He had completed secondary education in Yemen. He had been in Canada for three months. Upon arrival to Canada, Ellie started taking English courses at Algonquin College. He had finished pre-EAP 3A and was enrolled, during the training period, in pre-EAP 3B. Like the other participants from Algonquin College, Ellie was taking English courses to sit the academic IELTS exam before applying to Carleton University for a Bachelor's degree in Engineering.

Eva was a thirty-five-year old woman from China. She received her Bachelor's degree in Social work from a Chinese University. She had been in Canada for seven years. During the training period, Eva was not enrolled in any ESL courses. She volunteered in this study because she was preparing for the academic IELTS test.

Instruments

The instruments were designed to gauge the effectiveness of the focused instruction of formulaic sequences in upgrading L2 learners' academic writing skills. In this respect, these instruments can be divided into two types: instruments used for data collection and instruments used during the training period.

Instruments used for data collection. The researcher designed a prompt for data collection, which included a line graph similar to the first writing task in the academic IELTS test (See Appendix B). It depicted the different changes that five commodity prices underwent between 2010 and 2011. The line graph, originally developed by the UN's Food and Agricultural Organization (FAO), was taken from a business report that highlighted the changes in commodity prices between May 2010 and April 2011 all over the world (BBC News: Business, May 31, 2013). The same prompt was assigned for eliciting the pretest, the posttest, and the delayed posttest from each participant.

The repeated use of the same prompt for data collection at different time intervals is pertinent to the aims of this study, i.e., investigating any significant differences in the frequency and occurrence of the target formulaic sequences before and after the training period and finding out whether the formulaic sequences used in the posttest would be internalized into the participants' linguistic repertoire and deployed in their delayed posttest, fifteen to twenty days after the training period. In this respect, scrutinizing any significant improvements in the frequency and

occurrence of the target formulaic sequences entailed using the same prompt lest any changes in the prompt might positively or negatively interfere in the participants' tendency to utilize the formulaic sequences presented during the training period.

In addition to the repeated use of the same prompt, the researcher's decision to choose a line graph similar to the first writing task in the academic IELTS test was influenced by four main factors: the language used in response to these graphs, the control required for external factors, the heterogeneous nature of the study group, and the need for more participants to volunteer in this study.

First, the tasks based on similar graphic prompts are replete with predefined formulaic sequences about market prices, which L2 learners can practice and use interchangeably. These formulaic sequences include, but are not limited to, discourse organizers, sentence builders, and collocations, which can be successfully employed to write a coherent academic piece of writing while reporting the trends presented in the graph. Such a fact is in line with Boers and Lindstromberg's (2009) call for the importance of being selective when presenting formulaic sequences to L2 learners. To them, given the time constraints in the language classroom and the vast number of formulaic sequences worth teaching, language teachers should select the formulaic sequences which best account for L2 learners' needs in a specific genre. Accordingly, using a prompt on price comparisons facilitated the selection of the target formulaic sequences peculiar to Economics.

Second, since almost all participants in this study were taking English classes at different sites concurrently with the training period, there was a need to choose a prompt that is not likely to be tackled in the language classroom to lessen, as much as possible, any factors that might result in internal validity threats. According to Dornyei (2007), one of the internal validity threats

that can invalidate quantitative research results is related to “any uncontrolled factors that can significantly modify the results” (p. 53). Accordingly, the choice of this task can be seen as an attempt on the part of the researcher to control any unexpected ‘factors’ that might interfere in the participants’ performance, including learning how to write a summary in response to a graphic prompt in their language program, for example.

Third, the training was offered to a heterogeneous group of participants from different proficiency levels, different ages, and different educational and L1 backgrounds. Such obvious differences in the study sample required the researcher to choose a model that would suit the needs and levels of all the participants. In this respect, the line graph is a typical example of the first writing task that test takers, regardless of their proficiency level, age, level of education, or L1 background, need to accomplish successfully should they decide to take the academic IELTS test.

Fourth, since the researcher did not know any of the participants before the training period, she tried to choose a topic for practice that would motivate as many participants as possible to volunteer in this study. Thus, when giving a short oral description of the study in three ESLA classes (two ESLA 1300 and one ESLA 1900) at Carleton University, the researcher clearly stated that participating in the study would be beneficial for any students who wanted to improve their academic writing skills in general and for those who were interested in writing the Academic IELTS test to satisfy the language proficiency requirement without having to spend up to three terms studying English for Academic Purposes.

Instruments used during the training period. Prior to the training period, the researcher designed twelve worksheets to help her introduce the participants to the target formulaic sequences and provide them with the opportunity to practice their use in both contextualized

activities and decontextualized worksheets (see Appendix B). It is worth mentioning that the design of the twelve worksheets was mainly based on the activities suggested by prominent figures in the field of second language acquisition including Boers and Lindstromberg (2009) Celce-Murcia and Olshtain (2000), Conzett (2000), Hunt and Beglar (2002), Lewis (1997), Nation (2001), Nation and Newton (1997), Nattinger and DeCarrico (1992), and Wood (2009b, 2010a).

The first worksheet introduces the participants to the function of the topic sentence and provides them with two examples of a good topic sentence for similar writing tasks. The examples are followed by a practice part in which the participants are required to write a topic sentence for five different graphic charts. The second worksheet is very similar to the first one, but it focuses on concluding sentences instead. Like the first worksheet, it begins with an explanation of the function of the concluding sentence and provides the participants with two examples that are followed by a practice part, i.e., writing a concluding sentence for five different graphic charts.

These two worksheets have a dual function. First and foremost, they raise the participants' awareness of the importance and function of introductory and concluding parts in academic writing. They also provide the participants with formulaic sequences that can be manipulated and used in various contexts. As Nattinger and DeCarrico (1992) suggest, having L2 learners practice the creative use of some variable lexical phrases, including topic nomination and topic organization, can help them use these multiword sequences as a backbone for a coherent and well-structured academic piece of writing. Moreover, the fact that the participants are required to use the same formulaic sequences repeatedly is expected to promote their subsequent acquisition (Conzett, 2000).

The third worksheet contains a list of collocations that can be used to describe trends presented in various economic charts. According to Nation and Newton (1997), discussing the meaning of different collocations that share in common the same stem and identifying their common underlying meaning can maximize L2 learners' perception of these lexical items as collocations rather than individual entities and, hence, help their acquisition as wholes rather than individual words.

The fourth worksheet includes twelve collocations to be written (two for each graph) under six line graphs designed by the researcher. The manual creation of the line graphs was aimed at helping the participants better understand the different meanings of the collocations presented. At this point, it has been suggested that associating lexical entities with visual images may enhance deep processing which can ultimately result in acquiring lexical items and internalizing them in long-term memory (Boers & Lindstromberg, 2009; Celce-Murcia & Olshtain, 2000; Hunt & Beglar, 2002). Moreover, such a practice, which was adapted from Nation's (2001) delayed production activity that requires the retention and production of formulaic sequences as chunks, motivates the participants to retain the proposed formulaic sequences in their working memory while looking for the correct graphic chart that best describes the formulaic sequence, a fact that might enhance deep processing of the target collocations.

The fifth worksheet provides the participants with a number of sentence builders, discourse organizers, and formulaic sequences which can be deployed to avoid repetition and connect words and/or sentences together in order to produce a well-connected piece of writing. The function of each formulaic sequence in this worksheet is presented separately with an example clarifying its use in context. In some occasions, more than one example is provided to highlight the various possible structures of a particular formulaic sequence.

The sixth worksheet, in turn, is designed to have the participants practice the formulaic sequences presented in the previous worksheet. It is composed of ten decontextualized sentences which they have to rewrite following the instructions in brackets. These two worksheets resonate with the notions that introducing L2 learners to formulaic sequences and having them practice their creative functions in discourse can upgrade L2 learners' academic writing skills (Conzett, 2000; Nattinger & DeCarrico, 1992; Wood, 2010a).

The seventh worksheet presents a number of sentence heads and stance bundles that can be used to prime the topic of discussion, express attitudes, or present causative relations in a text. Similar to the fifth worksheet, each sentence head is followed by its function, with an example or two to highlight its use and its various possible structures in context.

The seventh worksheet is followed by a practice stage at which the participants are asked to match formulaic sequences provided in *List A* with three possible endings presented in *List B*. These two lists are presented in the eighth worksheet which goes beyond a mere matching task in the sense that the participants, in addition to matching items together, are required to add some words in order to form accurate sentences out of the elements provided in *List A* and *List B*. The presentation of sentence heads followed by a matching task can raise L2 learners' awareness of the creative use of formulaic sequences in discourse (Lewis, 1997; Nation & Newton, 1997; Nattinger & DeCaricco, 1992).

The ninth worksheet consists of a table of formulaic sequences which are arranged thematically. This table includes both formulaic sequences presented in the above-listed worksheets and new ones. The presentation of a thematic table, which Conzett (2000) considers a typical prewriting activity, aims to familiarize L2 learners with the formulaic sequences that can be manipulated to accomplish various tasks; moreover, the discussion of thematically

classified formulaic sequences is likely to enhance deep processing which can, in turn, increase the chances of their automatization (Boers & Lindstromberg, 2009; Celce-Murcia & Olshtain, 2000; Hunt & Beglar, 2002; Wood, 2010a).

The tenth worksheet includes a text, adapted from *Insight into IELTS*, with five questions that require the participants to underline the formulaic sequences and collocations presented in the text. This particular worksheet has a dual function as it makes the target formulaic sequences salient and serves as a model which can be followed when encountering similar tasks. It is worth noting that this activity can be considered a modified version of the outline task suggested by Nattinger and DeCaricco (1992) in that both activities make the forms and functions of formulaic sequences in discourse more salient.

A cloze activity is presented in the eleventh worksheet, in which the participants are instructed to complete a text, extracted from *Insight into IELTS* along with a small paragraph written by the researcher herself, using the formulaic sequences in brackets. Such a task is likely to promote the acquisition and internalization of formulaic sequences into L2 learners' linguistic repertoire (Lewis, 1997).

The last practice during the training period was a modified dictogloss in which the text was dictated once and the participants were, then, given the twelfth worksheet which included the text used for the dictation but with the target formulaic sequences missing. The participants were required to complete the twelfth worksheet using the information that they had jotted down from the dictation. Such an activity can prime L2 learners' noticing of the target formulaic sequences and urge them to retain spoken chunks in their working memory before producing them chunk by chunk in a written form, a fact that might help their subsequent acquisition (Nation, 2001; Wood, 2009b).

Selection of Formulaic Sequences

The target formulaic sequences for this exploratory study were extracted from two sources. The first source included lists deemed to be formulaic by prominent scholars in the field of discourse analysis, namely, Biber et al. (2004), Hyland (2008), Lewis (1997), Nattinger and DeCarrico (1992), and Simpson-Vlach and Ellis (2010), due to their high frequency and important functions in academic prose. In addition to these lists, the researcher added some formulaic sequences and collocations that were validated by checking their frequency in the Corpus of Contemporary American English, the first large, genre-balanced English corpus including 450 million words which are “evenly divided between spoken, fiction, popular magazines, newspapers, and academic journals” (Davies, 2010, p. 447). When consulting the COCA for the frequency of a particular formulaic sequence in academic contexts, a minimum of thirty occurrences of a sequence in one million words was considered formulaic. The actual frequency cutoff adopted in this study is based on a more conservative approach than the one identified in Biber and Conrad (1999) and Hyland (2008), twenty occurrences per one million words.

The target formulaic sequences and collocations, as their functions indicate, operate as discourse devices or building blocks that bind the text together and help L2 learners in their construction of the text meaning (Biber et al., 2004; Hyland, 2008; Nattinger & DeCarrico, 1992). Among those formulaic sequences are:

1. ***There is/was a/an/not; There are a number of*** are two of the most frequent formulaic sequences utilized in academic discourse, both spoken and written; they function as typical clause or sentence stems that indicate the existence of something or its introduction for the first time (Hyland, 2008; Nattinger & DeCarrico, 1992; Simpson-Vlach & Ellis, 2010).

2. *As well as (the)*, referred to as a topic elaboration bundle or a discourse marker, is used to either elaborate on a topic for explicit comparison or to connect and signal transitions between the constituents of a phrase (Biber et al., 2004; Hyland, 2008; Simpson-Vlach & Ellis, 2010).
3. *At the beginning of the/ At the end of* are two multiword sequences which are known as referential bundles or attribute specifiers and can, based on context, serve various functions, among which are introducing temporal and spatial reference points in discourse (Biber et al., 2004; Hyland, 2008; Simpson-Vlach & Ellis, 2010).
4. *Due to the fact that/ Due to the* are typically used in discourse to introduce a degree of certainty and are referred to as stance bundles or intangible framing attributes (Biber et al., 2004; Hyland, 2008; Simpson-Vlach & Ellis, 2010).
5. *It is important/necessary/clear/obvious* is a formulaic sequence that, in addition to its structural function as a sentence builder, helps the writer present his/her evaluation of previously introduced information and express impersonal obligations or directions to the reader to perform a certain task; it is classified as an obligation stance bundle or an evaluation formula (Biber et al., 2004; Hyland, 2008; Nattinger & DeCarrico, 1992; Simpson-Vlach & Ellis, 2010).
6. *It is possible to/It is possible that* are categorized as stance features or impersonal stance attributes that are mainly implemented to foreground the writer's attitudes towards events while, at the same time, remaining detached from them; they can also be used to express the writer's opinion without explicitly identifying its source (Biber et al., 2004; Hyland, 2008; Simpson-Vlach & Ellis, 2010).

7. *On the one hand/ On the other hand* are polyword phrases referred to as discourse organizers or transition signals that have a degree of idiomaticity and are mainly employed to contrast elements in a text (Biber et al., 2004; Hyland, 2008; Lewis, 1997; Simpson-Vlach & Ellis, 2010).

8. *At the same time* is a discourse marker that connects and signals transition between constituents or clauses and helps writers structure their experiences of the real world by introducing time and place. (Biber et al., 2004; Hyland, 2008; Simpson-Vlach & Ellis, 2010).

9. *As a result of* is a resultative signal or topic elaboration bundle which is utilized to make an explicit reference to cause and effect relations in a text (Biber et al., 2004; Hyland, 2008; Nattinger & DeCarrico, 1992; Simpson-Vlach & Ellis, 2010).

10. *As shown in (figure)* is a text reflexive bundle whose main function is to make direct references to elements introduced in a text and to organize stretches of discourse (Biber et al., 2004; Hyland, 2008; Simpson-Vlach & Ellis, 2010).

11. *According to the* is one the most frequent three-word bundles in academic discourse. It helps introduce claims and information presented by sources other than the writer (Hyland, 2008; Nattinger & DeCarrico, 1992; Simpson-Vlach & Ellis, 2010).

12. *It can be seen/observed (that)* is a clause or sentence stem that is used to structure a text meaning and engage readers directly in the writer's attitude towards and/or evaluation of certain information already presented in a text (Hyland, 2008; Simpson-Vlach & Ellis, 2010).

13. *It has been (adverb) noted/asserted/believed* is a sentence head utilized to prime topics of discussion without directly attributing them to the writer (Hyland, 2008;

Nattinger & DeCarrico, 1992). Due to its holistic nature, this formulaic sequence links elements of new information with less processing time and, thus, is extensively used in academic prose (Hyland, 2008).

14. ***On the basis of*** is a framing signal that is used to structure and organize arguments in a text by identifying limiting conditions (Hyland, 2008).

15. ***In addition (to)*** is a transition signal that not only organizes a text meaning but also establishes additive links between discourse elements (Hyland, 2008).

16. ***Based on the/a*** is a formulaic sequence of high frequency in academic prose and is employed to frame an attribute of the following noun phrase (Simpson-Vlach & Ellis, 2010).

17. ***Play(s) an important role in*** is one of the most frequent five-word bundles in academic prose (Hyland, 2008). As its meaning indicates, it is manipulated to present a causative link between elements in a text.

18. ***For example*** is discourse organizer used as an exemplifier to introduce new elements that can frame segments of discourse (Simpson-Vlach & Ellis, 2010).

19. ***Such as*** is a typical expository phrase whose main function is to introduce exemplifications (Simpson-Vlach & Ellis, 2010).

20. ***In other words*** is a discourse organizer that connects and signals transitions in academic discourse and, in some contexts, functions as an exemplifier (Nattinger & DeCarrico, 1992; Simpson-Vlach & Ellis, 2010).

21. ***The/A (N) of*** is a tangible framing attribute that consists of a noun phrase with an of- phrase fragment; it is manipulated to either introduce the physical or measurable characteristics of the following noun phrase or indicate the relations between the phrase

elements (Hyland, 2008; Simpson-Vlach & Ellis, 2010). This formulaic sequence is “the most common structure overall, comprising about a quarter of all forms in the corpus” (Hyland, 2008, p. 10).

22. *(A) and (B)/ Both (A) and (B)* are sentence builders used to connect and compare elements of discourse (Nattinger & DeCarrico, 1992; Simpson-Vlach & Ellis, 2010).

23. *Over a period of/ An increase in the* are tangible framing attributes that refer to the physical or measurable characteristics of the following noun phrase (Simpson-Vlach & Ellis, 2010).

24. *The (X) shows/demonstrates* is a sentence head that is mainly used to demonstrate the existence of certain elements in discourse (Nattinger & DeCarrico, 1992).

25. *In a nutshell* is a formulaic sequence that embodies the idiomatic sense of language and is mainly used as a summarizer to shape a text meaning (Hyland, 2008; Nattinger & DeCarrico, 1992).

26. *In summary/ In sum/ In conclusion/ To sum up/ To conclude* are closing lexical phrases whose main function, as summarizers, is to bring an argument to an end (Nattinger & DeCarrico, 1992).

In addition to the above-listed formulaic sequences, the worksheets designed for the training period included four other formulas and a number of collocations that were extracted from the Corpus of Contemporary American English (see Appendix C). The inclusion of these multiword sequences was intended to provide the participants with more formulaic sequences that serve similar discourse functions and equip them with a list of collocations that help describe the trends presented in economic charts. As mentioned earlier, the frequency of these formulaic

sequences in academic contexts was validated using the *CHART* display in the COCA, where each word string should occur at list thirty times in that context to be considered formulaic.

The first formulaic sequence is *the (X) present(s)*. According to the COCA, this multiword sequence is most frequent in academic contexts (37.70 per one million words). It is presented as an alternative sentence head to *the (X) shows/demonstrates* suggested by Nattinger and DeCarrico (1992). In addition, *for instance* is added to the above-listed exemplifiers, *for example* and *such as*. This formulaic sequence occurs 95.53 times per one million words in academic contexts, which has the highest frequency of all. *Between (A) and (B)* and *From (A) to (B)* are also added to the list of formulaic sequences due to their high frequency in academic discourse (451.88 and 300.33 per one million words respectively). They are presented to provide the participants with formulaic sequences that help them introduce periods of time, an essential element to be reported in such writing tasks.

As for collocations, the list was created using the *Word and Phrase Info.* function in the COCA. In accordance with Philip's (2008) suggestion that some collocations are not entirely fixed in that they might be composed of a fixed part and a variable lexical slot that can be filled with semantically-related words, the researcher first examined the frequency of the stem words in academic contexts using the *CHART* display in the COCA. She then checked the collocation stem by using the *Word and Phrase Info.* function, which lists the most frequent words that tend to co-occur with certain lexis. It also arranges them from the most frequent collocates to the least frequent ones (see Appendix C). This list includes:

1. *Increase* is most frequent in academic contexts (257.36 per one million words).

Increase (V) + dramatically and significantly

Increase (N) + significant, dramatic, sharp, and slight

2. *Decrease* is also most frequent in academic contexts (47.57 per one million words).

Decrease (V) + significantly, slightly, and dramatically

Decrease (N) + significant, dramatic, slight, and sharp

3. *Remain* occurs 125.02 times per million words in academic contexts and has the highest frequency of all.

Remain + same

4. *Stable* is most frequent in academic contexts (54.95 per one million words).

Stable + remain

5. *Decline* occurs 73.76 times in one million words and is most frequent in academic contexts.

Decline (N) + steady, sharp, significant, rapid, gradual, and dramatic

6. *Rise* occurs 90.93 times per one million words in academic contexts and has the second most frequent occurrence of all.

Rise (V) + fall

Rise (N) + fall

Along with the proposed collocations, *Fluctuate* and *fluctuation* were added to the list despite their low frequency as collocation stems in the COCA. The rationale behind their inclusion as collocations was based on the fact that although they did not satisfy the minimum threshold set for the collocation stem, 30 occurrences per one million words, their high frequency in academic discourse, particularly that of science and technology, and their tendency to repeatedly co-occur with *price* (the most frequent noun that co-occurs with *fluctuate* and the second most frequent noun used with *fluctuation*) contributed to viewing *fluctuate + price* and *fluctuation + price* as significant alternatives for the collocation *rise and fall*.

Procedure

The study took the form of focused instruction on formulaic sequences over a ten-week period of time. The researcher met with the participants on a weekly basis for approximately ninety minutes during which instruction was offered and data were collected. As for data collection, three timed paragraphs were elicited from each participant at three time points: at the beginning of the study before any instruction (a pretest); at the end of the study (a posttest); and fifteen to twenty days after the end of the study (a delayed posttest). At each stage of data collection, the participants were required to write a summary in response to the graphic prompt designed for data collection. The participants were given twenty minutes, the same time recommended for this task in the academic IELTS test. At the end of the time allowed, the researcher asked the participants to stop writing and hand in their paragraphs.

It is worth noting that the elicitation of three texts at three different time points was intended to probe the effectiveness of the focused instruction of formulaic sequences in enhancing their acquisition and internalization into the participants' linguistic repertoire. In other words, while the pretest was used to identify the formulaic sequences that the participants had acquired before the training period, the posttest was assigned to highlight any increase in the frequency and occurrence of the target formulaic sequences, which could be attributed to the training period. The elicitation of the delayed posttest, in turn, aimed to scrutinize the automatization of the target formulaic sequences, which would be evident in the participants' ability to successfully retrieve and implement them fifteen to twenty days after the training period.

After eliciting the pretest from the participants, the researcher started teaching formulaic sequences explicitly following consciousness-raising, which consists of a presentation stage followed by a practice stage (Ellis, 2002; Richards, 2002). Consciousness-raising, in which a linguistic form is isolated and explicitly presented to students to raise their awareness of that

particular form and help them realize its occurrence and features in discourse, is particularly applied to teaching language forms that have low saliency (Ellis, 2002; Richards, 2002).

Moreover, it has been suggested that when a language form has high communicative function but is not salient for L2 learners, resorting to metalanguage can foster explicit knowledge; this kind of knowledge is essential for communicating these common types successfully (Ellis, 2002; Richards, 2002; Swan, 2002). The complexity of some less salient linguistic forms resembles, to some extent, that of formulaic sequences which are considered problematic for many L2 learners due to their partial transparency or total opacity (Boers & Lindstromberg, 2009; Cortes, 2004; Ellis et al., 2008; Jones & Haywood, 2004).

Following the above-mentioned teaching method, each week the participants' attention was explicitly directed to the usage and functions of a number of formulaic sequences. In other words, the participants were made aware of the importance of formulaic sequences as powerful means to avoid repetition, produce a coherent and cohesive piece of writing, and approach the genre in question. The presentation stage was followed by a practice stage at which the participants had to do some activities offered in the designed worksheets. The participants also had to practice the use of the target formulaic sequences in uncontrolled production by using them in meaningful sentences (Nation, 2001, 2002). However, only nine formulaic sequences were not explicitly taught. They were rather made salient in the ninth worksheet, which was discussed orally, as alternatives to other formulaic sequences that were presented through focused instruction, and only two of these formulaic sequences were practiced at a later stage. It is worth noting that the maximum quantity of classroom time (approximately 80%) was devoted to presentation and practice with only 20% assigned for uncontrolled production.

The instruction sessions were held in the form of one-on-one tutoring as well as group teaching. The two distinct types of teaching were mainly determined by the participants' availability to volunteer in this study. Since all participants, except one, were enrolled in different ESL programs at different levels, they could attend the training on different days and at different times. Despite the difference in the start of the training period for most participants, the same amount of time and the same method of teaching were provided during the one-on-one tutoring and the group teaching.

After having presented all the target formulaic sequences selected for this study (40 formulaic sequences and 25 collocations in total), the posttest was elicited from the participants. Like the pretest, the participants were given twenty minutes to accomplish this task. Moreover, the researcher asked the participants not to refer back to any worksheets discussed during the training period but to rather depend on their memories when writing the second paragraph. Such a request was mainly influenced by the researcher's intention to investigate the number of formulaic sequences that had been successfully acquired as a result of the training offered.

On the last day of the focused instruction, the researcher scheduled the last meeting, fifteen to twenty days after the posttest. The final meeting was intended to elicit the delayed posttest in order to examine the number of formulaic sequences that the participants retained. Since the delayed posttest was an essential part of the data in this study, the researcher offered the participants different pedagogical inducements. These pedagogical inducements could be anything related to EAP including grammar teaching, proofreading, or some strategies before sitting the academic IELTS test. Thus, all participants were encouraged to contact the researcher before the final meeting to inform her of the topic they had chosen for consultation. At the last

meeting, the participants wrote the delayed posttest, for which twenty minutes were assigned. The remaining time of the session was devoted to consultation on the area they had chosen.

Data Analysis

Due to the fact that quantitative research studies do not “just categorize phenomena but measure them and express their relationships in terms of quantity” (Cress & Hesse, 2013, p. 93), the data for this study, thirty-six written texts produced by twelve participants at three different points in time, were analyzed using different quantitative data analysis methods to derive variables from the textual data and test the differences and associations between them statistically.

Before running any statistical tests, the textual data were quantified based on quantitative content analysis, a quantitative coding method which involves identifying the frequency and occurrence of words, phrases, or grammatical structures in a written sample (Cress & Hesse, 2013; Dornyëi, 2007; Jeong 2013). According to Jeong (2013), “quantitative content analysis focuses on quantifying textual data by using methods such as counting the frequencies of certain words present in texts or sets of texts” (p. 170). Moreover, unlike other coding methods in which “new categories have to be defined in an inductive way” (Cress & Hesse, 2013, p. 99), this method of textual analysis is driven by a ‘predefined coding schema’ to increase its reliability (Cress & Hesse, 2013; Dornyëi, 2007; Jeong 2013). In other words, the coding process should be guided by preexisting coding categories that, when used by other researchers to analyze the same data, should lead to the same results (Cress & Hesse, 2013). Furthermore, quantitative content analysis can be seen as a “manifest level analysis, because it is an objective and descriptive account of the surface meaning of data” (Dornyëi, 2007, p. 246).

The criteria determining the application of quantitative content analysis are in line with the objectives behind analyzing the data in this study, i.e., using a predefined list of categories to highlight the frequency, i.e., the total number of the target formulaic sequences used regardless of their repetition, and occurrence, the types of the formulaic sequences used in each text in that each type is counted once, of the target formulaic sequences at three different production levels. Accordingly, the researcher developed a predefined set of categories, as shown in Table 1 and Table 2 below, to identify the target formulaic sequences implemented in the elicited data.

Table 1	
<i>A Predefined Categories List</i>	
Collocations	Codes
Increase dramatically/ significantly	Using a collocation to describe trends presented in the line graph
Decrease significantly/ slightly/dramatically	
Rise and fall	
Prices fluctuate	
Remain the same/ Remain stable	
Significant/dramatic/sharp/slight increase	
Significant/dramatic/sharp/slight decrease	
Steady/sharp/significant/rapid/gradual/dramatic decline	
The rise and fall	
The fluctuation in the price	
<i>Note.</i> All the collocations in this table are extracted from the Corpus of Contemporary American English (COCA) using <i>Word and Info.</i> Function.	

Table 2	
<i>A Predefined Categories List</i>	
Formulaic Sequences	Codes
1. There is/are/was/were	Using a sentence stem to indicate that something exists
2. At the beginning/end of	Using a referential bundle to introduce a period of time
3. Due to the (fact that)	Using a stance bundle to introduce a degree of certainty
4. It is important/necessary/clear/obvious	Using a sentence builder to evaluate information presented earlier
5. It is possible to/that	Using a stance attribute to foreground the writer's attitude implicitly
6. On the one hand On the other hand	Using a transition signal to contrast elements in a text
7. At the same time	Using a discourse marker to signal transition between text elements
8. As a result of	Using a resultative signal to refer to cause and effect
9. As shown in (figure)	Using a reflexive bundle to make a direct reference to text elements
10. According to the	Using a three-word bundle to introduce information presented by sources other than the writer
11. It has been (adverb) believed/asserted/noted	Using a sentence head to prime topics of discussion
12. It can be seen/observed that	Using a sentence head to structure text meaning
13. On the basis of	Using a framing signal to organize arguments in a text
14. In addition to	Using a transition signal to organize text meaning and establish additive links between discourse elements
15. Based on a/the	Using a formula to frame an attribute of the following noun phrase

16. play(s) an important role in	Using a five-word bundle to present causative links between discourse elements
17. For example	Using a discourse organizer to introduce new elements that can frame segments of discourse
18. Such as	Using an expository phrase to introduce exemplifications
19. In other words	Using a discourse organizer to signal transitions between discourse elements
20. The/A (N) of	Using a tangible framing attribute to introduce the characteristics of the following noun phrase
21. (A) and (B) (A) as well as (B) Both (A) as well as (B)	Using a sentence builder to connect and compare elements of discourse.
22. The (X) shows/demonstrates	Using a sentence head to demonstrate the existence of discourse elements
23. In a nutshell	Using a multiword sequence to summarize and shape text meaning
24. In summary In sum In conclusion To sum up To conclude	Using a closing lexical phrase to bring an argument to an end
25. The (X) presents	Using a sentence head to demonstrate the existence of discourse elements
26. For instance	Using a discourse organizer to introduce new elements that can frame segments of discourse
27. From (A) to (B)	Using a formula to introduce a period of time
28 Between (A) and (B)	Using a formula to introduce a period of time
<p><i>Note.</i> The categories and codes used in this table are based on the lists provided by Biber et al. (2004), Hyland (2008), Lewis (1997), Nattinger and DeCarrico (1992) and Simpson-Vlach and Ellis (2010) except the last four formulaic sequences whose frequency was validated by the researcher using the COCA.</p>	

Quantifying the textual data by manually coding for the frequency and occurrence of the target formulaic sequences was an indispensable process before computing any statistical tests using the Statistical Package for the Social Sciences (SPSS) version 22.

In addition to the variables derived from the manual coding process, three judges, English language teachers who are experienced in teaching EAP, were asked to evaluate the thirty-six elicited texts. The evaluation of the judges was an elicitation instrument intended to obtain objective evaluation that could be used at the data analysis stage (Dornýei, 2007). Accordingly, the elicited texts from each participant were arranged randomly on one sheet and the three judges evaluated them individually on a scale from one to ten (one is the lowest and ten is the highest) without knowing which reports were written before or after the training period. The judges were not provided with any rubric or training for evaluation, but they rather evaluated the texts based on their overall content and form.

Since researchers cannot assume that variations identified in the raw scores reflect any ‘real’ difference without computing a statistical test to determine whether they “have got a generalizable result or whether the score is a random artefact of random variation” (Dornýei, 2007, p. 215), several statistical tests were computed on the raw scores for the frequency and occurrence of the target formulaic sequences and the three judges’ evaluation. These tests included, but were not limited to, measures of central tendency and variability, Cronbach’s alpha, paired-samples t-tests, and Pearson product-moment correlation coefficients.

First, measures of central tendency and measures of variability were computed to summarize and gauge the variance of the scores for the frequency and occurrence of the target formulaic sequences in the pretest, posttest, and delayed posttest and the evaluation of the three judges.

According to Dornyei (2007), these descriptive statistics “are indispensable when we want to share our results... and they also form the base of further inferential statistics” (p. 213).

Second, to measure the internal consistency of the scores for the judges’ evaluation and estimate their reliability, Cronbach’s alpha was computed. According to Dornyei (2007), Cronbach’s alpha is an internal consistency coefficient which indicates “how the different scores hang together” (p. 51). Hence, it is an essential reliability analysis that is computed to assess the degree to which the three different raters have made consistent estimates of the same test and to reveal any inconsistencies in the variables obtained using such elicitation instruments. It is worth noting here that the level of consistency should be a minimum of .60 to be acceptable (Dornyei, 2007; Gliem & Gliem, 2003). Dornyei (2007) suggests that “if the Cronbach Alpha of a scale does not reach 0.60, this should sound warning bells” (p. 207).

Third, paired-samples t-tests were computed to highlight any statistically significant increase in the scores for the variables derived based on the coding process and the three judges’ evaluation. Paired-samples t-tests, also known as matched t-tests or pairs t-tests, are particularly computed in the case of a repeated measures design to compare two variables within the same group and highlight any significant differences in the group’s performance before and after the treatment (Dornyei, 2007). Moreover, their results are robust in the presence of small samples and slight departures from normality (Bachman, 2004; Lumley, Diehr, Emerson, & Chen, 2002). As Bachman (2004) puts it, the t-test is robust in the case of normality violations because “recent research has shown that a violation of this assumption has almost no practical consequences for using the two-tailed t-test” (p. 236). Accordingly, five paired-samples t-tests were computed to compare the frequency and occurrence of the target formulaic sequences at each production stage as well as the three judges’ evaluation of the textual data.

It is worth noting that before computing paired-samples t-tests, tests of normality were carried out to investigate the distribution of each dependent variable (see Appendix D). To this end, the differences in the dependent variables of each pair were computed and tested statistically to investigate their distribution by calculating the skewness and Kurtosis z-values: if z-values are within +/- 1.96, the dependent variable is approximately normally distributed (Doane & Seward, 2011). Shapiro-Wilk tests of normality were also run to confirm or reject the null hypothesis for this test of normality: the data are normally distributed if $p > .05$ (Razali & Wah, 2011; Shapiro & Wilk, 1965).

Fourth, to “examine the relationship between variables ... and evaluate the strength and direction of their relationship” (Dornyei, 2007, p. 223), a correlation coefficient was computed to examine the association between the frequency and occurrence of the target formulaic sequences in the pretest, posttest, and delayed posttest and the evaluation of the three judges. As Dornyei (2007) puts it, a correlation coefficient “allows us to look at two variables and evaluate the strength and direction of their relationship or association with each other” (p. 223). Since both the frequency and occurrence of formulaic sequences in the textual data (i.e., independent variables) and the judges’ evaluation (i.e., dependent variables) are interval variables, three Pearson product-moment correlation coefficient matrices were generated to investigate whether an increase in one variable was associated with an increase in the other.

However, before conducting correlation coefficients, the researcher computed some statistical tests to verify that the data could be analyzed using a parametric test. Dornyei (2007) states that “parametric tests require interval data and ... the interval data also needs to be normally distributed” (p. 208). As mentioned above, the data are interval in that the scores for the frequency and occurrence of the target formulaic sequences as well as the judges’ evaluation

exist at an equal distance on a continuum. Moreover, to assess the normal distribution of the dependent variables, the evaluation of each judge was examined graphically by generating histograms along with measures of central tendency and variability (see Appendix D).

Approximately normally distributed variables were considered significant since “normality does not have to be perfect because most procedures work well with data that is only approximately normally distributed” (Dornyöi, 2007, p. 208).

Variables

In this study, thirty-five variables were analyzed using SPSS. While twenty variables were extracted from the manual coding process and the evaluation of the three judges, fifteen dependent variables were derived by computing the difference in the scores for each pair tested statistically based on a paired-samples t-test to investigate their distribution. The variables derived from the coding process and the judges' evaluation were:

- **The frequency of formulaic sequences in the pretest** is the first variable in this study and represents the number of formulaic sequences used by the twelve participants before the training.
- **The frequency of formulaic sequences in the posttest** is the second variable and indicates the total number of formulaic sequences utilized by the twelve participants in the posttest, which was elicited after the training.
- **The frequency of formulaic sequences in the delayed posttest** is the third variable and identifies the number of formulaic sequences implemented by the twelve participants in the delayed posttest, which was completed fifteen to twenty days after the training. The first three variables, which were derived based on the manual coding process, identify all

the target formulaic sequences deployed by the participants at three production stages, including the repeated ones.

- **The occurrence of formulaic sequences in the pretest** is the fourth variable and indicates the types used by the twelve participants in the texts elicited before the training.
- **The occurrence of formulaic sequences in the posttest** is the fifth variable and reveals the types used by the participants when they completed the posttest.
- **The occurrence of formulaic sequences in the delayed posttest** is the sixth variable that depicts the types of formulaic sequences utilized by the participants fifteen to twenty days after the training period. Variables four, five, and six represent the types used by the twelve participants at each production stage. In other words, each instance of the target formulaic sequences was counted once and all the repetitions of the same formulaic sequence in one text were disregarded.
- **Judge 1: Pretest** is the seventh variable and presents the first judge's evaluation of the textual data elicited from the twelve participants before the training.
- **Judge 1: Posttest** is the eighth variable and indicates the first judge's evaluation of the posttest completed by the twelve participants after the training period.
- **Judge 1: Delayed posttest** is the ninth variable and represents the scores for the first judge's evaluation of the delayed posttest.
- **Judge 2: Pretest** is the tenth variable and identifies the second judge's evaluation of the twelve participants' pretest.
- **Judge 2: Posttest** is the eleventh variable and illustrates the second judge's evaluation of the participants' posttest.

- **Judge 2: Delayed posttest** is the twelfth variable and represents the scores for the second judge's evaluation of the textual data elicited fifteen to twenty days after the training.
- **Judge 3: Pretest** is the thirteenth variable and presents the third judge's evaluation of the participants' pretest.
- **Judge 3: Posttest** is the fourteenth variable in this study and includes the scores for the third judge's evaluation of the posttest.
- **Judge 3: Delayed posttest** is the fifteen variable and represents the third judge's evaluation of the delayed posttest.
- **Frequency of formulaic sequences in all elicited texts** is the sixteenth variable and includes the frequency of the target formulaic sequences in the participants' pretest, posttest and delayed posttest.
- **Occurrence of formulaic sequences in all elicited texts** is the seventeenth variable. It incorporates the occurrence of the target formulaic sequences in the pretest, posttest, and delayed posttest of the study sample.
- **Judge 1 evaluation** is the eighteenth variable. It comprises the first judge's evaluation of the pretest, posttest and the delayed posttest.
- **Judge 2 evaluation** is the nineteenth variable and consists of the scores for the second judge's evaluation of all the elicited texts.
- **Judge 3 evaluation** is the twentieth variable which consists of the third judge's evaluation of the participants' pretest, posttest, and delayed posttest.

The variables computed on SPSS included:

- **Frequency difference: Pretest and posttest** is the first dependent variable derived by using the *Compute Variables* function on SPSS and subtracting the frequency of the

target formulaic sequences in the pretest from the frequency of the target formulaic sequences in the posttest.

- **Frequency difference: Pretest and delayed posttest** is the second dependent variable that is derived by subtracting the frequency of the target formulaic sequences in the pretest from the frequency of the target formulaic sequences in the delayed posttest.
- **Frequency difference: Posttest and delayed posttest** is the third dependent variable and represents the paired difference between the frequency of the target formulaic sequences in the posttest and the delayed posttest.
- **Occurrence difference: Pretest and posttest** is the fourth dependent variable. It is constructed by subtracting the variables that represent the occurrence of the target formulaic sequences in the pretest and the posttest.
- **Occurrence difference: Pretest and delayed posttest** is the fifth dependent variable. It is derived by calculating the difference in the scores for the occurrence of the target formulaic sequences in the pretest and the delayed posttest.
- **Occurrence difference: Posttest and delayed posttest** is the sixth dependent variable and is derived by subtracting the occurrence of the target formulaic sequences in the posttest from the occurrence of the target formulaic sequences in the delayed posttest.
- **Evaluation difference J1: Pretest and posttest** is the seventh dependent variable that is constructed by calculating the difference in the scores for the first judge's evaluation of the pretest and the posttest.
- **Evaluation difference J1: Pretest and delayed posttest** is the eighth dependent variable and represents the difference in the scores for the first judge's evaluation of the pretest and the delayed posttest.

- **Evaluation difference J1: Posttest and delayed posttest** is the ninth dependent variable and is derived by subtracting the evaluation of the posttest from the evaluation of the delayed posttest.
- **Evaluation difference J2: Pretest and posttest** is the tenth dependent variable that is constructed by calculating the difference in the scores for the second judge's evaluation of the pretest and the posttest.
- **Evaluation difference J2: Pretest and delayed posttest** is the eleventh dependent variable and represents the difference in the scores for the second judge's evaluation of the pretest and the delayed posttest.
- **Evaluation difference J2: Posttest and delayed posttest** is the twelfth dependent variable and is derived by subtracting the evaluation of the posttest from the evaluation of the delayed posttest.
- **Evaluation difference J3: Pretest and posttest** is the thirteenth dependent variable that is derived by subtracting the scores for the third judge's evaluation of the pretest from his evaluation of the posttest.
- **Evaluation difference J3: Pretest and delayed posttest** is the fourteenth dependent variable and represents the paired difference between the pretest evaluation and the delayed posttest evaluation.
- **Evaluation difference J3: Posttest and delayed posttest** is the fifteenth dependent variable. It is derived by calculating the difference in the scores for the third judge's evaluation of the posttest and the delayed posttest.

In sum, this chapter provided a thorough account of the research design. To answer the research questions that motivated this study, a quantitative research study was designed and

conducted over a ten-week period. The twelve participants received focused instruction on formulaic sequences and were required to write three timed texts to gauge the effectiveness of the focused instruction in augmenting L2 learners' academic writing skills. Throughout the training period, the researcher used different instruments to make the target formulaic sequences salient for the participants and provide them with the opportunity to practice their use in both contextualized activities and decontextualized worksheets. All the elicited texts were evaluated by three judges and coded by the researcher based on quantitative content analysis. The variables were statistically tested by computing descriptive statistics, Cronbach's alpha, paired-samples t-tests, and Pearson product-moment correlation coefficients.

Chapter Five

Findings

The present chapter addresses the areas of focus of this study: the tendency of L2 learners to increasingly integrate various formulaic sequences as a result of focused instruction; the better evaluation that they might receive after the focused instruction, and the possible association between the better evaluation and the increase in the usage of formulaic sequences. To this end, this chapter presents the findings of the quantitative analyses of the textual data elicited from the twelve participants. As mentioned in the previous chapter, the textual data were collected at three different points in time, i.e., at the beginning of the training period before any instruction (a pretest), at the end of the training period (a posttest), and fifteen to twenty days after the training period (a delayed posttest).

This chapter will start by highlighting the results of the manual coding process, which was conducted based on quantitative content analysis to identify the frequency and occurrence of the target formulaic sequences in the textual data. It will also offer a detailed account of all the statistical tests computed on the raw data to highlight any significant differences that can be attributed to the focused instruction on formulaic sequences.

The findings presented in this section are mainly based on quantitative content analysis and the statistical tests computed on the raw scores that were obtained from the manual coding and the three judges' evaluation of the textual data (see Appendix A). The statistical tests include measures of central tendency, measures of variability, Cronbach's alpha, paired-samples t-tests, and Pearson product-moment correlation coefficients.

Quantitative Content Analysis of Data

Despite the differences in their proficiency level, almost all the participants have exhibited a

tendency, as shown in Table 3 below, to increasingly manipulate formulaic sequences in the elicited posttest and delayed posttest.

Table 3			
<i>The Frequency of Formulaic Sequences Used in the Pretest, the Posttest, and the Delayed Posttest of Each Participant.</i>			
Participant	Pretest	Posttest	Delayed Posttest
Salem	5	14	18
Ellie	6	17	19
MATA	10	26	29
Jacob	12	35	38
Shireen	6	17	19
Abedeer	7	19	20
David	9	21	19
Mary	7	22	18
Eva	6	21	16
Ali	13	34	23
Moo	2	19	16
Gabriella	18	15	15
<i>Note.</i> This table presents the number of all the target formulaic sequences used by the participants, including the repeated ones.			

According to Table 3, all participants, except one, integrated more formulaic sequences in their posttest and delayed posttest than they did in their pretest. To start with, almost 50% of the participants' texts demonstrated a greater number of formulaic sequences in the posttest as compared to the pretest. These numbers slightly declined in the delayed posttest but remained

higher than those utilized in the pretest. For example, the frequency of formulaic sequences in Ali's production increased from 13 in the pretest to 34 in the posttest. It then declined to 23 in the delayed posttest, which is still higher than the pretest. Likewise, David's usage of formulaic sequences, 9 in the pretest, 21 in the posttest and 19 in the delayed posttest, corresponds to that of Ali's, where more formulaic sequences were detected after the training period despite the slight decline in the frequency of the target formulaic sequences in their delayed posttest.

Moreover, six participants, that is 50% of the study sample, manipulated more formulaic sequences in their delayed posttest than they did in their posttest. For example, Jacob's use of formulaic sequences increased from 12 in the pretest to 35 and then 38 in the posttest and delayed posttest, respectively. Similarly, MATA employed 26 formulaic sequences in the posttest and 29 in the delayed posttest, which shows an obvious increase in the number of formulaic sequences compared to only 10 formulaic sequences in the pretest.

However, only one participant, Gabriella, used fewer formulaic sequences in her posttest and delayed posttest than she did in her pretest: she implemented 18 formulaic sequences in the pretest, which then declined to 15 in both the posttest and the delayed posttest.

More importantly, this obvious increase in the frequency of the target formulaic sequences in the production of almost all the participants was not due to repeating the same formulaic sequences, but it was rather the result of utilizing a wider range of different types of formulaic sequences in their posttest and delayed posttest, as demonstrated in Table 4 below. It is worth noting that the scores in Table 4 represent the occurrence of the target formulaic sequences in the data, i.e., each instance is counted once, and all the repetitions of the same formulaic sequence are overlooked.

Table 4			
<i>The Occurrence of Formulaic Sequences Used in the Pretest, the Posttest, and the Delayed Posttest of Each Participant.</i>			
Participant	Pretest	Posttest	Delayed Posttest
Salem	4	11	13
Ellie	4	13	14
MATA	6	19	18
Jacob	7	17	17
Shireen	5	13	14
Abedeer	3	15	14
David	7	14	17
Mary	4	12	11
Eva	4	12	11
Ali	8	20	15
Moo	2	15	12
Gabriella	9	12	12
<i>Note.</i> In this table, each target formulaic sequence is counted once no matter how many times it has occurred in each text.			

Based on Table 4, all the participants augmented their writing with more types of formulaic sequences after the training period. The tendency to manipulate the greatest variation of types in the posttest was evident in the production of six participants, that is, 50% of the investigated sample. The types used by Abedeer, for example, increased from 3 in the pretest to 15 in the posttest, and this increase remained relatively stable in the delayed posttest, i.e., 14 types. Likewise, Eva's usage of types rose from 4 in the pretest to 12 and 11 in the posttest and the delayed posttest, respectively.

In addition, four participants, namely, Salem, Ellie, Shireen, and David, showed the greatest variation in the use of formulaic sequences in the delayed posttest. For instance, the types used by Salem increased from 4 in the pretest to 11 in the posttest, and it then again increased to 13 in the delayed posttest. Shireen as well used more types in the delayed posttest, 14, than she did in the pretest and posttest, 5 and 13 respectively.

Last but not least, two participants, namely, Jacob and Gabriella, implemented the same number of types in their posttest and delayed posttest, which is remarkably higher than that in the pretest. While Jacob deployed 17 types in the posttest and delayed posttest, that is 10 types more than the ones integrated in the pretest, Gabriella utilized 12 types in the posttest and 12 types in the delayed posttest after using 9 in the pretest.

Descriptive Statistics

Measures of central tendency and measures of variability were computed to describe the data set and investigate the dispersion of the raw scores which represent the frequency and occurrence of formulaic sequences in the pretest, posttest, and delayed posttest of the twelve participants as well as the evaluation of the three judges.

Measures of central tendency and variability for the frequency of the target formulaic sequences in the pretest, posttest, and delayed posttest of the total study sample are shown in Table 5 below. They summarize the scores for the frequency of formulaic sequences at three production stages and display their variability. They also reveal the minimum and maximum numbers of the target formulaic sequences utilized at each production stage.

Table 5

Descriptive Statistics: The Frequency of Formulaic Sequences at Three Production Stages

		Frequency: Pretest	Frequency: Posttest	Frequency: Delayed Posttest
N	Valid	12	12	12
	Missing	0	0	0
Mean		8.4167	21.6667	20.8333
Median		7.0000	20.0000	19.0000
Mode		6.00	17.00 ^a	19.00
Std. Deviation		4.29499	6.81353	6.56206
Variance		18.447	46.424	43.061
Range		16.00	21.00	23.00
Minimum		2.00	14.00	15.00
Maximum		18.00	35.00	38.00

a. Multiple modes exist. The smallest value is shown

According to the mean scores in Table 5, almost all the participants implemented a greater number of the target formulaic sequences in the posttest ($M = 21.66$, $SD = 6.81$) and the delayed posttest ($M = 20.83$, $SD = 6.56$) than they did in the pretest ($M = 8.41$, $SD = 4.29$). Moreover, the scores for the minimum frequency of formulaic sequences in the pretest, posttest, and the delayed posttest indicate that the lowest frequency of the target formulaic sequences increased from 2 in Moo's pretest to 14 in Salem's posttest and 15 in Gabriella's delayed posttest. Likewise, the maximum number of the target formulaic sequences increased from 18 in Gabriella's pretest to 35 and then 38 in Jacob's posttest and delayed posttest, respectively. In this respect, the mean scores along with the minimum and maximum scores for the frequency of the target formulaic sequences demonstrate that almost all the participants utilized a significantly greater number of formulaic sequences in their writing after the training period.

Table 6 below summarizes the scores for the occurrence of the target formulaic sequences in

the pretest, the posttest, and the delayed posttest of the total sample and displays their variability. It also reveals their minimum and maximum scores for their occurrence at each production stage.

Table 6

Descriptive Statistics: The Occurrence of Formulaic Sequences at Three Production Stages

		Occurrence: Pretest	Occurrence: Posttest	Types: Delayed posttest
N	Valid	12	12	12
	Missing	0	0	0
Mean		5.2500	14.4167	14.0000
Median		4.5000	13.5000	14.0000
Mode		4.00	12.00	14.00
Std. Deviation		2.13733	2.90637	2.37410
Variance		4.568	8.447	5.636
Range		7.00	9.00	7.00
Minimum		2.00	11.00	11.00
Maximum		9.00	20.00	18.00

All the participants, as shown in Table 6, augmented their writing with more types of the target formulaic sequences in the posttest ($M = 14.41$, $SD = 2.90$) and the delayed posttest ($M = 14$, $SD = 2.37$), than they did in the pretest ($M = 5.25$, $SD = 2.13$). As well, the scores for the minimum occurrence of the target formulaic sequences increased from 2 types (used by Moo in the pretest) to 11 types in both the posttest (implemented by Salem) and the delayed posttest (employed by Mary and Eva). The scores for the maximum occurrence of the target formulaic sequences also increased from 9 in Gabriella's pretest to 20 in Ali's posttest. This occurrence then slightly decreased in the delayed posttest, 18 types used by MATA, which was remarkably higher than that in the pretest.

Measures of central tendency and variability were also computed on the scores for the three judges' evaluation of the textual data. Table 7 below summarizes these scores.

Table 7

Descriptive Statistics: The Evaluation of the First Judge

		Judge 1: Pretest	Judge 1: Posttest	Judge 1: Delayed posttest
N	Valid	12	12	12
	Missing	0	0	0
Mean		5.9167	7.0833	7.2917
Median		6.0000	7.2500	7.2500
Mode		6.00	7.50	7.00 ^a
Std. Deviation		.63365	.59671	.68948
Variance		.402	.356	.475
Range		2.50	2.00	3.00
Minimum		5.00	6.00	6.00
Maximum		7.50	8.00	9.00

a. Multiple modes exist. The smallest value is shown

The increase in the mean scores from ($M = 5.91$, $SD = .63$) for the pretest to ($M = 7.08$, $SD = .59$) for the posttest and, then, to ($M = 7.29$, $SD = .68$) for the delayed posttest demonstrated that the texts which were elicited after the focused instruction on formulaic sequences received better evaluation from the first judge. The higher evaluation, which almost all the participants received for their production after the training, was as well reflected in the minimum and maximum scores for the evaluation of the three written samples. The minimum scores for the evaluation of the elicited texts increased from 5 for Moo's pretest to 6 for both Shireen's posttest and Gabriella's delayed posttest. The maximum scores, in turn, increased from 7.5 for Ali's pretest to 8 for Abedeer's posttest and 9 for Ali's delayed posttest.

In addition to the first judge's evaluation, the scores for the evaluation of the second judge are summarized in Table 8 below.

		Judge 2: pretest	Judge 2: Posttest	Judge 2: Delayed Posttest
N	Valid	12	12	12
	Missing	0	0	0
Mean		5.3333	6.9167	6.8750
Median		5.5000	7.2500	7.0000
Mode		5.50 ^a	5.50 ^a	7.00
Std. Deviation		1.15470	1.53495	1.15059
Variance		1.333	2.356	1.324
Range		3.50	5.00	4.00
Minimum		3.50	4.00	5.00
Maximum		7.00	9.00	9.00

As shown in Table 8, the increase in the mean scores from ($M = 5.33$, $SD = 1.15$) for the pretest to ($M = 6.91$, $SD = 1.53$) for the posttest revealed that the evaluation of the posttest was significantly higher than that of the pretest. Moreover, the slight decrease in the mean scores for the delayed posttest ($M = 6.87$, $SD = 1.15$), which remained relatively higher than the mean scores for the pretest (a difference of -1.54), demonstrated that although the second judge's overall evaluation of the delayed posttest was not as high as that of the posttest, his evaluation of the delayed posttest was statistically significantly higher than that of the pretest. The better evaluation that the participants received for their writing after the training period can as well be detected in the minimum and maximum scores for the evaluation of each test. The minimum scores for the second judge's evaluation increased from 3.5 for Salem's pretest to 4 and then to 5 for Shireen's posttest and delayed posttest, respectively. The maximum scores also increased from 7 for MATA's pretest to 9 for both Abedeer's posttest and MATA's delayed posttest.

The last measures of central tendency and variability were computed on the scores for the third judge's evaluation to summarize them and display their variability.

Table 9

Descriptive Statistics: The Evaluation of the Third Judge

		Judge 3: pretest	Judge 3: Posttest	Judge 3: Delayed Posttest
N	Valid	12	12	12
	Missing	0	0	0
Mean		6.2917	7.0000	6.9583
Median		6.5000	7.0000	7.0000
Mode		6.50	6.50 ^a	6.50
Std. Deviation		.58225	.76871	.49810
Variance		.339	.591	.248
Range		1.50	2.50	1.50
Minimum		5.50	6.00	6.50
Maximum		7.00	8.50	8.00

a. Multiple modes exist. The smallest value is shown

According to Table 9, there was a relatively slight increase in the mean scores from ($M = 6.29$, $SD = .58$) for the pretest to ($M = 7$, $SD = .76$) for the posttest; this increase was followed by a very slight decline in the mean scores for the delayed posttest ($M = 6.95$, $SD = .49$), which was still higher than the pretest mean. These results denoted that the texts elicited after the training period received better evaluation from the third judge. Such a fact was also reflected in the minimum and maximum scores for each test. The minimum score for the third judge's evaluation rose from 5.5 for Salem's, MATA's, and Moo's pretest to 6 for Salem's and MATA's posttest and then to 6.5 for Salem's, MATA's, Jacob's, Shireen's, and Ali's posttest. The maximum scores as well increased from 7 for Ellie's, Eva's, and Gabriella's pretest to 8.5 for Eva's posttest. They then slightly declined to 8 for Eva's delayed posttest.

Based on the results obtained from the measures of central tendency and measures of variability of the frequency and occurrence of the target formulaic sequences in the pretest, the posttest, and the delayed posttest of the study sample, it is clear that offering the participants focused instruction on the target formulaic sequences enhanced their acquisition and internalization into the participants' linguistic repertoire. Such a fact is evident in the participants' ability to augment their writing with a greater number of various formulaic sequences after the training period. Moreover, the descriptive statistics computed on the scores for the three judges' evaluation revealed the effectiveness of the focused instruction on formulaic sequences in augmenting the participants' writing proficiency, as reflected in the better evaluation of the texts elicited after the training period.

Cronbach's Alpha

A coefficient of internal consistency was run on the raw scores for the three judges' evaluation of the elicited texts to statistically test their internal consistency reliability.

<i>Reliability Statistics</i>		
Cronbach's Alpha	Cronbach's Alpha Based on Standardized Items	N of Items
.668	.704	3

<i>Item-Total Statistics</i>					
	Scale Mean if Item Deleted	Scale Variance if Item Deleted	Corrected Item-Total Correlation	Squared Multiple Correlation	Cronbach's Alpha if Item Deleted
J1 Evaluation	13.1250	3.634	.507	.271	.562
J2 Evaluation	13.5139	1.607	.634	.403	.453
J3 Evaluation	13.1389	4.223	.487	.257	.628

Table 10 displays an acceptable level of internal consistency ($\alpha = .66, n = 3$) for the three judges' evaluation. Moreover, the highest level of internal consistency, as shown in Table 11, is achieved when the evaluation of the three judges is included.

Paired-Samples T-Tests

The initial findings of the quantitative content analysis as well as the evaluation of the three judges were statistically tested by conducting several paired-samples t-tests. Five paired-samples t-tests were computed on the above-mentioned raw scores to investigate any statistically significant differences in the frequency and occurrence of the target formulaic sequences as well as the judges' evaluation before and after the training period.

First, a paired-samples t-test was computed on the scores for the frequency of the target formulaic sequences at the three production stages to highlight any statistically significant difference in the number of formulaic sequences utilized by the participants before and after the training period.

Table 12

Paired Samples Statistics: The Frequency of Formulaic Sequences at Three Production Stages

	Mean	N	Std. Deviation	Std. Error Mean	
Pair 1	Frequency: Pretest	8.4167	12	4.29499	1.23986
	Frequency: Posttest	21.6667	12	6.81353	1.96690
Pair 2	Frequency: Pretest	8.4167	12	4.29499	1.23986
	Frequency: Delayed Posttest	20.8333	12	6.56206	1.89430
Pair 3	Frequency: Posttest	21.6667	12	6.81353	1.96690
	Frequency: Delayed Posttest	20.8333	12	6.56206	1.89430

	Paired Differences					t	df	Sig. (2-tailed)
	Mean	Std. Deviation	Std. Error Mean	95% Confidence interval of the difference				
				Lower	Upper			
Pretest-Posttest	-13.250	6.59373	1.90345	-17.43946	-9.06054	-6.961	11	.000
Pretest-Delayed Posttest	-12.416	6.66686	1.92456	-16.65258	-8.18075	-6.452	11	.000
Posttest-Delayed Posttest	.83333	4.36585	1.26031	-1.94059	3.60726	.661	11	.522

* $p < .05$

As shown in Tables 12 and 13, there was a statistically significant difference in the scores for the frequency of the target formulaic sequences in the pretest ($M = 8.41$, $SD = 4.29$) and the posttest ($M = 21.66$, $SD = 6.81$); $t(11) = -6.96$, $p < .05$, and the magnitude of the difference in the means was very large (eta squared = .81). Such a statistically significant increase in the frequency of formulaic sequences in the posttest, an improvement of 13.25, highlighted the effectiveness of the focused instruction on formulaic sequences in promoting their acquisition. Moreover, there was a statistically significant difference in the scores for the pretest ($M = 8.41$, $SD = 4.29$) and the delayed posttest ($M = 20.83$, $SD = 6.56$); $t(11) = -6.45$, $p < .05$, with an improvement of 12.41 and a large effect size (eta squared = .79), suggesting that almost all participants could successfully internalize the target formulaic sequences into their linguistic repertoire for later usage as a result of the training period. However, there was no significant difference in the scores for the posttest ($M = 21.66$, $SD = 6.81$) and the delayed posttest ($M = 20.83$, $SD = 6.56$); $t(11) = .66$, $p > .05$. In other words, the mean difference of the two variables and the direction of the t

value demonstrated that almost all the participant utilized approximately the same number of the target formulaic sequences in the posttest and the delayed posttest.

Second, a paired-samples t-test was computed on the scores for the occurrence of the target formulaic sequences in the pretest, posttest, and delayed posttest to reveal any statistically significant difference in the variation of types before and after the training period.

Table 14

Paired Samples Statistics: The Occurrence of Formulaic Sequences at Three Production Stages

		Mean	N	Std. Deviation	Std. Error Mean
Pair 1	Occurrence: Pretest	5.2500	12	2.13733	.61699
	Occurrence: Posttest	14.4167	12	2.90637	.83900
Pair 2	Occurrence: Pretest	5.2500	12	2.13733	.61699
	Occurrence: Delayed posttest	14.0000	12	2.37410	.68534
Pair 3	Occurrence: Posttest	14.4167	12	2.90637	.83900
	Occurrence: Delayed posttest	14.0000	12	2.37410	.68534

Table 15

Paired Samples Test: The Occurrence of Formulaic Sequences at Three Production Stages

	Paired Differences					t	df	Sig. (2-tailed)
	Mean	Std. Deviation	Std. Error Mean	95% Confidence interval of the difference				
				Lower	Upper			
Pretest - Posttest	-9.1666	2.97973	.86017	-11.0599	-9.05539	-10.657	11	.000
Pretest - Delayed posttest	-8.7500	2.41680	.69767	-11.0599	-7.21444	-12.542	11	.000
Posttest - Delayed posttest	.41667	2.15146	0.62107	-.95031	-1.78364	.671	11	.516

* $p < .05$, two-tailed.

As shown in Tables 14 and 15, the differences in the scores for the occurrence of the target formulaic sequences in the pretest ($M = 5.25$, $SD = 2.13$) and the posttest ($M = 14.41$, $SD = 2.90$); $t(11) = -10.65$, $p < .05$ was statistically significant with an improvement of 9.16 and a large effect size (eta squared = .91). In addition to the significant increase in the occurrence of the target formulaic sequences in the posttest, there was a statistically significant difference between the scores for the occurrence of formulaic sequences in the pretest ($M = 5.25$, $SD = 2.13$) and the delayed posttest ($M = 14$, $SD = 2.37$); $t(11) = -12.54$, $p < .05$, eta squared = .93. The mean difference of the two scores ($M = -8.75$) and the direction of the t value ($t(11) = -12.54$) revealed the participants' tendency to augment their writing with a greater number of types in the delayed posttest. This ability demonstrated that various types of the target formulaic sequences were successfully internalized into the participants' linguistic repertoire. However, according to the mean difference ($M = .41$) and the direction of the t value ($t(11) = .671$), there was no significant difference in the scores for the occurrence of the target formulaic sequences in the posttest ($M = 14.41$, $SD = 2.90$) and the delayed posttest ($M = 14$, $SD = 2.37$); $t(11) = .671$, $p > .05$. These results indicated that almost all the participants employed approximately the same number of types in the posttest and the delayed posttest.

Third, a paired-samples t-test was computed on the raw scores for the first judge's evaluation of the pretest, posttest, and delayed posttest to highlight any statistically significant difference in her evaluation of the elicited texts.

Table 16

Paired Samples Statistics: The First Judge's Evaluation

	Mean	N	Std. Deviation	Std. Error Mean
Pair 1 Judge 1: Pretest	5.9167	12	.63365	.18292
Judge 1: Posttest	7.0833	12	.59671	.17225
Pair 2 Judge 1: Pretest	5.9167	12	.63365	.18292
Judge 1: Delayed posttest	7.2917	12	.68948	.19903
Pair 3 Judge 1: Posttest	7.0833	12	.59671	.17225
Judge 1: Delayed posttest	7.2917	12	.68948	.19903

Table 17

Paired Samples Test: The First Judge's Evaluation

	Paired Differences					t	df	Sig. (2-tailed)
	Mean	Std. Deviation	Std. Error Mean	95% Confidence interval of the difference				
				Lower	Upper			
Judge1: Pretest - Judge1: Posttest	-1.16667	.86164	.24873	-1.71413	-0.61921	-4.690	11	.001
Judge1: Pretest - Judge1: Delayed posttest	-1.37500	.56909	.16428	-1.73658	-1.01342	-8.370	11	.000
Judge1: Posttest - Judge1: Delayed posttest	-.20833	.89082	.25716	-.77433	0.35767	-.810	11	.435

* $p < .05$, two-tailed.

There was a statistically significant difference, as shown in Tables 16 and 17, in the scores for the first judge's evaluation of the pretest ($M = 5.91$, $SD = .63$) and the posttest ($M = 7.08$, $SD = .59$); $t(11) = -4.69$, $p < .05$, with an improvement of 1.16 and a large effect size (eta squared = .66). This statistically significant improvement in the scores for the posttest suggests that the

participants received better evaluation after the training period. Moreover, there was a statistically significant increase in the evaluation scores from ($M = 5.91, SD = .63$) for the pretest to ($M = 7.29, SD = .68$) for the delayed posttest; $t(11) = -8.37, p < .05$, with an improvement of 1.37 and a large effect size (eta squared = .86), indicating that the evaluation of the participants' delayed posttest is significantly higher than that of the pretest. Furthermore, there was no statistically significant difference in the scores for the posttest ($M = 7.08, SD = .59$) and the delayed posttest ($M = 7.29, SD = .68$); $t(11) = -.81, p > .05$, showing that the first judge's evaluation of the participants' posttest and the delayed posttest was very similar.

Fourth, a paired-samples t-test was computed on the raw scores for the second judge's evaluation of the textual data to reveal any statistically significant increase in his evaluation.

	Mean	N	Std. Deviation	Std. Error Mean
Pair 1 Judge 2: Pretest	5.3333	12	1.15470	.33333
Judge 2: Posttest	6.9167	12	1.53495	.44310
Pair 2 Judge 2: Pretest	5.3333	12	1.15470	.33333
Judge 2: Delayed posttest	6.8750	12	1.15059	.33215
Pair 3 Judge 2: Posttest	6.9167	12	1.53495	.44310
Judge 2: Delayed posttest	6.8750	12	1.15059	.33215

	Paired Differences					t	df	Sig. (2-tailed)
	Mean	Std. Deviation	Std. Error Mean	95% Confidence interval of the difference				
				Lower	Upper			
Judge2: Pretest - Judge2: Posttest	-1.58333	1.54968	.44735	-2.56795	-0.59871	-3.539	11	.005
Judge2: Pretest - Judge2: Delayed posttest	-1.54167	0.89082	.25716	-2.10767	-0.97567	-5.995	11	.000
Judge2: Posttest - Judge2: Delayed posttest	.04167	1.19579	.34520	-.71810	0.80144	.121	11	.906

* $p < .05$, two-tailed.

Tables 18 and 19, in turn, reveal the statistically significant increase in the scores for the second judge's evaluation from ($M = 5.33$, $SD = 1.15$) for the pretest to ($M = 6.91$, $SD = 1.53$) for the posttest; $t(11) = -3.53$, $p < .05$, with an improvement of 1.58 and a large effect size (eta squared = .53). Another significant difference was detected between the scores for the pretest ($M = 5.33$, $SD = 1.15$) and those for the delayed posttest ($M = 6.87$, $SD = 1.15$); $t(11) = -5.99$, $p < .05$, a large effect size (eta squared = .76), and an improvement of 1.54. However, there was no statistically significant difference in the scores for the posttest ($M = 6.91$, $SD = 1.53$) and the delayed posttest ($M = 6.87$, $SD = 1.15$); $t(11) = .12$, $p > .05$. In this respect, the results displayed in the above-mentioned tables indicate that the texts elicited after the training period received better evaluation from the second judge, as evident in the statistically significant difference between the pretest and the posttest and the pretest and the delayed posttest. Moreover, the fact that there was not any significant difference between the posttest and the delayed posttest

demonstrates that the second judge's evaluation of the written samples elicited after the focused instruction on formulaic sequences was very similar.

Fifth, the last paired-samples t-test was computed on the third judge's evaluation of the participants' three productions.

Table 20

Paired Samples Statistics: The Third Judge's Evaluation

		Mean	N	Std. Deviation	Std. Error Mean
Pair 1	Judge 3: pretest	6.2917	12	.58225	.16808
	Judge 3: Posttest	7.0000	12	.76871	.22191
Pair 2	Judge 3: pretest	6.2917	12	.58225	.16808
	Judge 3: Delayed Posttest	6.9583	12	.49810	.14379
Pair 3	Judge 3: Posttest	7.0000	12	.76871	.22191
	Judge 3: Delayed Posttest	6.9583	12	.49810	.14379

Table 21

Paired Samples Test: Third Judge's Evaluation

	Paired Differences					t	df	Sig. (2-tailed)
	Mean	Std. Deviation	Std. Error Mean	95% Confidence interval of the difference				
				Lower	Upper			
Judge3: Pretest - Judge3: Posttest	-.70833	.33428	.09650	-.92072	-.49594	-7.340	11	.000
Judge3: Pretest - Judge3: Delayed posttest	-.66667	.44381	.12812	-.94865	-.38468	-5.204	11	.000
Judge3: Posttest - Judge3: Delayed posttest	.04167	.49810	.14379	-.27481	.35815	.290	11	.777

* $p < .05$, two-tailed.

The results of the paired-samples t-test computed on the raw scores for the third judge's evaluation are displayed in Tables 20 and 21 above. To start with, there was a statistically significant difference in the scores for the pretest ($M = 6.29$, $SD = .58$) and the posttest ($M = 7$, $SD = .76$); $t(11) = -7.34$, $p < .05$, and the magnitude of the difference in the means was very large (eta squared = .83). Another statistically significant difference was detected in the scores for the pretest ($M = 6.29$, $SD = .58$) and the delayed posttest ($M = 6.95$, $SD = .49$); $t(11) = -5.20$, $p < .05$, and the magnitude of the difference in the means was as well very large (eta squared = .71). In this respect, the differences in the means for the pretest and the posttest and the pretest and the delayed posttest demonstrated an improvement of .70 and .66, respectively. There was, however, no statistically significant difference in the scores for the posttest ($M = 7$, $SD = .76$) and the delayed posttest ($M = 6.95$, $SD = .49$); $t(11) = .29$, $p > .05$, indicating that the participants' written samples elicited after the training received almost the same evaluation by the third judge.

Based on the results obtained from the paired-samples t-tests, it is clear that the focused instruction on formulaic sequences resulted in their successful internalization into the participants' linguistic repertoire. This is inferred from the statistically significant differences in the results of the t-tests computed on the frequency and occurrence of the target formulaic sequences in the pretest and the posttest as well as the pretest and the delayed posttest. Moreover, the fact that there was no statistically significant difference in the tests' results that compared the frequency and occurrence of the target formulaic sequences in the posttest and the delayed posttest adduce further evidence in support of the effectiveness of the focused instruction of formulaic sequences: the focused instruction offered helped the automatization of the target formulaic sequences, and, hence, the participants' use of this language phenomenon after the training period was relatively stable. Moreover, the t-test results demonstrated the positive

effects of the training period in promoting the participants' writing proficiency: the scores for the three judges' evaluation of the posttest and the delayed posttest were significantly higher than those of the pretest.

Correlation Coefficients

Three Pearson product-moment correlation coefficients were computed to highlight any possible correlations between the frequency and occurrence of the target formulaic sequences and the evaluation of the three judges.

The first Pearson-product moment correlation coefficient was carried out on the raw scores for the frequency and occurrence of the target formulaic sequences in all the elicited texts and the first judge's evaluation.

		Frequency	Occurrence	J1 Evaluation
Frequency	Pearson Correlation	1	.905**	.673**
	Sig. (2-tailed)		.000	.000
	N	36	36	36
Occurrence	Pearson Correlation	.905**	1	.716**
	Sig. (2-tailed)	.000		.000
	N	36	36	36
J1 Evaluation	Pearson Correlation	.673**	.716**	1
	Sig. (2-tailed)	.000	.000	
	N	36	36	36

The Pearson product-moment correlation coefficient results revealed, as demonstrated in Table 22, a strong positive correlation between the frequency of the target formulaic sequences and the

first judge's evaluation, which was statistically significant ($r = .67, p < .05, n = 36$), suggesting that 44% of the variance in the first judge's evaluation is accounted for by the variance in the frequency of the target formulaic sequences. There was as well a statistically significant positive correlation between the occurrence of the target formulaic sequences and the first judge's evaluation, ($r = .71, p < .05, n = 36$), indicating that 50% of the higher evaluation was associated with the increase in the types used by the participants.

The second Pearson product-moment correlation coefficient was computed on the scores for the frequency and occurrence of the target formulaic sequences and the second judge's evaluation of the textual data.

Table 23				
<i>A Correlation Matrix: The Association between the Frequency and Occurrence of the Target Formulaic Sequences and the Second Judge's Evaluation</i>				
		J2 Evaluation	Frequency	Occurrence
J2 Evaluation	Pearson Correlation	1	.636**	.602**
	Sig. (2-tailed)		.000	.000
	N	36	36	36
Frequency	Pearson Correlation	.636**	1	.905**
	Sig. (2-tailed)	.000		.000
	N	36	36	36
Occurrence	Pearson Correlation	.602**	.905**	1
	Sig. (2-tailed)	.000	.000	
	N	36	36	36

* $p < .05$, two-tailed.

Based on the results presented in Table 23, there was a statistically significant positive correlation between the scores for the second judge's evaluation and the frequency of the target formulaic sequences ($r = .63, p < .05, n = 36$), indicating that 39% of the variance in the second judge's evaluation is accounted for by the increase in the frequency of the target formulaic

sequences. There was another statistically significant positive correlation between the scores for the second judge's evaluation and those for the occurrence of the target formulaic sequences ($r = .60, p < .05, n = 36$), suggesting that 36% of the increase in the occurrence of the target formulaic sequences was associated with better evaluation.

The third Pearson-product moment correlation coefficient was conducted on the scores for the frequency and occurrence of the target formulaic sequences and the evaluation of the third judge to reveal any statistically significant association between these variables.

		J3 Evaluation	Frequency	Occurrence
J2 Evaluation	Pearson Correlation	1	.255	.315
	Sig. (2-tailed)		.134	.062
	N	36	36	36
Frequency	Pearson Correlation	.255	1	.905**
	Sig. (2-tailed)	.134		.000
	N	36	36	36
Occurrence	Pearson Correlation	.315	.905**	1
	Sig. (2-tailed)	.062	.000	
	N	36	36	36

* $p < .05$, two-tailed.

An examination of the correlation coefficient results in Table 24 demonstrated that there was no statistically significant association between the frequency of the target formulaic sequences and the third judge's evaluation ($p > .05$), nor was there any statistically significant correlation between the occurrence of the target formulaic sequences and the third judge's evaluation ($p > .05$).

The results presented in the three correlation matrices suggest that while the higher scores of the first and second judge were associated, to some extent, with the increase in the frequency and occurrence of the target formulaic sequences, the better evaluation that the participants received from the third judge for their posttest and delayed posttest did not correlate with the increase in the frequency and occurrence of the target formulaic sequences.

All in all, this chapter has presented a detailed account of the data analysis findings, including quantitative content analysis, descriptive statistics, and inferential statistics. The textual data were first quantified based on quantitative content analysis and the variables derived from the data along with the evaluation of the three judges were statistically tested to reveal any significant differences that could be attributed to the training period. The results of the paired-samples t-tests indicated that the focused instruction resulted in the acquisition and automatization of the target formulaic sequences and promoted the participants' writing proficiency. Such facts were evident in the statistically significant increase in the scores for both the frequency and occurrence of the target formulaic sequences in and the three judges' evaluation of the posttest and the delayed posttest.

Furthermore, Pearson product-moment correlation coefficients results demonstrated that the increase in the frequency and occurrence of the target formulaic sequences was associated with better evaluation in the case of the first and the second judge. However, there was no statistically significant association between the evaluation of the third judge and the frequency and occurrence of the target formulaic sequences.

Chapter Six

Discussions of the Quantitative Analysis Results

The present chapter sets forth an interpretation of the results of the quantitative data analysis in order to answer the research questions that have guided this study. To this end, the discussions, which will integrate the findings with the theoretical background for this study, are divided into four sections and each is directly related to one research question: the effectiveness of the focused instruction of formulaic sequences in enhancing L2 learners' tendency to successfully implement them in their writing, the positive impact that formulaic sequences might have on L2 learners' writing proficiency, the possible association between the increase in the frequency and occurrence of formulaic sequences and L2 learners' better achievements, and L2 learners' obvious inclination to utilize formulaic sequences as 'zones of safety' as well as links to construct proficient, coherent, and cohesive pieces of writing. Examples from the data as needed to provide further clarification are integrated into the discussions.

Focused Instruction of Formulaic Sequences

The descriptive statistics and the paired-samples t-tests computed on the raw scores for the frequency and occurrence of formulaic sequences demonstrated the effectiveness of the focused instruction of formulaic sequences in motivating the participants to increasingly employ this language phenomenon in their writing. More importantly, the focused instruction resulted in a growing tendency on the part of the participants to augment their writing with different types of formulaic sequences rather than repeating the same ones. Such findings resonate with the calls of several experts in the field of second language acquisition for the importance of explicitly teaching formulaic sequences to L2 learners. For example, Boers and Lindstromberg (2009), Cortes (2004), Granger (1998), Hyland (2006), Jones and Haywood (2004), and Paquot (2008)

point out that focused instruction, which provides L2 learners with the opportunity to rehearse and, thus, internalize formulaic sequences into their linguistic repertoire, is needed to promote the appropriate use of this language phenomenon in L2 learners' writing.

As for the frequency of the target formulaic sequences, the results of the above-mentioned quantitative tests show that the focused instruction provided during the training period resulted in a statistically significant increase in the frequency of the target formulaic sequences in the posttest as compared to the pretest. Moreover, the ability of all the participants to employ almost the same number of the target formulaic sequences in both the posttest and the delayed posttest, which took place fifteen to twenty days after the training, suggests that the focused instruction has led to the successful acquisition and internalization of the target formulaic sequences into the participants' linguistic repertoire. This is reflected in the statistically significant difference between the pretest and the delayed posttest and the absence of any statistically significant difference between the posttest and the delayed posttest. These results can be clearly linked to Cortes's (2004) hypothesis that L2 learners' limited use of formulaic sequences in their writing can be ascribed to the fact that they were never explicitly taught how to utilize and implement this language phenomenon when approaching a writing task.

In addition to the increase in the frequency of the target formulaic sequences, the focused instruction fostered the participants' ability to augment their written products with different types of formulaic sequences. Thus, instead of repeating the same formulaic sequences, all the participants could successfully implement more types in the posttest and the delayed posttest as compared to the pretest. The participants' manipulation of a wider range of types is manifested in the statistically significant results of the paired-samples t-tests computed to compare the pretest and the posttest as well as the pretest and the delayed posttest. Moreover, the fact that there was

no significant difference in the occurrence of formulaic sequences in the posttest and the delayed posttest demonstrates that various types were successfully internalized into the participants' linguistic repertoire as a result of the focused instruction offered.

The participants' growing tendency to utilize formulaic sequences in their academic writing can be further clarified with reference to some examples from the participants' written production. The most obvious variation was in the formulaic sequences used to describe the trends presented in the prompt for data collection. In the pretest, almost all the participants used two or three lexical words (*increase*, *decrease*, and *change*) to describe the price differences in the prompt; however, after the training period they utilized different types of formulaic sequences to report the changes in the trends more accurately. The texts elicited from MATA display the highest variation; that is, the formulaic sequences utilized by MATA in the posttest and the delayed posttest included *increased dramatically*, *reached a peak*, *slight increase*, *remained the same*, *rise and fall*, *the fluctuation of the prices*, *increased gradually*, and *slight increase*. Such variation was not detected in the pretest, in which *changes*, *differences*, and *reached a peak* were repeatedly deployed to present the trends in the prompt. Salem's writing, in turn, is another typical example of this variation. After employing *increase*, *decrease*, and *change* in his pretest, Salem supplemented his posttest and delayed posttest with the target formulaic sequences that are typical of such a task. Among the formulaic sequences used in his posttest and delayed posttest were *the rise and fall*, *remain the same*, *slightly decreased*, and *dramatic increase*.

In addition to price comparisons, the participants augmented their texts with different types of formulaic sequences to introduce periods of time. Jacob, for example, utilized three different formulaic sequences in the posttest and the delayed posttest to introduce a period of time,

namely, *from (A) to (B)*, *between (A) and (B)*, and *over a period of*, after repeatedly using one formulaic sequence to perform the same function in the pretest. In other words, *from (A) to (B)* was implemented four times in Jacob's pretest. In addition to Jacob, Moo could successfully implement the three above-mentioned formulaic sequences in the posttest and the delayed posttest after refraining from introducing almost all the periods of time associated with the differences in the trends in his pretest. Here, it is worth mentioning that Moo, in his pretest, used *from (A) to (B)* once to present a price range, and he only referred to one date, using *in Month, year* structure. The participants' tendency to augment their writing with various types of formulaic sequences to introduce periods of time has been detected in almost all the texts elicited after the training period (see Appendix E). Another example of this variation was found in Shireen's posttest and delayed posttest, in which three different exemplifiers were used interchangeably after the training period to list examples, i.e., *such as*, *for example*, and *for instance* after using only two exemplifiers in the pretest.

Furthermore, the manual coding of the types used in the elicited texts has revealed crystal clear similarities in the use of the target formulaic sequences. Some formulaic sequences were implemented by almost all the participants in both the posttest and the delayed posttest. They include, but are not limited to, *the graph shows*, *the rise and fall*, *dramatic increase*, *remain the same*, *in conclusion*, *both (A) and (B)*, and *the (A) of the*. The obvious similarities in the participants' production after the training period can be interpreted with reference to the significant role of the focused instruction of formulaic sequences in familiarizing L2 learners with the use and functions of this language phenomenon (Boers & Lindstromberg, 2009; Jones & Haywood, 2004; Lewis, 2000b). It can be suggested that the focused instruction, along with the intensive practice provided to the participants during the training period, did not only promote

the acquisition and internalization of the target formulaic sequences into the participants' linguistic repertoire, but it also helped the participants fully grasp the use and function of each formulaic sequence, and, thus, they relied on the ones that best helped them express the text meaning rather than randomly implementing the formulaic sequences which they have already learned and practiced. In this respect, one might argue that the focused instruction promoted the participants' ability to deploy the target formulaic sequences accurately and appropriately in their writing, an action that is considered essential for proficient writing skills (Handl, 2008). As Boers and Lindstromberg (2009), Cortes (2004), Ellis et al. (2008), Granger (1998), O'Keeffe (2012), and Wood (2009a) put it, focused instruction is needed to elevate the saliency of formulaic sequences, raise L2 learners' awareness of their different discourse functions, promote L2 learners' ability to accurately implement them in their language production, and augment L2 learners' implicit knowledge of the target language.

Last but not least, the textual data elicited from the participants have adduced further evidence in support of the effectiveness of practice as far as formulaic sequences are concerned. During the training period, seven formulaic sequences, namely, *there is*, *there are*, *at the beginning of*, *at the end of*, *on the basis of*, *based on the*, and *as shown in*, were not explicitly taught to the participants, but they were only referred to orally when discussing the ninth worksheet, a table that includes all the target formulaic sequences. Although they are alternatives to other formulaic sequences that were repeatedly used in the posttest and the delayed posttest, they were scarcely utilized in the texts collected after the training period, when compared to their alternatives. The formulaic sequences were *there is* in Abedeer's delayed posttest, *there are* in both Ali's posttest and David's delayed posttest, *based on the* in Jacob's posttest, *as shown in* in MATA's posttest, and *at the end of* in both Ellie's posttest and Gabriella's delayed posttest.

Other formulaic sequences, namely, *according to the* and *over a period of*, like the former ones, were not located in the pretest of any of the participants, except Mata's, and were only presented orally to the participants. However, unlike the above-listed formulaic sequences, they were practiced in cloze and dictogloss activities and were located in almost all the texts elicited after the training period.

At this point, it can be suggested that the practice stage might have promoted the acquisition of the formulaic sequences that were made salient at the presentation stage. Such an inclination to use the formulaic sequences that were previously practiced can be interpreted with reference to Schmitt's (2000) emphasis on the importance of providing L2 learners with a robust support, through activities which elevate deep processing, to facilitate the acquisition and automatization of this language phenomenon. It is worth mentioning here that this particular interpretation is not meant to generalize that raising L2 learners' awareness does not result in successful acquisition but to indicate that making discourse elements salient, without augmenting them with appropriate practice, can, in several cases, result in upgrading receptive knowledge rather than both receptive and productive knowledge (Schmitt, 2000).

Focused Instruction of Formulaic Sequences and L2 Writing Proficiency

The focused instruction on the target formulaic sequences provided the twelve participants with the opportunity to develop their writing proficiency. In other words, most of the participants could produce better quality reports after the training period. This training period, in which the participants practiced writing introductory and concluding sentences, raised their awareness of the importance of these two essential components of academic writing. Abedeer, for example, did not write an introductory sentence to her pretest, as illustrated in the example given below which was the first sentence she started her paragraph with:

In this graph we can notice that one of the product like sugar keep growing during the year and the consumers still using it or in need to it on the other hand we can see that the dairy product stayed at the same range of price and the number of the consumers decrease.

(Appendix E, p. 239)

However, Abedeer wrote introductory sentences very similar to the models she practiced during the training period for both her posttest and delayed posttest:

This graph shows the rise and fall in prices of food between May 2010 and April 2011 and the behaviour issue of the buyer. (Appendix E, P. 240)

The graph shows the rise and fall in food prices for the period of May 2010 to April 2011.

(Appendix E, p. 241)

Shireen, in turn, did not write a concluding sentence for her pretest, but she concluded both the posttest and the delayed posttest with a summary of the information presented in the text proceeded by a formulaic sequence to mark the beginning of the conclusion:

To sum up, the increase in food prices has a direct effect on our health and consumption habits. (Appendix E, p. 236)

To conclude, the fluctuation in the food prices has a direct effect on the consumption habit for both families and individuals. (Appendix E, p. 237)

In other cases, having the participants practice model introductory and concluding sentences resulted in overall increased linguistic accuracy. Salem, for instance, could write better quality introductory and concluding sentences after the training period. The examples provided below present Salem's production at the pretest, the posttest, and the delayed posttest, respectively. Since the main focus here is on the introductory and concluding sentences, the body of each paragraph is replaced by three dots:

It show how are some materials change throw around one year. ... In fact, when the price became higher thats give pad effect. For example. The poor people will be increase, and many people will do criminal thing like theft. (Appendix E, p. 193)

The line graph presents the rise and fall in the general price of commodity between May 2010 and April 2011. ... In conclusion, it can be seen that the commodity prices have generally increased wich leads to a decline in the purchasing power of the concumers.

(Appendix E, p. 193)

The graph shows how the commodity prices change over a one year. ... In conclusion, it is clear from the information of graph that price of commodity have generally increased between May 2010 to April 2011. (Appendix E, p. 194)

The examples illustrated above are in line with Lewis's (1997) and Nattinger and DeCarrico's (1992) beliefs in the effectiveness of having L2 learners practice the use of some formulaic sequences that function as sentence builders or frames. To them, L2 learners can use these frames to produce better quality sentences that are constructed through a chunking process rather than being composed from scratch.

Moreover, the lexical errors relating to a lack of formulaic sequences knowledge were corrected in the posttest and the delayed posttest. For instance, in the pretest, Ali made the same error four times when he implemented a formulaic sequence to introduce the price range of several products, “*between 150 to 250” (Appendix E, p. 212). However, that error was corrected in Ali's posttest. Similarly, Mary used a formulaic sequence inaccurately in the pretest to introduce a period of time, “*during May 2010 to April 2010” (Appendix E, p. 204). However, this formulaic sequence, which was twice constructed inaccurately in Mary's pretest, was employed correctly in her posttest and delayed posttest. The participants' errors in these two

particular formulaic sequences echo with Boers and Lindstromberg's (2009) claims that although some formulaic sequences occur frequently in input, they pass unnoticed by many L2 learners. According to the Corpus of Contemporary American English, *Between (A) and (B)* and *From (A) to (B)* are highly frequent in all types of discourse, especially the academic one (451.88 and 300.33 per one million words, respectively). Ali's and Mary's repeated errors in these two formulaic sequences are typical examples of the failure of some L2 learners to perceive the structure of some formulaic sequences despite their high frequency in input. This failure suggests that they did not have full grasp of their correct structure by exposure alone. Moreover, the fact that they could use the target formulaic sequences accurately in both the posttest and the delayed posttest indicates that introducing L2 learners to formulaic sequences can indeed promote their lexical accuracy (Boers & Lindstromberg, 2009; Coxhead & Byrd, 2007; Schmitt, 2000).

Formulaic Sequences and L2 Learners' Better Achievement

As mentioned earlier, the participants' pretest, posttest, and, delayed posttest were evaluated by three judges who are experienced in teaching English for academic purposes. The raw scores for the three judges' evaluation indicated that almost all the texts that were produced after the training period received better evaluation from the three judges, and the results of the three paired-samples t-tests demonstrated that the increase in each rater's scores for the posttest and the delayed posttest was statistically significant when compared to the pretest. More importantly, the absence of any statistically significant difference in the scores for the posttest and the delayed posttest suggests that each judge's evaluation of the posttest and the delayed posttest was very similar.

The paired-samples t-test results, which demonstrated that the participants could achieve higher scores after the training period, motivated carrying out correlation coefficients to gain in-

depth understanding of the association between the frequency and occurrence of the target formulaic sequences and the judges' evaluation. On the one hand, the scores for the evaluation of the first and the second judge positively correlated with the frequency and occurrence of the target formulaic sequences, i.e., the increase in the number as well as types of the target formulaic sequences was associated with better evaluation. As previously discussed in Chapter Two, implementing formulaic sequences that are typical of a particular genre might be positively associated with L2 learners' better achievements in academic contexts (Boers & Lindstromberg, 2009; Coxhead & Byrd, 2007; Ellis et al., 2008; Jones & Haywood, 2004; Lewis, 1997). To Li and Schmitt (2009), "to be a successful academic writer, an L2 learner is required to be competent at using these conventional sequences which characterize the learner's discipline" (p. 86).

On the other hand, there was no statistically significant correlation between the evaluation of the third judge and the frequency and occurrence of the target formulaic sequences. In other words, despite the statistically significant increase in the scores for the third judge's evaluation of both the posttest and the delayed posttest, this increase was not associated with the increase in the frequency and occurrence of the target formulaic sequences. These results can be viewed as an indication that the use of the target formulaic sequences has implicitly enhanced the overall quality of the participants' writing and helped them perform at a level acceptable in the academic context without being directly associated with better achievements. As Cortes (2004), Granger (1998), and Hyland (2008) explicate, the successful internalization of formulaic sequences might augment L2 learners' implicit knowledge of the target language and help define them as proficient writers.

The Use of Formulaic Sequences as Zones of Safety as well as Links

It has been suggested that formulaic sequences can serve as ‘zones of safety,’ ‘islands of reliability,’ or ‘frames’ on which L2 users tend to depend, in writing or in speech, to mitigate their linguistic errors and communicate their ideas successfully (Boers et al., 2006; Boers & Lindstromberg, 2009; Hill, 2000; Lewis, 2000b; Nattinger & DeCarrico, 1992; Richards & Rodgers, 2001; Wood, 2009a, 2009b). The data for this study exhibited the participants’ strong inclination to resort to formulaic sequences when composing their sentences. At this point, some of the sentences generated after the training period were constructed through, what Lewis (1997) calls, a chunking process, i.e., composing a whole sentence by joining formulaic sequences together. It should be noted here that these sentences were detected in several texts regardless of the participants’ proficiency level. The two examples extracted from Salem’s posttest and delayed posttest can best describe this tendency:

The line graph presents the rise and fall in the general price of commodity between May 2010 and April 2011. (Appendix E, p. 193)

In addition, Both oils and cereals increased gradually. (Appendix E, p. 194)

In addition to Salem, Ellie’s topic sentences for both the posttest and the delayed posttest were entirely formulaic:

The graph shows fluctuation in food prices between May 2010 and April 2011.

(Appendix E, p. 196)

The graph shows the fluctuation in commodity prices from May 2011 to April 2011.

(Appendix E, p. 198)

Ali also composed several sentences in both the posttest and the delayed posttest by using formulaic sequences to describe the trends presented in the chart:

On the other hand, the price of cereals and oils fluctuated between 150 and 250 over this

period of time. It can be seen that the price of both dairy and meat remain stable over these twelve months. (Appendix E, p. 213).

In addition, the price of meat increased slightly over these months. (Appendix E, p. 215).

The above-listed examples, in which all the underlined strings of words represent target formulaic sequences, reveal the participants' inclination to rely heavily on the target formulaic sequences in their posttest and delayed posttest. Other examples from the data, although not entirely formulaic, disclose the same tendency on the part of almost all the participants. The underlined words in Mary's posttest and delayed posttest are all formulaic:

according to the graph the price of meat has slightly increased during the studied period.

When it comes about cereals and oils, it can be seen that both have experienced a gradual increase from May to December 2010, after that the prices of cereals and oil have slightly decreased. (Appendix E, p. 205)

The strong impact was the double increase of sugar prices from May 2010 to December 2010. In conclusion, the graph represents the prices' fluctuation of sugar, cereals, oils, dairies, and meat during one year. (Appendix E, p. 206)

Another prominent example of this tendency is that of MATA's posttest and delayed posttest:

Similarly, the prices of cereals, oils and dairy had dramatic increase from May 2010 to the end of the same year. On the other hand, meat's prices remained the same in the first seven monthes with just a slight increase from Dec 2010 to April 2011. (Appendix E, p. 218).

In conclusion, as it is clear from the data given, the prices of commodity had rise and fall over a period of one year. It has been believed that changes in the food prices have negative and positive effects on the consumption of any product. (Appendix E, p. 220)

Based on the examples provided above, it seems that the participants viewed formulaic sequences as what Boers et al. (2006) call ‘zones of safety’ that can be utilized to compose sentences, while, at the same time, mitigate the number of linguistic errors they might make when writing in the target language.

In addition, some of the examples in the posttest and the delayed posttest revealed the participants’ tendency to use, in some occasions, the target formulaic sequences creatively in order to express the text meaning. In other words, some participants amalgamated two formulaic sequences to form a noun phrase. For instance, Ali, as shown in the example above, combined a tangible framing attribute (*the (A) of*) with a sentence builder (*Both (A) and (B)*) to construct a noun phrase: *the price of both dairy and meat*. Eva, in turn, merged a collocation (*the rise and fall*) with a tangible framing attribute (*the (A) of*) to construct a noun phrase: *the rise and fall of food price*.

Other participants changed the structure of the target formulaic sequences in different ways. David, in the delayed posttest, used a sentence head creatively by inserting an adverb into it to express an impersonal evaluation, i.e., *it is quite obvious*. Similarly, Eva deployed a sentence stem creatively in the delayed posttest (*it is clearly seen*). Moreover, almost all the participants used the target formulaic sequences creatively in order to describe the trends presented in the chart by either merging the constituents of two formulaic sequences to form a new one or converting their part of speech from nouns to verbs and vice versa. Among the formulaic sequences that were constructed creatively are *increased gradually, decreased gradually, increased slightly, remained in the same price, sharply increased, dramatic fall, fluctuate between (A) and (B), decreased sharply, stayed stable, and rose rapidly*.

All the above listed examples reveal the participants' endeavors to rely on the target formulaic sequences in their construction of the text meaning rather than composing their sentences out of individual lexical items. Thus, it can be noted here that the participants might have viewed the target formulaic sequences as trustworthy frames which they could use creatively to lessen the possibilities of breaching the linguistic rules of the target language. This can be seen in Nattinger and DeCarrico's (1992) and Schmitt's (2000) notions about the acquisition of formulaic sequences. After the successful acquisition of a considerable number of formulaic sequences, L2 learners might use formulaic sequences creatively as a backbone for a well-structured and accurate academic piece of writing. The participants' tendency to creatively utilize the target formulaic sequences can also be interpreted with reference to the syntactic and semantic flexibility of formulaic sequences which allows changing the order or their constituents and inserting semantically related words without breaching the syntactic rules of language (Erman & Warren, 2000).

Finally, the quantitative content analysis pinpointed the participants' increased usage of formulaic sequences which function as discourse organizers that connect the constituents of their texts together after the training period. David, Mary, Ali, MATA, Jacob, and Abedeer, for instance, implemented very few discourse organizers, if any, in their pretest. However, after the training period, they used such formulaic sequences as *according to the*, *in addition to*, *at the same time*, *on the one hand*, and *on the other hand* as links to help them establish a smooth transition between the text elements and, thus, construct coherent and cohesive texts. In this respect, raising L2 learners' awareness of the important role of these formulaic sequences in the composition of a good academic piece of writing appears to promote their tendency to use them (Biber et al., 2004; Hyland, 2008; Nattinger & DeCarrico, 1992; Simpson-Vlach & Ellis, 2010).

To sum up, the findings of this study suggested that the explicit instruction resulted in the acquisition and internalization of the target formulaic sequences into the participants' linguistic repertoire as evident in the significant increase in the number of formulaic sequences in both the posttest and delayed posttest regardless of the participants' proficiency levels. Moreover, the participants did not merely memorize the target formulaic sequences and use them as wholes in their writing, but they rather had the tendency to manipulate the ones which best suited their needs to mitigate their linguistic errors while constructing the text meaning. The extensive use of the target formulaic sequences augmented L2 learners' writing skills, which can be inferred from the statistically significant increase in the three judges' evaluation of the posttest and the delayed posttest.

Chapter Seven

Conclusion

The present chapter is devoted to briefly restate the findings of this research study, suggest some implications for both curriculum design and teaching methodology, highlight the limitations of the study, and present some directions for future research.

Quantitative Results

This exploratory study investigated the effects of focused instruction of formulaic sequences on L2 learners' academic writing skills. The study has shown that explicit instruction of formulaic sequences can lead to an increase in the implementation of different types of formulaic sequences, which might, in turn, augment L2 learners' writing proficiency. It has also demonstrated that focused instruction of formulaic sequences may help L2 learners construct better quality sentences, reduce the number of their lexical errors, utilize the target formulaic sequences holistically and creatively to accomplish a writing task with fewer mistakes, and produce well-connected pieces of writing by using formulaic sequences as links that bind their texts together. Furthermore, the study has revealed the possible association between the increase in the frequency and occurrence of the target formulaic sequences and the better evaluation that L2 learners might achieve in academic contexts.

Formulaic Sequences and Language Teaching

This quantitative research study highlighted the relationship between formulaic sequences and L2 learners' academic writing skills, and, hence, it suggests some pedagogical implications for both curriculum design and teaching methodology in academic contexts. It has been generally believed that writing is the most difficult skill for many L2 learners to master and become proficient in (Cook & Bassetti, 2005; Ferris, 2002; Leki, 2006; Richards & Renandya, 2002;

Seow, 2002). Such a difficulty has been predominantly ascribed to the complex nature of writing, which requires the mastery of several aspects of language, including lexicon and structure, as well as the awareness of the expressions typical of the discipline in question (Biber, 2006; Celce-Murcia & Olshtain, 2000; Corson, 1995; Coxhead & Byrd, 2007; Ellis et al., 2008; Hinkel, 2002; Li & Schmitt, 2009; Martinez & Schmitt, 2012). As Richards and Renandya (2002) elucidate,

The skills involved in writing are highly complex. L2 writers have to pay attention to higher level skills of planning and organizing as well as lower level skills of spelling, punctuation, word choice and so on. The difficulty becomes even more pronounced if their language proficiency is weak. (p. 303)

According to the findings of this study, formulaic sequences can be seen as the skeleton of L2 learners' proficient academic writing. In this respect, it can be suggested that, when designing curricula for English for Academic Purposes programs, curriculum designers should integrate formulaic sequences that are peculiar to different disciplines and arrange them thematically. In so doing, L2 learners' might gain in-depth knowledge of the use and functions of formulaic sequences in particular fields. After all, proficient writers are expected to both abide by the linguistic rules of language and comply with the intended readers' expectations by implementing the expressive potential formulaic sequences of the discourse in question (Hyland, 2008). Moreover, it is recommended that curriculum designers enhance the target formulaic sequences with writing activities that promote noticing and deep processing, two essential processes for the successful acquisition and automatization of this language phenomenon (Boers & Lindstromberg, 2009; Cortes, 2004; Coxhead & Byrd, 2007; Nation, 2002). As suggested by

Schmitt (2000) and demonstrated in the findings of this study, noticing alone does not always result in developing both the receptive and productive knowledge of L2 learners.

In addition to curriculum design, language instructors, being the medium and the trustworthy resource of information for almost all L2 learners, have the strongest contribution to the development of L2 learners' academic writing skills. As illustrated in Chapter Two, most of the academic evaluation, if not all, is based on the students' ability to generate and organize their ideas in academic texts in order to demonstrate their in-depth understanding of their academic disciplines (Connor, 2003; Leki, 2006; Llach, 2011; Nation, 1990; Paltridge, 2004; Zhu, 2006). However, elevating L2 learners' expressive skills from everyday informal language to the specialized academic language is not always achieved through practice alone. To Reppen (2002), "simply allowing students to write a lot will not necessarily provide sufficient practice in the types of writing valued for academic learning" (p. 321). As a solution to this problem, Reppen (2002) suggests that writing teachers provide L2 learners with "the language necessary to talk about texts ... to make a piece of writing more effective and appropriate to the communicative purpose" (p. 322).

In fact, Reppen's (2002) arguments go hand in hand with the previously-mentioned scholars' intensive accounts of the effectiveness of formulaic sequences in augmenting L2 learners' academic writing skills. Hence, formulaic sequences can be viewed as the springboard for L2 learners' proficient writing. Writing teachers can make these language items salient for L2 learners by implementing some activities that are replete with different types of formulaic sequences. They should also foster their students' expressive skills by teaching them how to use some sentence frames creatively to account for their needs. By doing so, L2 learners will expand their linguistic repertoires with academic expressions that can upgrade their academic writing

skills. At the same time, not only will L2 learners use formulaic sequences as wholes, but they will also, as suggested by Boers et al. (2006), Boers and Lindstromberg (2009), and Nattinger and DeCarrico (1992) and confirmed in the results of this study, utilize them as ‘zones of safety’ and links to construct more accurate and coherent pieces of writing.

Furthermore, since English for Academic Purposes (EAP) programs are designed to build on L2 learners’ everyday communicative skills by providing them with the language necessary for their successful operation at the academic level, language instructors may be able to upgrade L2 learners’ productive skills by introducing different formulaic sequences that serve various functions and by raising L2 learners’ awareness of the importance of this language phenomenon in academic contexts. Such an approach to teaching English for academic purposes can be highly beneficial for L2 learners’ performance because mastering the expressions that are typical of different academic fields is an obstacle they need to overcome in order to perform at an acceptable level in academic contexts (Biber, 2006; Corson, 1995; Coxhead & Byrd, 2007; Ellis et al., 2008; Hinkel, 2002; Li & Schmitt, 2009; Martinez & Schmitt, 2012). Thus, such an approach to language teaching will upgrade L2 learners’ academic skills.

Limitations

This quantitative research study has five obvious limitations. The first limitation relates to the heterogeneous study sample that does not reach the specified threshold for any generalization. Dornyi (2007) suggests that generalizing the findings of a research study mainly relates to the sample size and the sample’s representativeness of the whole population under investigation. However, this exploratory study, which was conducted on twelve participants, falls below the threshold for generalizability, a minimum of thirty participants for correlational studies and fifteen participants for comparative studies (Dornyi, 2007). Thus, despite the significant

findings of this research study, its results cannot be generalized because the heterogeneous nature of the study sample and the number of the participants make the study sample far from representing the targeted population.

The second limitation relates to the nature of correlational studies which cannot identify cause and effect. Accordingly, it is not clear whether the main reason behind the better evaluation that the participants received for their posttest and delayed posttest was the increase in the frequency and occurrence of the target formulaic sequences, or there were other factors that contributed to better evaluation.

The third limitation is related to the reasons behind the increased structural and grammatical accuracy in some of the texts elicited after the training period, which might have, to some extent, affected the three judges' evaluation. It is worth noting here that the increased structural and grammatical accuracy was not accounted for in the sixth chapter for two reasons. The first reason pertains to the obvious inconsistency in the participants' increased accuracy. For example, they self-corrected some grammatical and structural mistakes in two or three sentences, but they, again, made identical mistakes in other sentences. Among the prominent examples of such a tendency were the comma splice after a prepositional phrase and dropping the third person -S. Second, due to the strong instinct of a language teacher, the researcher could not totally control her tendency to provide the participants with some explicit oral error correction, which was sometimes accompanied by metalinguistic feedback, when they made major errors. In this respect, the fewer mistakes which some participants made in the posttest and the delayed posttest might have been the result of the explicit error correction offered during the training period rather than the acquisition of the target formulaic sequences, and, hence, the oral error correction

can be seen as an uncontrolled factor in this study that might have, to some extent, interfered with the participants' performance and overall achievements.

Fourth, there was no rubric for the three judges to measure and evaluate the participants' performance objectively. Nor were they provided with any training prior to their evaluation of the elicited texts. Such a fact has resulted in moderate internal consistency among the raters and the lack of adequate information about the basis on which each judge has evaluated the elicited texts.

Finally, there were not any follow up interviews with the participants or the three judges to reflect on the whole process. While conducting interviews with the participants would have provided in-depth understanding of the reasons behind their extensive use of the target formulaic sequences in both the posttest and the delayed posttest, follow up interviews with the judges would have augmented this study with several objective viewpoints on the participants' overall achievements and provided insights into the reasons behind their evaluation.

Conclusion and Suggestions for Future Research

This quantitative research study is novel in that it has explored the role of the focused instruction of formulaic sequences in upgrading L2 learners' writing skills in paragraph writing. It has as well highlighted the way in which L2 learners might utilize formulaic sequences holistically and creatively in order to compose better quality texts that approach academic discourse. Moreover, it has empirically expanded the view of formulaic sequences, which were mainly tackled as a cornerstone in enhancing speech fluency, to encompass both productive skills, i.e., speaking and writing. Furthermore, the study has demonstrated empirically that focused instruction of formulaic sequences may result in L2 learners' better achievements, and it has provided insights into curriculum design and teaching methodology in academic contexts.

By highlighting the effectiveness of focused instruction of formulaic sequences in augmenting L2 learners' academic writing skills, this thesis ushers the way for future research to scrutinize such a role empirically on a larger sample to come up with generalizable results that can provide insights into teaching academic writing in various contexts. Moreover, the results of this study shed light on one of the major areas that should be the focus of future research.

Since focused instruction of formulaic sequences appear to lead to an increase in the holistic and creative use of formulaic sequences in a specific genre, more empirical studies are needed to examine any possible tendency on the part of L2 learners to transfer acquired formulaic sequences from one genre to another. If such a tendency is empirically proven, it will contribute to the development of genuine classroom practices in EAP programs by providing L2 learners with frames onto which they cling their ideas in different genres.

As for the relationship between formulaic sequences and increased linguistic accuracy, this particular notion should be treated with caution because, although very recurrent in research, it has not yet been empirically proven. At this point, empirical research might clarify whether formulaic sequences alone upgrade L2 learners' linguistic accuracy, or they need to be accompanied by corrective feedback. Thus, conducting empirical, controlled studies in which different groups receive focused instruction on formulaic sequences with one controlled factor, corrective feedback, can reveal whether L2 learners' increased accuracy, if any, is the result of mastering formulaic sequences which have semantic and syntactic functions in discourse (Pawley & Syder, 1983), or is due to a particular type of corrective feedback.

In a nutshell, the importance of this quantitative research study can be mainly attributed to its overarching theme, i.e. upgrading L2 learners' academic writing skills. The study has revealed the effectiveness of the focused instruction of formulaic sequences in augmenting L2 learners

writing skills. It has shown that providing L2 learners with the practice necessary for the acquisition and automatization of formulaic sequences typical of a genre in controlled activities may promote their use in paragraph writing and will, hence, upgrade L2 learners' academic writing skills.

References

- Arnaud, P. J. L., & Ferragne, E. Lewis, D. M., & Maniez, F. (2008). Adjective + Noun sequences in attributive or NP-final position: Observations on lexicalization. In S. Granger & F. Meunier (Eds.), *Phraseology: An interdisciplinary perspective* (pp. 111-125). Amsterdam: John Benjamins.
- Arnaud, P. J. L., & Savignon, S. J. (1997). Rare word, complex lexical units and the advanced learner. In J. Coady & T. Huckin (Eds.), *Second language vocabulary acquisition: A rationale for pedagogy* (pp. 157-173). Cambridge: Cambridge University Press.
- Bacha, N. N. (2002). Developing learners' academic writing skills in higher education: A study for educational reform. *Language and Education*, 16(3), 161-177.
- Bachman, L. F. (2004). *Statistical analyses for language assessment*. Cambridge: Cambridge University Press.
- Barlow, M. (2000). Usage, blends and grammar. In M. Barlow & S. Kemmer (Eds.), *Usage based models of language* (pp. 315-346). Stanford: CSLI Publications.
- Beckner, C., Blythe, R., Bybee, J., Christiansen, M. H., Croft, W., Ellis, N. C., Holland, J., Ke, J., Larsen-Freeman, D., & Schoenemann, T. (2009). Language is a complex adaptive system: Position paper. *Language learning*, 59(1), 1-26.
- Biber, D. (2006). *University language: A corpus-based study of spoken and written registers*. Amsterdam: John Benjamins.
- Biber, D., & Conrad, S. (1999). Lexical bundles in conversation and academic prose. In H. Hasselgard & S. Oskserjell (Eds.), *Out of Corpora: Studies in honour of Stig Johansson* (pp. 181-190). Amsterdam: Rodopi.
- Biber, D., Conrad, S., & Cortes, V. (2004). If you look at...: Lexical bundles in university teaching and textbooks. *Applied Linguistics*, 25(3), 371-405.

- Biber, D., Conrad, S., & Reppen, R. (1998). *Corpus linguistics: Investigating language structure and use*. Cambridge: Cambridge University Press.
- Boers, F., Eyckmans, J., Kappel, J., Stengers, H., & Demecheleer, M. (2006). Formulaic sequences and perceived oral proficiency: Putting a lexical approach to the test. *Language Teaching Research*, 10(3), 245-261.
- Boers, F., & Lindstromberg, S. (2009). *Optimizing a lexical approach to instructed second language acquisition*. Basingstoke: Palgrave Macmillan.
- Celce-Murcia, M., & Olshtain, E. (2000). *Discourse and context in language teaching: A guide for language teachers*. Cambridge: Cambridge University Press.
- Chen, X. (2009, December). *The role of formulaic sequence in speech fluency development in ESL*. Paper presented at the International Conference on Computational Intelligence and Software Engineering, Wuhan.
- Chowdhury, S. A., & Tarannum, S. (2010). Speaking skills: Teaching and learning. *Glottodidactica: Biannual Journal of Applied Linguistics*, 1(2), 27-34.
- Coady, J. (1997). L2 vocabulary acquisition through extensive reading. In J. Coady & T. Huckin (Eds.), *Second language vocabulary acquisition: A rationale for pedagogy* (pp. 225-237). Cambridge: Cambridge University Press.
- Conklin, K., & Schmitt, N. (2008). Formulaic sequences: Are they processed more quickly than nonformulaic language by native and nonnative speakers? *Applied Linguistics*, 29(1), 72-89.
- Conklin, K., & Schmitt, N. (2012). The processing of formulaic language. *Annual Review of Applied Linguistics*, 32, 45-61.

- Connor, U. (2003). Changing currents in contrastive rhetoric: Implications for teaching and research. In B. Kroll (Ed.), *Exploring the dynamics of second language writing* (pp. 218-241). Cambridge: Cambridge University Press.
- Conzett, J. (2000). Integrating collocation into a reading and writing course. In M. Lewis (Ed.), *Teaching collocations: Further developments in the lexical approach* (pp. 70-87). Boston, MI: Heinle.
- Cook, V., & Bassetti, B. (2005). An introduction to researching second language writing systems. In V. Cook & B. Bessetti (Eds.), *Second language writing systems* (pp. 1-70). Clevedon: Multilingual Matters.
- Corson, D. J. (1995). *Using English words*. Dordrecht: Kluwer Academic Publishers.
- Cortes, V. (2004). Lexical bundles in published and student disciplinary writing: Examples from history and biology. *English for Specific Purposes*, 23(4), 397-423.
- Coxhead, A. (1998). *An academic word list*. English language institute occasional publication No 18. New Zealand: Victoria University of Wellington.
- Coxhead, A. (2000). A new academic word list. *TESOL Quarterly*, 34(2), 213-238.
- Coxhead, A. (2008). Phraseology and English for academic purposes: Challenges and opportunities. In F. Meunier & S. Granger (Eds.), *Phraseology in foreign language learning and teaching* (pp. 149-161). Amsterdam: John Benjamins.
- Coxhead, A., & Byrd, P. (2007). Preparing writing teachers to teach the vocabulary and grammar of academic prose. *Journal of Second Language Writing*, 16(3), 129-147.
- Cress, U., & Hesse, F. W. (2013). Quantitative methods for studying small groups. In C. E. Hmelo-Silver, C. A. Chinn, C. K. Chan & A. O'Donnell (Eds.), *The international handbook of collaborative learning* (pp. 93-111). New York, NY: Routledge.

- Davies, M. (2010). The Corpus of Contemporary American English as the first reliable monitor corpus of English. *Literary and Linguistic Computing*, 25(4), 447-464.
- Ding, Y. (2007). Text memorization and imitation: The practices of successful Chinese learners of English. *System*, 35(2), 271-280.
- Doane, D. P., & Seward, L. E. (2011). Measuring skewness, *Journal of Statistics Education*, 19(2), 1-18.
- Dornyi, Z. (2007) *Research methods in applied linguistics*. Oxford: Oxford University Press.
- Ellis, N. C. (1996). Sequencing in SLA. *Studies in Second Language Acquisition*, 18(1), 91-126.
- Ellis, N. C. (2007). The Associative-cognitive CREED. In B. VanPatten & J. Williams (Eds.), *Theories in second language acquisition: An introduction* (pp. 77-95). Mahwah, NJ: Lawrence Erlbaum.
- Ellis, N. C., & Simpson-Vlach, R. (2009). Formulaic language in native speakers: Triangulating psycholinguistics, corpus linguistics, and education. *Corpus Linguistics and Linguistic Theory*, 5(1), 61-78.
- Ellis, N. C., Simpson-Vlach, R., & Maynard, C. (2008). Formulaic language in native and second language speakers: Psycholinguistics, corpus linguistics, and TESOL. *TESOL Quarterly*, 42(3), 375-396.
- Ellis, R. (2002). Grammar teaching- Practice or consciousness raising? In J. C. Richards & W. A. Renandya (Eds.), *Methodology in language teaching: An anthology of current practice* (pp. 167-174). Cambridge: Cambridge University Press.
- Erman, B., & Warren, B. (2000). The idiom principle and the open choice principle. *Text*, 20(1), 29-62.

- Ferris, D. (2002). Teaching students to self-edit. In J. C. Richards & W. A. Renandya (Eds.), *Methodology in language teaching: An anthology of current practice* (pp. 328-334). Cambridge: Cambridge University Press.
- Fitzpatrick, T. (2005). Formulaic sequences. *System*, 33(2), 377-380.
- Gentil, G. (2006). EAP and technical writing without borders: The impact of departmentalization on the teaching and learning of academic writing in a first and second language. In P. K. Matsuda, C. Ortmeire-Hooper & X. You (Eds.), *The politics of second language writing: In search of the promised land* (pp.147-167). West Lafayette, IN: Parlor Press.
- Gibbs, R. W., & Gonzales, G. P. (1985). Syntactic frozenness in processing and remembering idioms. *Cognition*, 20(3), 243-259.
- Gliem, J. A., & Gliem, R. R. (2003). *Calculating, interpreting, and reporting Cronbach's alpha reliability coefficient for Likert-type scales*. Paper presented at the Proceedings of the 2011 Midwest Research-to-Practice Conference in Adult, Continuing, and Community Education, The Ohio State University.
- Granger, S. (1998). Prefabricated patterns in advanced EFL writing: Collocations and formulae. In A. P. Cowie (Ed.), *Phraseology: Theory, analysis, and applications*. Oxford: Clarendon Press.
- Gregory, M. L., Raymond, W. D., Bell, A., Fosler-Lussier, E., & Jurafsky, D. (1999). The effects of collocational strength and contextual predictability in lexical production. *Chicago Linguistic Society*, (35), 151-166.
- Hall, C. J., Smith, P. H., & Wicaksono, R. (2011). *Mapping applied linguistics: A guide for students and practitioners*. London: Routledge.

- Handl, S. (2008). Essential collocations for learners of English: The role of collocational direction and weight. In F. Meunier & S. Granger (Eds.), *Phraseology in foreign language learning and teaching* (pp. 43-66). Amsterdam: John Benjamins.
- Hill, J. (2000). Revising priorities: From grammatical failure to collocational success. In M. Lewis (Ed.), *Teaching collocations: Further developments in the lexical approach* (pp. 47-69). Boston, MI: Heinle.
- Hinkel, E. (2002). *Second language writers' text: Linguistics and rhetoric features*. Mahwah, NJ: Erlbaum.
- Howarth, P. (1998). Phraseology and second language proficiency. *Applied linguistics*, 19(1), 24-44.
- Hunston, S. (2002). *Corpora in applied linguistics*. Cambridge: Cambridge University Press.
- Hunt, A., & Beglar, D. (2002). Current research and practice in teaching vocabulary. In J. C. Richards & W. A. Renandya (Eds.), *Methodology in language teaching: An anthology of current practice* (pp. 258-266). Cambridge: Cambridge University Press.
- Hyland, K. (2003). *Second language writing*. Cambridge: Cambridge University Press.
- Hyland, K. (2006). *English for academic purposes: An advanced resource book*. Abingdon: Routledge.
- Hyland, K. (2007). Genre pedagogy: Language, literacy and L2 writing instruction. *Journal of Second Language Writing*, 16(3), 148-164.
- Hyland, K. (2008). As can be seen: Lexical bundles and disciplinary variation. *English for Specific Purposes*, 27(1), 4-21.

- Hyland, K. (2011). Disciplinary specificity: Discourse, context, and ESP. In D. Belcher, A. Johns & B. Paltridge (Eds.), *New directions in English for specific purposes* (pp. 6-24). Ann Arbor, MI: University of Michigan Press.
- Jakeman, V., & McDowell, C. (1999). *Insight into IELTS: Updated edition*. Cambridge: Cambridge University Press.
- Jeong, H. (2013). Verbal data analysis for understanding interactions. In C. E. Hmelo-Silver, C. A. Chinn, C. K. Chan & A. O'Donnell (Eds.), *The international handbook of collaborative learning* (pp. 168-183). New York, NY: Routledge.
- Jones, M., & Haywood, S. (2004). Facilitating the acquisition of formulaic sequences: An exploratory study in an EAP context. In N. Schmitt (Ed.), *Formulaic sequences: Acquisition, processing, and use* (pp. 269-300). Amsterdam: John Benjamins.
- Kennedy, G. (2008). Phraseology and language pedagogy: Semantic preference associated with English verbs in the British National Corpus. In F. Meunier & S. Granger (Eds.), *Phraseology in foreign language learning and teaching* (pp. 21-41). Amsterdam: John Benjamins.
- Kormos, J., & Dénes, M. (2004). Exploring measures and perceptions of fluency in the speech of second language learners. *System*, 32(2), 145-164.
- Kress, G. (1994). *Learning to write*. London: Routledge.
- Leki, I. (2006). The legacy of first-year composition. In P. K. Matsuda, C. Ortmeire-Hooper & X. You (Eds.), *The politics of second language writing: In search of the promised land* (pp. 59-74). West Lafayette, IN: Parlor Press.

- Lewis, M. (1997). Pedagogical implications of the lexical approach. In J. Coady & T. Huckin (Eds.), *Second language vocabulary acquisition: A rationale for pedagogy* (pp. 255-270). Cambridge: Cambridge University Press.
- Lewis, M. (2000a). Language in the lexical approach. In M. Lewis (Ed.), *Teaching collocations: Further developments in the lexical approach* (pp. 126-154). Boston, MI: Heinle.
- Lewis, M. (2000b). Learning in the lexical approach. In M. Lewis (Ed.), *Teaching collocations: Further developments in the lexical approach* (pp. 155-185). Boston, MI: Heinle.
- Lewis, M. (2000c). Materials and resources for teaching collocation. In M. Lewis (Ed.), *Teaching collocations: Further developments in the lexical approach* (pp. 186-204). Boston, MI: Heinle.
- Li, J., & Schmitt, N. (2009). The acquisition of lexical phrases in academic writing: A longitudinal case study. *Journal of Second Language Writing, 18*(2), 85-102.
- Littlewood, W. (2004). Second language learning. In A. Davies & C. Elder (Eds.), *The handbook of applied linguistics* (pp. 501-524). Malden, MA: Blackwell Publishing.
- Liu, W., & Huo, Y. (2011). On the role of formulaic sequences in second language acquisition. *US-China Foreign Language, 9*(11), 31-36.
- Llach, A. (2011). *Lexical errors and accuracy in foreign language writing*. Bristol: Multilingual Matters.
- Lumley, T., Diehr, P., Emerson, S., & Chen, L. (2002). The importance of the normality assumption in large public health data sets. *Annu Rev Public Health, 23*, 151-169.
- Martin, W. (2008). A unified approach to semantic frames and collocational patterns. In S. Granger & F. Meunier (Eds.), *Phraseology: An interdisciplinary perspective* (pp. 51-65). Amsterdam: John Benjamins.

- Martinez, R., & Schmitt, N. (2012). A phrasal expressions list. *Applied linguistics*, 33(3), 299-320.
- Miller, J., & Weinert, R. (1998). *Spontaneous spoken language: Syntax and discourse*. Oxford: Clarendon Press.
- Moon, R. (1997). Vocabulary connection: Multi-word items in English. In N. Schmitt & M. McCarthy (Eds.), *Vocabulary: Description, acquisition, and pedagogy* (pp. 40-63). Cambridge: Cambridge University Press.
- Nation, I. S. P. (1990). *Teaching and learning vocabulary*. Boston, MA: Heinle & Heinle Publishers.
- Nation, I. S. P. (2001). *Learning vocabulary in another language*. Cambridge: Cambridge University Press.
- Nation, I. S. P. (2002). Best practice in vocabulary teaching and learning. In J. C. Richards & W. A. Renandya (Eds.), *Methodology in language teaching: An anthology of current practice* (pp. 267-272). Cambridge: Cambridge University Press.
- Nation, I. S. P., & Newton, J. (1997). Teaching vocabulary. In J. Coady & T. Huckin (Eds.), *Second language vocabulary acquisition: A rationale for pedagogy* (pp. 238-254). Cambridge: Cambridge University Press.
- Nattinger, J. R., & DeCarrico, J. S. (1992). *Lexical phrases and language teaching*. Oxford: Oxford University Press.
- Oakey, D. (2002). Formulaic language in English academic writing: A corpus-based study of the formal and functional variation of a lexical phrase in different academic disciplines. In R. Reppen, S. M. Fitzmaurice & D. Biber (Eds.), *Using corpora to explore linguistic variation: Studies in corpus linguistics* (pp. 111-130). Amsterdam: John Benjamins.

- O'Keefe, A. (2012). Vocabulary instruction. In A. Burns & J. Richards (Eds.), *The Cambridge guide to pedagogy and practice in second language teaching* (pp. 236-245). Cambridge: Cambridge University Press.
- Omazic, M. (2008). Processing of idioms and idiom modifications: A view from cognitive linguistics. In S. Granger & F. Meunier (Eds.), *Phraseology: An interdisciplinary perspective* (pp. 67-79). Amsterdam: John Benjamins.
- Osborne, J. (2008). Phraseology effects as a trigger for errors in L2 English: The case of more advanced learners. In F. Meunier & S. Granger (Eds.), *Phraseology in foreign language learning and teaching* (pp. 67-83). Amsterdam: John Benjamins.
- Paltridge, B. (2004). Academic writing. *Language teaching*, 37(2), 87-105.
- Paquot, M. (2008). Exemplification in learning writing: A cross-linguistic perspective. In F. Meunier & S. Granger (Eds.), *Phraseology in foreign language learning and teaching* (pp. 101-119). Amsterdam: John Benjamins.
- Paré, A. (2009). What we know about writing, and why it matters. *Compendium 2*, 2(1), 1-12.
- Pawley, A., & Syder, F. H. (1983). Two puzzles for linguistic theory: Nativelike selection and nativelike fluency. In J. C. Richards & R. W. Schmidt (Eds.), *Language and communication* (pp. 191-226). New York, NY: Longman.
- Philip, G. (2008). Reassessing the canon: 'Fixed' phrases in general reference corpora. In S. Granger & F. Meunier (Eds.), *Phraseology: An interdisciplinary perspective* (pp. 95-108). Amsterdam: John Benjamins.
- Qi, Y., & Ding, Y. (2011). Use of formulaic sequences in monologues of Chinese EFL learners. *System*, 39(2), 164-174.

- Raimes, A. (2002). The steps in planning a writing course and training teachers of writing. In J. C. Richards & W. A. Renandya (Eds.), *Methodology in language teaching: An anthology of current practice* (pp. 306-314). Cambridge: Cambridge University Press.
- Razali, N. M., & Wah, Y. B. (2011). Power comparisons of Shapiro-Wilk, Kolmogorov-Smirnov, Lilliefors and Anderson-Darling tests. *Journal of Statistical Modeling and Analytics*, 2(1), 21-33.
- Reppen, R. (2002). A genre-based approach to content writing instruction. In J. C. Richards & W. A. Renandya (Eds.), *Methodology in language teaching: An anthology of current practice* (pp. 321-327). Cambridge: Cambridge University Press.
- Richards, J. C. (2002). Addressing the grammar gap in task work. In J. C. Richards & W. A. Renandya (Eds.), *Methodology in language teaching: An anthology of current practice* (pp. 153-166). Cambridge: Cambridge University Press.
- Richards, J. C., & Renandya, W. A. (Eds.). (2002). *Methodology in language teaching: An anthology of current practice*. Cambridge: Cambridge University Press.
- Richards, J. C., & Rodgers, T. S. (2001). *Approachers and methods in language teaching* (2nd ed.). Cambridge: Cambridge University Press.
- Schmitt, N. (2000). *Vocabulary in language teaching*. Cambridge: Cambridge University Press.
- Schmitt, N., & Carter, R. (2004). Formulaic sequences in action: An introduction. In N. Schmitt (Ed.), *Formulaic sequences: Acquisition, processing, and use* (pp. 1-22). Amsterdam: John Benjamins.
- Schmitt, N., Dornyei, Z., Adolphs, S., & Durow, V. (2004). Knowledge and acquisition of formulaic sequences: A longitudinal study. In Schmitt, N. (Ed.), *Formulaic sequences: Acquisition, processing, and use* (pp. 55-86). Amsterdam: John Benjamins.

- Schmitt, N., & Underwood, G. (2004). Exploring the processing of formulaic sequences through a self-paced reading task. In N. Schmitt (Ed.), *Formulaic sequences: Acquisition, processing and use*, (pp. 173-190). Amsterdam: John Benjamins.
- Seow, A. (2002). The writing process and process writing. In J. C. Richards & W. A. Renandya (Eds.), *Methodology in language teaching: An anthology of current practice* (pp. 315-320). Cambridge: Cambridge University Press.
- Shapiro, S. S., & Wilk, M. B. (1965). An analysis of variance test for normality (complete samples). *Biometrika*, 52(3), 591-611.
- Silva, T. (1993). Toward an understanding of the distinct nature of L2 writing: The ESL research and its implications. *TESOL Quarterly*, 27(4), 657-677.
- Simpson-Vlach, R., & Ellis, N. C. (2010). An academic formulas list: New methods in phraseology research. *Applied Linguistics*, 31(4), 487-512.
- Sinclair, J. (1991) *Corpus, concordance, collocation*. Oxford: Oxford University Press.
- Stengers, H., Boers, F., Housen, A., & Eyckmans, J. (2011). Formulaic sequences and L2 oral proficiency: Does the type of target language influence the association? *International Review of Applied Linguistics in Language Teaching*, 49(4), 321-343.
- Swan, M. (2002). Seven bad reasons for teaching grammar- And two good ones. In J. C. Richards & W. A. Renandya (Eds.), *Methodology in language teaching: An anthology of current practice* (pp. 148-152). Cambridge: Cambridge University Press.
- Tabossi, P., Fanari, R., & Wolf, K. (2009). Why are idioms recognized fast? *Memory & Cognition*, 37(4), 529-540.

- Tremblay, A., & Baayen, H. (2010). Holistic processing of regular four-word sequences: A behavioural and ERP study of the effects of structure, frequency, and probability of immediate free recall. In D. Wood (Ed.), *Perspectives on formulaic language: Acquisition and communication* (pp. 151-173). London: Continuum.
- VanPatten, B., & Williams, J. (2007). Early theories in second language acquisition. In B. VanPatten & J. Williams (Eds.), *Theories in second language acquisition: An introduction* (pp. 17-35). Mahwah, NJ: Lawrence Erlbaum.
- Wajnryb, R. (1990). *Grammar dictation*. Oxford: Oxford University Press.
- Weinert, R. (1995). The role of formulaic language in second language acquisition: A review. *Applied Linguistics*, 16(2), 180-205.
- Williams, G. C. (2008). The Good Lord and his works: A corpus-driven study of collocational resonance. In S. Granger & F. Meunier (Eds.), *Phraseology: An interdisciplinary perspective* (pp. 159-173). Amsterdam: John Benjamins.
- Wood, D. (2001). In search of fluency: What is it and how can we teach it? *Canadian Modern Language Review*, 57(4), 573-589.
- Wood, D. (2005). *Formulaic language in speech fluency development in English as a second language*. (Unpublished doctoral dissertation). University of Ottawa, Ottawa.
- Wood, D. (2006). Uses and functions of formulaic sequences in second language speech: An exploration of the foundations of fluency. *Canadian Modern Language Review*, 63(1), 13-33.
- Wood, D. (2009a). Effects of focused instruction of formulaic sequences on fluent expression in second language narratives: A case study. *Canadian Journal of Applied Linguistics*, 12(1), 39-57.

- Wood, D. (2009b). Formulaic language in acquisition and production: Implications for teaching. *TESL Canada Journal*, 20(1), 1-15.
- Wood, D. (2010a). *Formulaic language and second language speech fluency: Background, evidence and classroom applications*. London: Continuum International Publishing Group.
- Wood, D. (2010b). *Perspectives on formulaic language: Acquisition and communication*. London: Continuum.
- Woolard, G. (2000). Collocations- Encouraging learner independence. In M. Lewis (Ed.), *Teaching collocations: Further developments in the lexical approach* (pp. 28-46). Boston, MI: Heinle.
- Wray, A. (1999). Formulaic language in learners and native speakers. *Language Teaching*, 32(4), 213-231.
- Wray, A. (2000). Formulaic sequences in second language teaching: Principle and practice. *Applied Linguistics*, 21(4), 463-489.
- Wray, A. (2002). *Formulaic language and the lexicon*. Cambridge: Cambridge University Press.
- Wray, A. (2012). What do we (think we) know about formulaic language? An evaluation of the current state of play. *Annual Review of Applied Linguistics*, 32, 231–254.
- Wray, A., & Fitzpatrick, T. (2008). Why can't you just leave it alone? Deviations from memorized language as a gauge of nativelike competence. In F. Meunier & S. Granger (Eds.), *Phraseology in foreign language learning and teaching* (pp. 123-147). Amsterdam: John Benjamins.
- Wray, A., & Fitzpatrick, T. (2010). Pushing learners to the extreme: The artificial use of prefabricated material in conversation. *Innovation in language learning and teaching*, 4(1), 37-51.

Wray, A., & Perkins, M. R. (2000). The functions of formulaic language: An integrated model

Language and Communication, 20(1), 1-28.

Zhu, W. (2006). Understanding context for writing in university content classrooms. In P. K.

Matsuda, C. Ortmeire-Hooper & X. You (Eds.), *The politics of second language writing: In search of the promised land* (pp. 129-146). West Lafayette, IN: Parlor Press.

Zimmerman, C. B. (1997). Historical trends in second language vocabulary instruction. In J.

Coady & T. Huckin (Eds.), *Second language vocabulary acquisition: A rationale for pedagogy* (pp. 5-19). Cambridge: Cambridge University Press.

Appendix A
The Three Judges' Evaluation

Table A1			
<i>The First Judge's Evaluation of the Participants' Three Elicited Texts</i>			
Participants	Pretest	Posttest	Delayed Posttest
Salem	5.5	7.5	7.5
Ellie	6	7	7
MATA	6	6.5	7.5
Jacob	6	7.5	7.5
Shireen	6.5	6	7.5
Abedeer	5.5	8	7
David	6	7.5	7
Mary	5.5	6.5	7.5
Eva	5.5	7.5	7
Ali	7.5	7.5	9
Moo	5	6.5	7
Gabriella	6	7	6

Note. The first judge did not follow any rubric, but she rather evaluated the participants' written responses based on the overall content and structure.

Table A2

The Second Judge's Evaluation of the Participants' Three Elicited Texts

Participants	Pretest	Posttest	Delayed Posttest
Salem	3.5	5.5	5.5
Ellie	5	7	6
MATA	7	7.5	9
Jacob	5.5	8	7
Shireen	4	4	5
Abedeer	6.5	9	7
David	6.5	6	8
Mary	5.5	8	7.5
Eva	4.5	9	8
Ali	5.5	7.5	7
Moo	4	6	6
Gabriella	6.5	5.5	6.5

Note. The third judge did not follow any rubric, but he rather evaluated the participants' written responses based on the overall content and structure.

Table A3

The Third Judge's Evaluation of the Participants' Three Elicited Texts

Participants	Pretest	Posttest	Delayed Posttest
Salem	5.5	6	6.5
Ellie	7	8	7
MATA	5.5	6	6.5
Jacob	6	6.5	6.5
Shireen	6	6.5	6.5
Abedeer	6.5	7	7
David	6.5	7	7.5
Mary	6.5	7.5	7
Eva	7	8.5	8
Ali	6.5	7	6.5
Moo	5.5	6.5	7
Gabriella	7	7.5	7.5

Note. The second judge did not follow any rubric, but he rather evaluated the participants' written responses based on the overall content and structure.

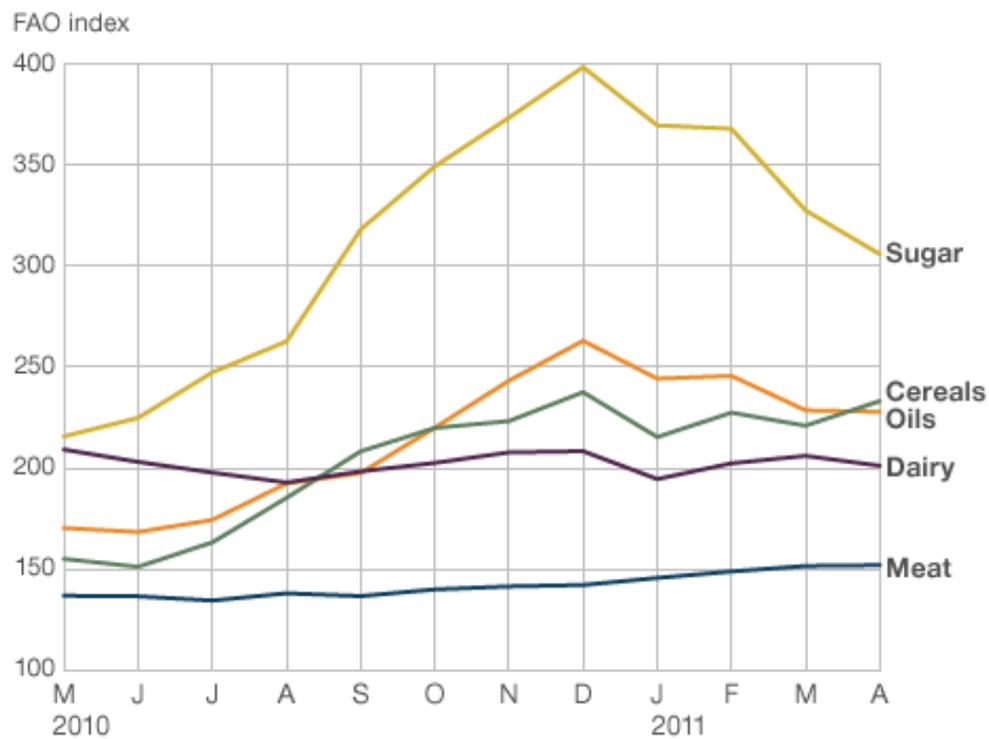
Appendix B

Instruments

Prompt Used for Data Collection

The graph below gives information about the changes in commodity prices over a one-year period. Summarize the information provided in no more than 150 words then explain how such changes in prices might affect consumers.

Commodity prices over the last 12 months



Source: FAO

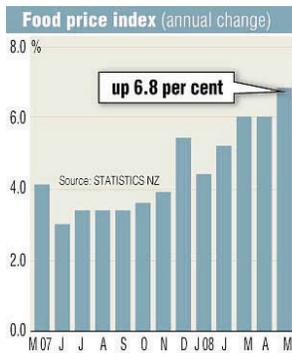
Retrieved from BBC News: Business

<http://www.bbc.co.uk/news/business-11177214>

Worksheet 1: Topic Sentences

The function of a topic sentence is to give the reader a general idea about the whole paragraph.

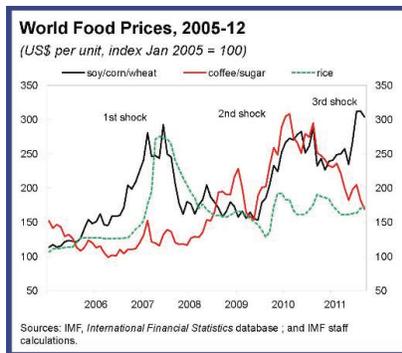
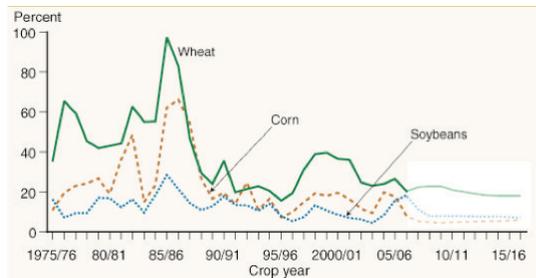
Note that, in academic writing, a topic sentence usually appears at the beginning; it also includes all the information that is going to be presented. Examples:

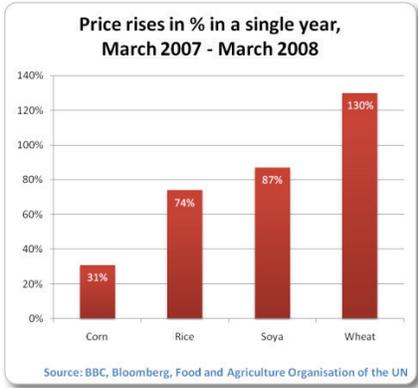


The graph presents the rise and fall in the general prices of food over a period of twelve months.

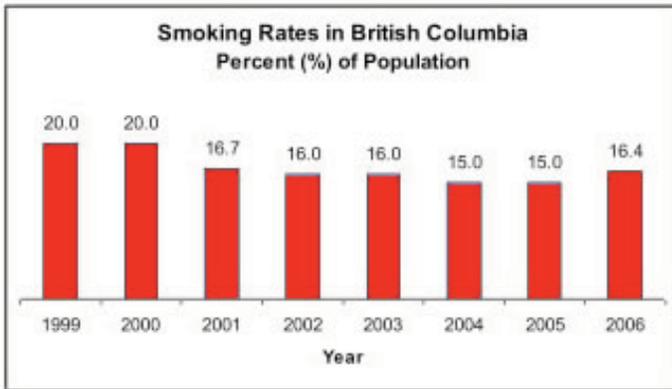
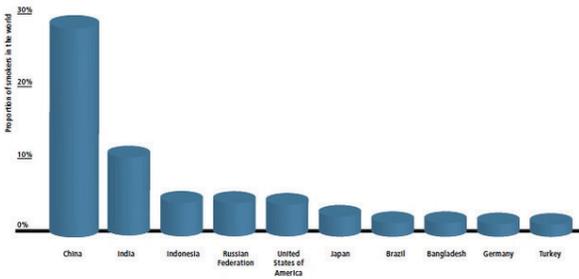
OR

The graph shows the rise and fall in the general prices of food between May 2007 and May 2008.





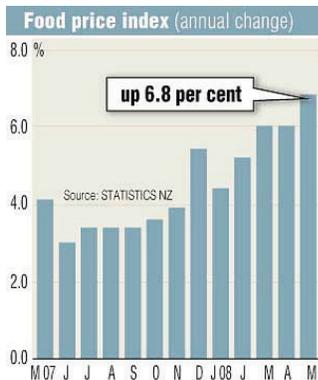
NEARLY TWO THIRDS OF THE WORLD'S SMOKERS LIVE IN 10 COUNTRIES



Worksheet 2: Concluding Sentences

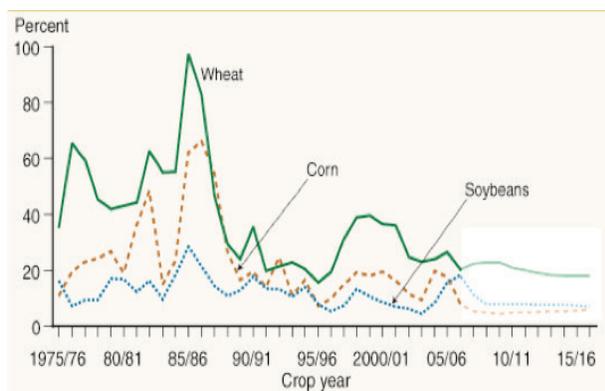
The function of a concluding sentence is to summarize all the ideas presented by the writer. Note that, in academic writing, a concluding sentence always appears at the end; we usually start our concluding sentences with such phrases as in conclusion, in sum, to sum up, to conclude, etc.

Examples:

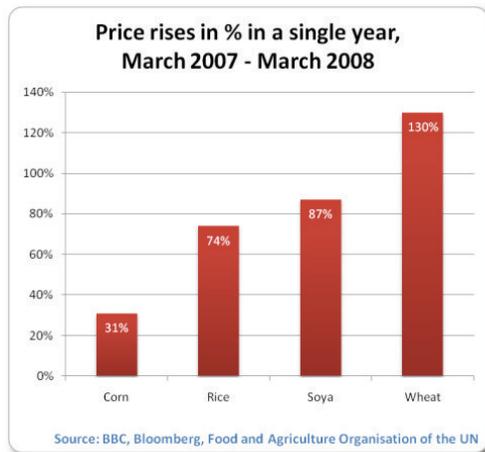
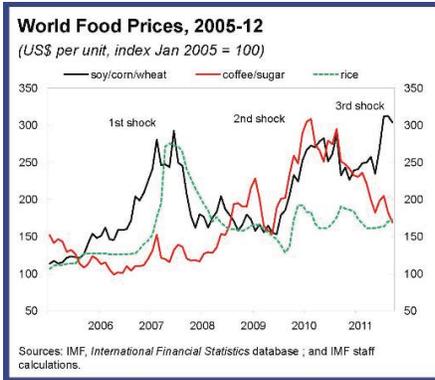


In conclusion, it is clear from the data given that food prices generally increased between May 2007 and May 2008

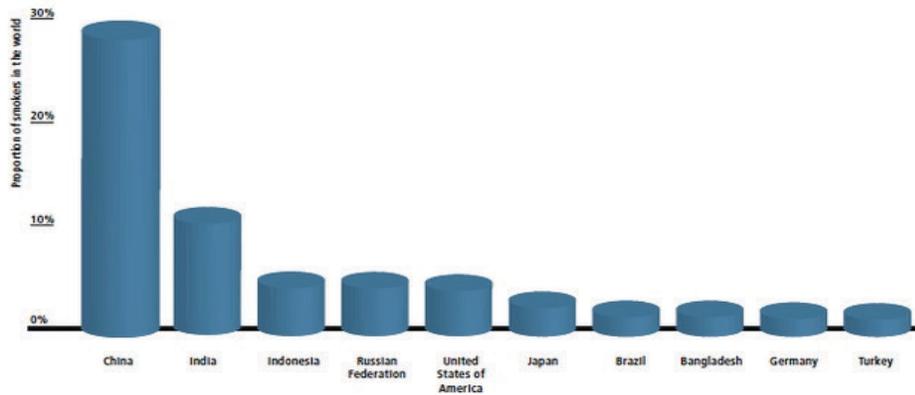
OR: To sum up, it can be seen that the prices generally increased between May 2007 and May 2008 despite the slight decrease in June 2007 and January 2008.

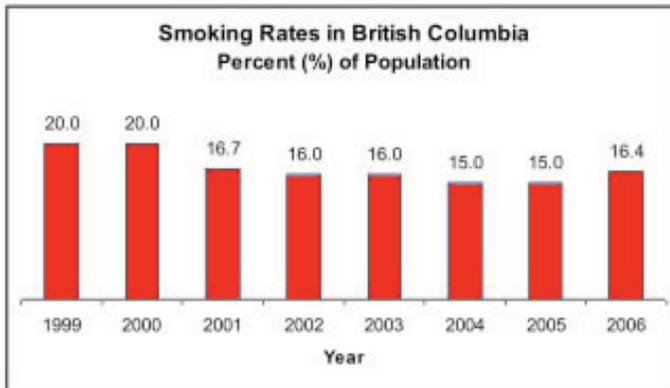


.....



NEARLY TWO THIRDS OF THE WORLD'S SMOKERS LIVE IN 10 COUNTRIES



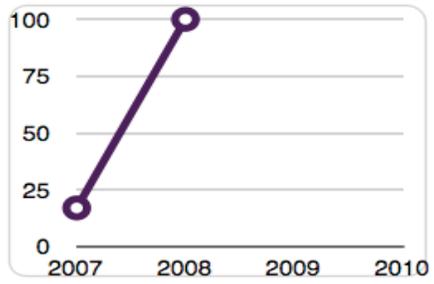
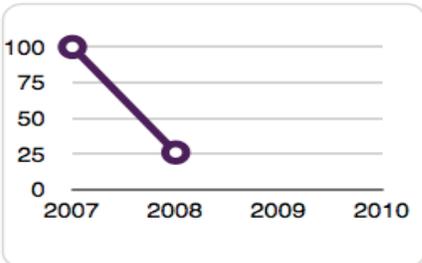
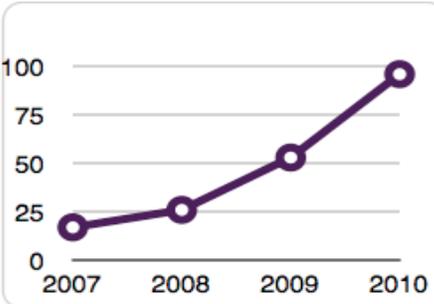
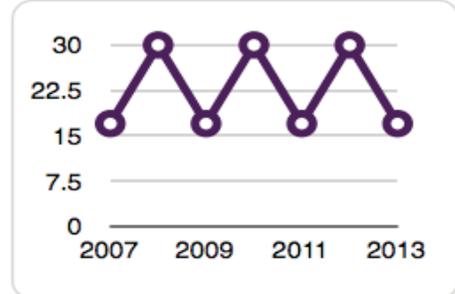
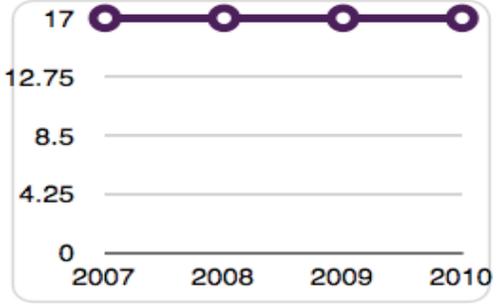
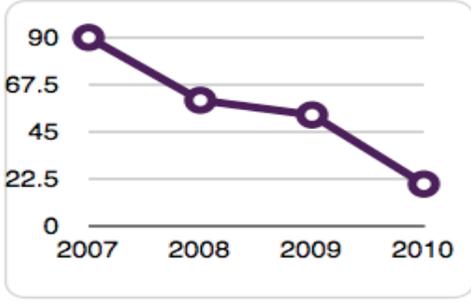


Worksheet 3: List of Collocations

Verbs	Nouns
Increase dramatically Increase significantly	Significant increase Dramatic increase Sharp increase Slight increase
Decrease significantly Decrease slightly Decrease dramatically	Significant decrease Dramatic decrease Sharp decrease Slight decrease
	Steady decline Sharp decline Significant decline Rapid decline Gradual decline Dramatic decline
Rise and fall Prices fluctuate	The rise and fall The fluctuation in the price
Remain the same Remain stable	

Worksheet 4: Collocation Practice

Describe the following patterns using the phrases between brackets (Remain the same, gradual increase, sharp fall, slightly decreased, increased dramatically, fluctuate, remain stable, increased gradually, dramatic increase, dramatic decrease, gradual decline, rise and fall evenly).

 <p>..... OR.....</p>	 <p>..... OR.....</p>
 <p>..... OR.....</p>	 <p>..... OR.....</p>
 <p>..... OR.....</p>	 <p>..... OR.....</p>

Worksheet 5: Formulaic Sequences

These phrases are beneficial for your academic writing as they help you write a well-connected paragraph and/or essay

- ***On the one hand/ on the other hand***

They are used to introduce two opposing ideas.

Examples:

-On the one hand, I'd like more money, but on the other hand, I'm not prepared to work the extra hours in order to get it.

If your sentences are long, separate them with a full stop.

-On the one hand, the prices of Honda, Ford, and Chevrolet have increased dramatically. On the other hand, the prices of Kia and Mercedes remained the same.

- ***Both (A) and (B)***

It is usually used to avoid repetition in writing when two sentences present the same information about two different subjects.

Examples:

Instead of writing:

Russia is cold in winter. Canada is cold in winter.

You can write

Both Russia and Canada are cold in winter.

- ***(A) and (B)***

It is usually used to avoid repetition in writing when two sentences present the same information about two different subjects.

Example:

Instead of writing:

Russia is cold in winter. Canada is cold in winter.

You can write

Russia and Canada are cold in winter.

- ***as well as***

It is used to connect two items together.

Example:

The service was good as well as efficient.

- ***In other words***

This is usually used at the beginning of a sentence to restate what you have just written in different words

Example:

We are broke. In other words, we do not have money.

- ***In addition to***

In addition to is usually used to help smooth transition from one sentence into the other.

Example:

The price of rice has slightly increased. In addition to rice, the price of wheat has also increased from \$ 2.15 to \$ 2.24 per k.g.

- ***For example/ For instance/ Such as***

They are usually used to list examples of something that has already been introduced.

Examples:

I can play several musical instruments, for example, the flute, the guitar, and the piano.

I can play several musical instruments, for instance, the flute, the guitar, and the piano.

I can play several musical instruments such as the flute, the guitar, and the piano.

Note:

If you decide to list examples in a separate sentence **ONLY** for example and for instance can be used.

I can play several musical instruments. For example, I can play the flute, the guitar, and the piano.

I can play several musical instruments. For instance, I can play the flute, the guitar, and the piano.

In this case, ‘such as’ cannot be used.

Between (A) and (B)/ From (A) to (B)

These are generally used to introduce periods of time. They can be used at the beginning of a sentence or at the end.

Examples:

-From 1999 to 2009, Canada’s food prices increased dramatically. Or

Canada’s food prices increased dramatically from 1999 to 2008.

-Between 1999 and 2009, Canada’s food prices increased dramatically. Or

Canada’s food prices increased dramatically between 1999 and 2008.

The (A) of...

It is most commonly used to construct a noun phrase.

Examples:

The prices of several products have increased.

The price of this dress is so expensive.

Worksheet 6: Practice

Rewrite the following sentences using the phrases provided.

1. Vegetables are good for your health. Fruits are good for your health. (Both *A* and *B*)

.....

2. I have read several novels this year. I have read *Pride and Prejudice* and *Sons and Lovers*. (for instance)

.....

3. The price of milk has increased dramatically since 1994. The price of oil has decreased considerably over the last few years. (On the one hand/ On the other hand)

.....

4. Wear simple things. Wear a skirt and a blouse. (such as)

.....

5. My father loves going to restaurants that serve Chinese food. Last week he went to a restaurant which serves delicious dumplings. (for example)

.....

6. The price of rice has remained stable. The price of oil has remained the same. (In addition to)

.....

7. Tourism decreased due to the economic recession (add time period using between and/2003-2011).

.....

8. He has another appointment on Thursday. I don't think he will be attending your gathering. (In other words).

.....

9. He is strong. He is courageous. (as well as)

.....

10. I lived in Paris (add time period using from to/2007-2010).

.....

11. Sugar prices have increased slightly, but rice prices have decreased (rewrite the underlined words using the (A) of).

.....

Worksheet 7: Formulaic Sequences

Expressing Opinions:

In academic context, we might be asked to express our opinions in writing. In such cases, we use certain phrases/clauses to help us express our viewpoints.

- **It has been generally believed/noted**

This clause is usually used when we write about things that have been accepted as facts or commonsense. It usually comes at the beginning of a sentence and is followed by that + full sentence.

Examples:

It has been generally believed that smoking causes heart and lung diseases.

It has been generally noted that regular exercise is good for your health.

- **It is possible**

This clause is usually used when we write about things that are possible. It usually comes at the beginning of a sentence and is followed by either to + V or that + full sentence. *Examples:*

It is possible to help smokers in early motivational stages to quit.

It is possible that he will have to return to the hospital.

- **It is necessary to.....**

This clause is used to indicate that something is required or essential. It must be followed by a verb. *Example:*

It is necessary to buy your tickets in advance.

Other Useful Phrases

- **..... plays an important role in**

This phrase is usually used to indicate the effect of something on something else.

Example:

Regular exercise plays an important role in the health, well-being, and quality of life of Canadians.

-as a result of/due to the fact that....

These phrases are usually used to express cause and effect. *Examples:*

The tower has collapsed as a result of safety violations.

John decided to move to the United States due to the fact that he could not find a job in Canada.

Note that these phrases can also come at the beginning of a sentence. *Examples*

As a result of safety violations, the tower has collapsed.

Due to the fact that he could not find a job in Canada, John decided to move to the United States.

Worksheet 8: Practice

The first part of each sentence in *List 1* can be completed with the group of endings given in *List 2*

2. Match the first parts with the endings and add as many words as you need to form complete sentences.

List 1

1. It has been generally believed/noted

2. It is possible

3. play(s) an important role in

4.as a result of.....

5. due to the fact that

6. It is important/necessary to.....

List 2

a. social injustice/crime
b. smoking/lung cancer
c. exercise/good health

a. cars/by hand
b. book online
c. men/on Mars

a. practice/learning
b. persistence/success
c. technology/lives

a. couldn't walk/ accident
b. failed the exam/indifference
c. unemployment/economic depression

a. very difficult/different educational system
b. failed the exam/ did not study hard
c. EAP classes/required mark

a. work hard/improve your language skills
b. study hard/pass the test
c. finish/project on time

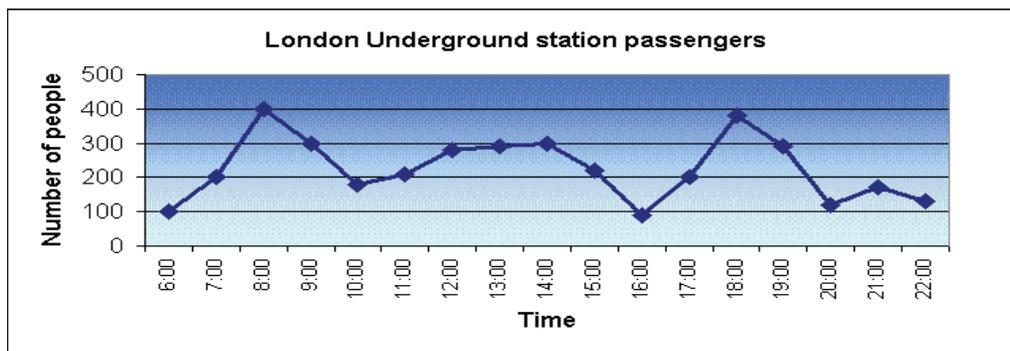
Worksheet 9: A Table of Formulaic Sequences

Listing Examples	<ul style="list-style-type: none"> -For instance/ For example - Such as
Expressing Time	<ul style="list-style-type: none"> -Between (A) and (B)/ From (A) to (B)/ Over a period of - At the beginning of/ At the end of
Expressing Opinion	<ul style="list-style-type: none"> -It has been generally believed/noted/ (A) plays an important role in the -It is necessary to/ As a result of -Due to the fact that/ Due to the -It is possible that -It can be seen/ It can be observed
Conclusion	<ul style="list-style-type: none"> -In conclusion/ In sum/ In summary/ To sum up/ To conclude -It can be seen that/ It can be observed that
Linking Phrases	<p>According to the/ Based on the/ As shown in / On the basis of</p> <ul style="list-style-type: none"> -On the one hand, -On the other hand, -At the same time, -In addition to - In other words,
Sentence Heads	<ul style="list-style-type: none"> -There is/ are - There was/were - The (A) of - Both (A) and (B) - (A) as well as (B) - (A) and (B)
Introduction	<ul style="list-style-type: none"> - The graph shows -The graph demonstrates -The graph presents

Worksheet 10: Practice Sheet

Read the description of the graph below and answer the following questions

This graph shows the fluctuation in the number of people at a London underground station over the course of a day. According to the graph, there is a sharp increase between 6:00 and 8:00 in the morning, with 400 people using the station at 8 o'clock. The numbers fall dramatically to less than 200 at 10 o'clock. Between 11:00 and 3:00, the number of people rises and falls evenly with a plateau around lunch time of just under 300 people using the station. Numbers then decline, with the lowest number being recorded at 4:00 in the afternoon. There is then a rapid rise between 4.00 and 6 p.m during the evening rush hour with a peak of 380 people at 6 p.m. After 7 p.m. numbers fall significantly, with only a slight increase again just after 8 p.m, tailing off after 9 p.m. In conclusion, based on the graph presented the station is most crowded in the early morning and early evening rush hour periods. (adapted from Jakeman & McDowell, 1999, p. 182).



1. What verbs did the writer use to introduce the differences in people numbers?
2. How did the writer refer to durations of time?
3. How did the writer start his/her paragraph?
4. How did the writer conclude his/her paragraph?
5. Underline the linking words/phrases used by the writer to connect sentences?

Worksheet 11: A Cloze Activity

Complete the following text using the phrases between brackets.

(in addition to/ between 1900 and 2000 / over a period of one century/ as well as/ the graph shows/ in conclusion/ according to the)

..... how the amount of water used worldwide changed

..... graph, the largest quantity of water was used for agricultural purposes, and this increased dramatically from about 500 km³ to around 3,000 km³ in the year 2000.

..... agricultural purposes, water used in the industrial domestic sectors also increased, but consumption was minimal until mid-century. From 1950 onwards, industrial use grew steadily to just over 1,000 km³, while domestic use rose more slowly to only 300 km³, both far below the levels of consumption by agriculture.

....., the consumption of water for different purposes has increased

.....

Adapted from IELTS Sample Tests

http://www.ielts-exam.net/academic_writing_samples_task_1/589/

Worksheet 12: Full Text for Dictogloss

The graph shows the differences between men and women's achievements in running events. According to the graph, men are faster than women in such events. On the one hand, the best male runners can cover 640 miles in the six-day running event. On the other hand, the fastest woman can cover only 548 miles. This represents a difference of 16.8 per cent. This difference becomes less significant in the shorter events; for example, in the 100 m, there is only a 6.1 per cent difference in performance between men and women.

(Jakeman & McDowell, 1999, p. 182)

It has been generally believed that men run faster than women due to the fact that men's bodies are stronger and they have more muscles than fat as compared to women. In addition to this, men's lungs are bigger, and, thus, they can breathe faster than women. Furthermore, it is possible that the key male sex hormone, testosterone, promotes men's performance in such contests. In conclusion, as this graph shows, men usually run faster than women, especially in long distance running events.

www.oprah.com/health/Are-Men-and-Women-Different_1/12

Appendix C

Formulaic Sequences Extracted from the COCA

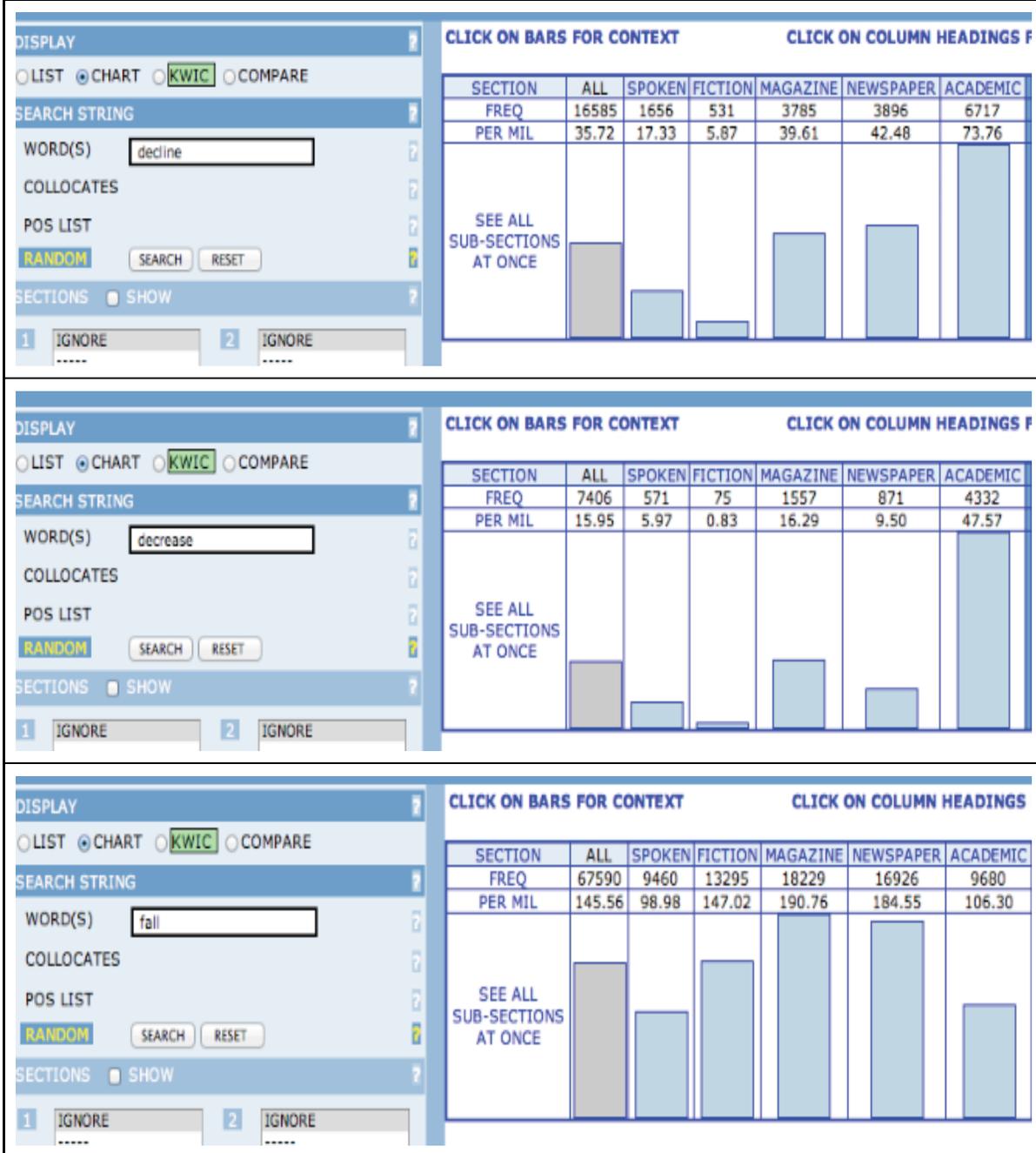
Table C1

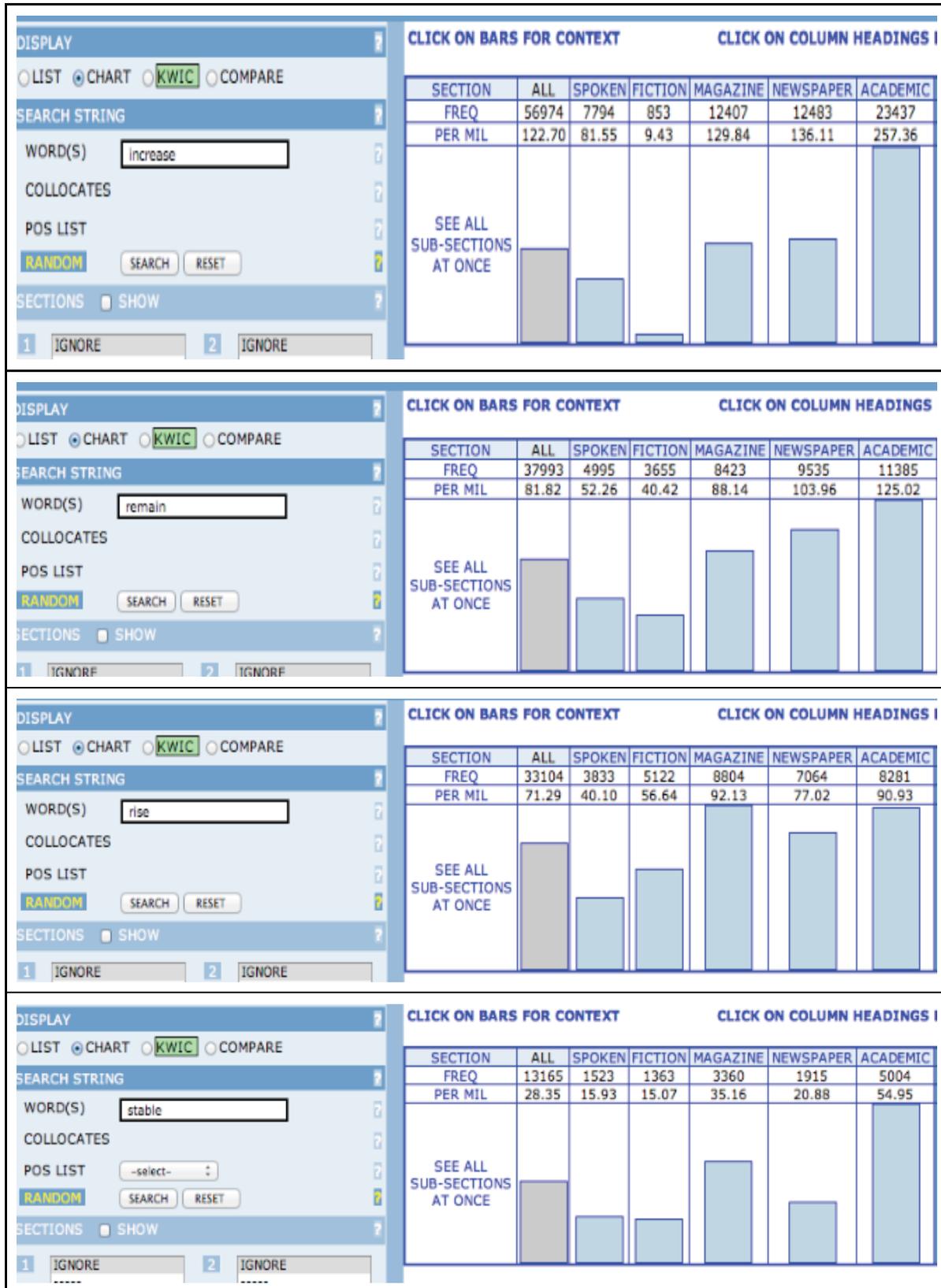
The Formulaic Sequences Extracted from the COCA



Table C2

The Frequency of the Stem Collocations in the COCA





DISPLAY

LIST CHART **KWIC** COMPARE

SEARCH STRING

WORD(S)

COLLOCATES

POS LIST

RANDOM SEARCH RESET

SECTIONS SHOW

1 IGNORE

SPOKEN

FICTION

2 IGNORE

SPOKEN

FICTION

CLICK ON BARS FOR CONTEXT **CLICK ON COLUMN**

SECTION	ALL	SPOKEN	FICTION	MAGAZINE	NEWSPAPER	ACADEMIC	1990-1994	1995-1999	2000
FREQ	454	56	15	135	73	175	109	116	1
PER MIL	0.98	0.59	0.17	1.41	0.80	1.92	1.05	1.12	0

SEE ALL SUB-SECTIONS AT ONCE

	History	Education	Geog/SocSc	Law/PolSci	Humanities	Phil/Rel	Sci/Tech
	18	13	43	12	13	5	59
1.47	1.38	2.66	1.40	1.09	0.74	4.19	

DISPLAY

LIST CHART **KWIC** COMPARE

SEARCH STRING

WORD(S)

COLLOCATES

POS LIST

RANDOM SEARCH RESET

SECTIONS SHOW

1 IGNORE

SPOKEN

FICTION

MAGAZINE

NEWSPAPER

2 IGNORE

SPOKEN

FICTION

MAGAZINE

NEWSPAPER

CLICK ON BARS FOR CONTEXT **CLICK ON COLUMN**

SECTION	ALL	SPOKEN	FICTION	MAGAZINE	NEWSPAPER	ACADEMIC	1990-1994	1995-1999	2000
FREQ	348	16	31	62	40	199	87	95	7
PER MIL	0.75	0.17	0.34	0.65	0.44	2.19	0.84	0.92	0.

SEE ALL SUB-SECTIONS AT ONCE

	History	Education	Geog/SocSc	Law/PolSci	Humanities	Phil/Rel	Sci/Tech
	17	9	29	7	10	16	91
1.39	0.95	1.79	0.81	0.84	2.37	6.47	

Table C3

The Collocations Extracted from the COCA

WORD:

LIST FROM #: (1-60,000)

PART OF SPEECH: NOUN VERB ADJ ADV MISC

[6 / 11] HIDE HELP << < > >>

The **collocates** show you the nearby words (shown here for *moonlight*), and these provide useful information on the meaning and use of a word. They are grouped by part of speech and then listed in order (most frequent first). Note that any collocates in parentheses (mainly for very low frequency words) are tentative.

SYNONYMS (click to see) [?] SEE ALSO: MORE SPECIFIC (18) MORE GENERAL (7) MULTI-WORD (1)

deterioration
2572 drop
2763 decline

SINGLE ENTRY: DISPLAYED IN FRAME BELOW

RANK #	POS	WORD	TOTAL	SPOKEN	FICTION	MAGAZINE	NEWSPAPER	ACADEMIC
1	N	DECLINE	12849	1266	324	2818	3024	5417

DECLINE *n* (RANK 2763, FREQ 12849)

	SPOKEN	FICTION	MAGAZINE	NEWSPAPER	ACADEMIC
CLICK BAR TO LIMIT					
STORED	12	4	46	50	90
MORE	1266	324	2818	3024	5417

DEFINITIONS (WORDNET) (BAD ENTRY?)
1. change toward something smaller or lower **2.** a condition inferior to an earlier condition **3.** a gradual decrease **4.** a downward slope or bend

COLLOCATES (click to see with DECLINE)
adj economic, steady, sharp, steep, significant, rapid, gradual, dramatic, general **noun** percent, rate, year, number, population, price, market, stock, reason, value **verb** show, lead, cause, experience, contribute, suffer, reverse, occur, appear, reveal

WORD:

LIST FROM #: (1-60,000)

PART OF SPEECH: NOUN VERB ADJ ADV MISC

[6 / 11] HIDE HELP << < > >>

The **collocates** show you the nearby words (shown here for *moonlight*), and these provide useful information on the meaning and use of a word. They are grouped by part of speech and then listed in order (most frequent first). Note that any collocates in parentheses (mainly for very low frequency words) are tentative.

SYNONYMS (click to see) [?] SEE ALSO: MORE SPECIFIC (31) MORE GENERAL (7) MULTI-WORD (1)

refuse
2010 decline

SINGLE ENTRY: DISPLAYED IN FRAME BELOW

RANK #	POS	WORD	TOTAL	SPOKEN	FICTION	MAGAZINE	NEWSPAPER	ACADEMIC
1	V	DECLINE	19593	2114	914	3996	7277	5292

DECLINE *v* (RANK 2010, FREQ 19593)

	SPOKEN	FICTION	MAGAZINE	NEWSPAPER	ACADEMIC
CLICK BAR TO LIMIT					
STORED	11	7	34	81	69
MORE	2114	914	3996	7277	5292

DEFINITIONS (WORDNET) (BAD ENTRY?)
(5/6) **1.** grow worse **2.** show unwillingness towards **3.** refuse to accept **4.** grow smaller **5.** of nouns, pronouns, and adjectives

COLLOCATES (click to see with DECLINE)
noun percent, year, rate, population, number, request, price, official, company, offer, stock, level, invitation, spokesman, age **misc** comment, interview, discuss, continue, begin

WORD:

LIST FROM #: (1-60,000)

PART OF SPEECH: NOUN VERB ADJ ADV MISC

[6 / 11] HIDE HELP << < > >>

The **collocates** show you the nearby words (shown here for *moonlight*), and these provide useful information on the meaning and use of a word. They are grouped by part of speech and then listed in order (most frequent first). Note that any collocates in parentheses (mainly for very low frequency words) are tentative.

SYNONYMS (click to see) [?] SEE ALSO: MORE SPECIFIC (52) MORE GENERAL (5)

reduction
1011 fail
1501 out

SINGLE ENTRY: DISPLAYED IN FRAME BELOW

RANK #	POS	WORD	TOTAL	SPOKEN	FICTION	MAGAZINE	NEWSPAPER	ACADEMIC
1	N	DECREASE	3818	248	22	578	475	2495

DECREASE *n* (RANK 6391, FREQ 3818)

	SPOKEN	FICTION	MAGAZINE	NEWSPAPER	ACADEMIC
CLICK BAR TO LIMIT					
STORED	11	1	38	23	129
MORE	248	22	578	475	2495

DEFINITIONS (WORDNET) (BAD ENTRY?)
1. a change downward **2.** a process of becoming smaller or shorter **3.** the act of decreasing or reducing something **4.** the amount by which something decreases

COLLOCATES (click to see with DECREASE)
adj significant, dramatic, slight, marked, overall, substantial, corresponding, average, total, sharp **noun** percent, number, rate, level, behavior, increase, result, price, crime, blood **verb** show, result, report, lead, cause, increase, experience, associate, indicate, occur

WORD: decrease

LIST FROM #: (1-60,000)

PART OF SPEECH: NOUN VERB ADJ ADV MISC

[6 / 11] HIDE HELP

The **collocates** show you the nearby words (shown here for *moonlight*), and these provide useful information on the meaning and use of a word. They are grouped by part of speech and then listed in order (most frequent first). Note that any collocates in parentheses (mainly for very low frequency words) are tentative.

SINGLE ENTRY: DISPLAYED IN FRAME BELOW

RANK #	POS	WORD	TOTAL	SPOKEN	FICTION	MAGAZINE	NEWSPAPER	ACADEMIC
3393	V	DECREASE	10025	617	177	2196	1011	6024

SYNONYMS (click to see) [?] SEE ALSO: MORE SPECIFIC (64) MORE GENERAL (2)

reduce

414 cut
701 reduce
2010 decline

DECREASE v (RANK 3393, FREQ 10025)

	SPOKEN	FICTION	MAGAZINE	NEWSPAPER	ACADEMIC
CLICK BAR TO LIMIT					
STORED	15	4	41	30	112
MORE	617	177	2196	1011	6024

DEFINITIONS (WORDNET) (BAD ENTRY?)

1. decrease in size, extent, or range 2. make smaller

COLLOCATES (click to see with DECREASE)

noun percent, number, rate, level, risk, amount, cost, behavior, size, pressure, activity, chance, likelihood, population, body **misc** significantly, during, actually, slightly, dramatically

WORD: fall

LIST FROM #: (1-60,000)

PART OF SPEECH: NOUN VERB ADJ ADV MISC

[6 / 11] HIDE HELP

The **collocates** show you the nearby words (shown here for *moonlight*), and these provide useful information on the meaning and use of a word. They are grouped by part of speech and then listed in order (most frequent first). Note that any collocates in parentheses (mainly for very low frequency words) are tentative.

SINGLE ENTRY: DISPLAYED IN FRAME BELOW

RANK #	POS	WORD	TOTAL	SPOKEN	FICTION	MAGAZINE	NEWSPAPER	ACADEMIC
410	V	FALL	105127	13462	38151	21248	20176	12090

SYNONYMS (click to see) [?] SEE ALSO: MORE SPECIFIC (67) MORE GENERAL (44) MULTI-WORD (29)

decrease

701 reduce

FALL v (RANK 410, FREQ 105127)

	SPOKEN	FICTION	MAGAZINE	NEWSPAPER	ACADEMIC
CLICK BAR TO LIMIT					
STORED	28	76	36	40	22
MORE	13462	38151	21248	20176	12090

DEFINITIONS (WORDNET) (BAD ENTRY?)

(S/32) 1. descend in free fall under the influence of gravity 2. move downward and lower, but not necessarily all the way 3. pass suddenly and passively into a state of body or mind 4. come under, be classified or included 5. decrease in size, extent, or range

COLLOCATES (click to see with FALL)

noun love, percent, price, category, snow, rain, ground, floor, rate, hair, face, tree, stock, knee, victim **misc** into, asleep, apart, short, behind, below

off horse, cliff, chair, bike, wheel, ladder, sharply, hat **down** dead, well, drunk, bridge, onto, pant, tear **back** into, onto, head,

WORD: fluctuate

LIST FROM #: (1-60,000)

PART OF SPEECH: NOUN VERB ADJ ADV MISC

[1 / 11] HIDE HELP

At the most basic level, you can search for a specific word (e.g. *stream, block, smooth, skew, jolt, plush, inveigh, chicanery, addled*). You can also search for words that match a certain pattern, e.g. words starting with *soft** or *wind**, words ending in **ism* or **ship*, words with the root **back** or **heart**, words with the pattern *J-g**, or *hyphenated* words). You can also see a randomly-selected word by clicking on

SINGLE ENTRY: DISPLAYED IN FRAME BELOW

RANK #	POS	WORD	TOTAL	SPOKEN	FICTION	MAGAZINE	NEWSPAPER	ACADEMIC
13061	V	FLUCTUATE	1032	99	58	278	164	433

SYNONYMS (click to see) [?] SEE ALSO: MORE GENERAL (4)

vary

2386 vary
2469 swing
2906 alter

FLUCTUATE v (RANK 13061, FREQ 1032)

	SPOKEN	FICTION	MAGAZINE	NEWSPAPER	ACADEMIC
CLICK BAR TO LIMIT					
STORED	19	7	62	34	80
MORE	99	58	278	164	433

DEFINITIONS (WORDNET) (BAD ENTRY?)

1. be unstable 2. move or sway in a rising and falling or wavelike pattern 3. cause to fluctuate or move in a wave-like pattern

COLLOCATES (click to see with FLUCTUATE)

noun price, rate, year, level, temperature, number, market, weight, population, interest, value, stock, season, month, period **misc** between, wildly, over, widely, according, during, greatly, allow, depending, around, dramatically, throughout, tend, naturally, whose

WORD: fluctuation

LIST FROM #: (1-60,000)

PART OF SPEECH: NOUN VERB ADJ ADV MISC

[6 / 11] HIDE HELP

The **collocates** show you the nearby words (shown here for *moonlight*), and these provide useful information on the meaning and use of a word. They are grouped by part of speech and then listed in order (most frequent first). Note that any collocates in parentheses (mainly for very low frequency words) are tentative.

SINGLE ENTRY: DISPLAYED IN FRAME BELOW

RANK #	POS	WORD	TOTAL	SPOKEN	FICTION	MAGAZINE	NEWSPAPER	ACADEMIC
8620	N	FLUCTUATION	2219	76	113	634	191	1205

SYNONYMS (click to see) [?] SEE ALSO: MORE SPECIFIC (13) MORE GENERAL (7)

variation

7031 instability
8558 variability
8620 fluctuation

FLUCTUATION n (RANK 8620, FREQ 2219)

	SPOKEN	FICTION	MAGAZINE	NEWSPAPER	ACADEMIC
CLICK BAR TO LIMIT					
STORED	7	15	50	14	116
MORE	76	113	634	191	1205

DEFINITIONS (WORDNET) (BAD ENTRY?)

1. the quality of being unsteady and subject to fluctuations 2. an instance of change 3. a wave motion

COLLOCATES (click to see with FLUCTUATION)

adj natural, small, random, large, short-term, economic, significant, daily, thermal, wild **noun** temperature, price, quantum, market, currency, population, rate, density, level, energy **verb** cause, occur, measure, result, reflect, detect, experience, observe, reveal, correspond

WORD: SEARCH RESET

LIST FROM # (1-60,000)

PART OF SPEECH NOUN VERB ADJ ADV MISC

[6 / 11] HIDE HELP

The collocates show you the nearby words (shown here for moonlight), and these provide useful information on the meaning and use of a word. They are grouped by part of speech and then listed in order (most frequent first). Note that any collocates in parentheses (mainly for very low frequency words) are tentative.

SINGLE ENTRY: DISPLAYED IN FRAME BELOW

RANK #	POS	WORD	TOTAL	SPOKEN	FICTION	MAGAZINE	NEWSPAPER	ACADEMIC
1	N	INCREASE	33639	5178	388	6221	8443	13409

SYNONYMS (click to see) [?]

SEE ALSO:
 MORE SPECIFIC (56)
 MORE GENERAL (5)
 MULTI-WORD (5)

upsurge

1265 increase
1970 rise

INCREASE n (RANK 1265, FREQ 33639)

	SPOKEN	FICTION	MAGAZINE	NEWSPAPER	ACADEMIC
CLICK BAR TO LIMIT					
STORED	24	2	38	47	91
MORE	5178	388	6221	8443	13409

DEFINITIONS (WORDNET) (BAD ENTRY?)

1. a quantity that is added 2. a change resulting in an increase 3. a process of becoming larger or longer or more numerous or more important 4. the act of increasing something 5. the amount by which something increases

COLLOCATES (click to see with INCREASE)

adj significant, large, dramatic, substantial, huge, rapid, sharp, average, slight **noun** percent, tax, rate, price, number, population, cost, wage, spending, production **verb** show, lead, report, result, cause, experience, indicate, occur, vote, observe

WORD: SEARCH RESET

LIST FROM # (1-60,000)

PART OF SPEECH NOUN VERB ADJ ADV MISC

[6 / 11] HIDE HELP

The collocates show you the nearby words (shown here for moonlight), and these provide useful information on the meaning and use of a word. They are grouped by part of speech and then listed in order (most frequent first). Note that any collocates in parentheses (mainly for very low frequency words) are tentative.

SINGLE ENTRY: DISPLAYED IN FRAME BELOW

RANK #	POS	WORD	TOTAL	SPOKEN	FICTION	MAGAZINE	NEWSPAPER	ACADEMIC
1	V	INCREASE	65431	6773	1726	15535	12208	29189

SYNONYMS (click to see) [?]

SEE ALSO:
 MORE SPECIFIC (64)
 MORE GENERAL (2)
 MULTI-WORD (5)

enlarge

433 raise
660 increase

INCREASE v (RANK 660, FREQ 65431)

	SPOKEN	FICTION	MAGAZINE	NEWSPAPER	ACADEMIC
CLICK BAR TO LIMIT					
STORED	19	5	40	33	105
MORE	6773	1726	15535	12208	29189

DEFINITIONS (WORDNET) (BAD ENTRY?)

1. become bigger or greater in amount 2. make bigger or more

COLLOCATES (click to see with INCREASE)

noun number, risk, rate, year, tax, price, cost, amount, production, chance, likelihood, population, pressure, size, power **misc** dramatically, significantly, continue, greatly, substantially

WORD: SEARCH RESET

LIST FROM # (1-60,000)

PART OF SPEECH NOUN VERB ADJ ADV MISC

[2 / 11] HIDE HELP

Unlike other sites, everything here is based on frequency in the corpus. For example, you can find words that have high, medium, and low frequency. Just indicate in the form above what word you want to start with (1-60,000). You can also see a randomly-selected range (between 4,000 and 15,000) by clicking on .

SINGLE ENTRY: DISPLAYED IN FRAME BELOW

RANK #	POS	WORD	TOTAL	SPOKEN	FICTION	MAGAZINE	NEWSPAPER	ACADEMIC
1	J	STABLE	11976	1403	1189	3033	1642	4709

SYNONYMS (click to see) [?]

calm

2931 steady
4769 calm
11815 even
14593 collected
16985 settled

STABLE J (RANK 2920, FREQ 11976)

	SPOKEN	FICTION	MAGAZINE	NEWSPAPER	ACADEMIC
CLICK BAR TO LIMIT					
STORED	25	18	50	28	81
MORE	1403	1189	3033	1642	4709

DEFINITIONS (WORDNET) (BAD ENTRY?)

1. resistant to change of position or condition 2. showing little if any change 3. maintaining equilibrium 4. not taking part readily in chemical change 5. firm and dependable

COLLOCATES (click to see with STABLE)

noun condition, environment, family, government, relationship, price, system, life, democracy, population, rate, economy **misc** more, remain, relatively, less, provide, enough, create, fairy

WORD: SEARCH RESET

LIST FROM # (1-60,000)

PART OF SPEECH NOUN VERB ADJ ADV MISC

[2 / 11] HIDE HELP

Unlike other sites, everything here is based on frequency in the corpus. For example, you can find words that have high, medium, and low frequency. Just indicate in the form above what word you want to start with (1-60,000). You can also see a randomly-selected range (between 4,000 and 15,000) by clicking on .

SINGLE ENTRY: DISPLAYED IN FRAME BELOW

RANK #	POS	WORD	TOTAL	SPOKEN	FICTION	MAGAZINE	NEWSPAPER	ACADEMIC
1	V	REMAIN	98934	10623	12688	22353	22911	30359

SYNONYMS (click to see) [?]

continue

426 remain
3677 endure

REMAIN v (RANK 426, FREQ 98934)

	SPOKEN	FICTION	MAGAZINE	NEWSPAPER	ACADEMIC
CLICK BAR TO LIMIT					
STORED	19	23	53	38	69
MORE	10623	12688	22353	22911	30359

DEFINITIONS (WORDNET) (BAD ENTRY?)

1. stay the same 2. continue in a place, position, or situation 3. be left 4. stay behind

COLLOCATES (click to see with REMAIN)

noun question, mystery, second, challenge, calm, minute, troop, decade, focus, core, doubt, secret, priority, inflation, gap **misc** while, until, same, silent, however, open, constant, unchanged

WORD: SEARCH RESET

LIST FROM #

PART OF SPEECH: NOUN VERB ADJ ADV MISC

[6 / 11] HIDE HELP

The **collocates** show you the nearby words (shown here for *moonlight*), and these provide useful information on the meaning and use of a word. They are grouped by part of speech and then listed in order (most frequent first). Note that any collocates in parentheses (mainly for very low frequency words) are tentative.

SINGLE ENTRY: DISPLAYED IN FRAME BELOW

RANK #	POS	WORD	TOTAL	SPOKEN	FICTION	MAGAZINE	NEWSPAPER	ACADEMIC	
1	1970	N	RISE	19602	2020	2428	5068	3921	6165

SYNONYMS (click to see) [?]

SEE ALSO:
 MORE SPECIFIC (25)
 MORE GENERAL (16)
 MULTI-WORD (3)

climb

6790 elevation
 7998 climb

RISE n (RANK 1970, FREQ 19602)

	SPOKEN	FICTION	MAGAZINE	NEWSPAPER	ACADEMIC
CLICK BAR TO LIMIT					
STORED	20	29	60	37	56
MORE	2020	2428	5068	3921	6165

DEFINITIONS (WORDNET) (BAD ENTRY?)

(5/9) **1.** a growth in strength or number or importance **2.** the act of changing location in an upward direction **3.** the property possessed by a slope or surface that rises **4.** the amount a salary is increased **5.** a movement upward

COLLOCATES (click to see with RISE)

adj rapid, sharp, dramatic, sea-level, meteoric, recent, steady **noun** fall, price, power, level, percent, rate, temperature, sea, crime **verb** give, let, cause, contribute, explain, witness, experience

WORD: SEARCH RESET

LIST FROM #

PART OF SPEECH: NOUN VERB ADJ ADV MISC

[6 / 11] HIDE HELP

The **collocates** show you the nearby words (shown here for *moonlight*), and these provide useful information on the meaning and use of a word. They are grouped by part of speech and then listed in order (most frequent first). Note that any collocates in parentheses (mainly for very low frequency words) are tentative.

SINGLE ENTRY: DISPLAYED IN FRAME BELOW

RANK #	POS	WORD	TOTAL	SPOKEN	FICTION	MAGAZINE	NEWSPAPER	ACADEMIC	
1	727	V	RISE	58380	5946	18695	13348	13067	7324

SYNONYMS (click to see) [?]

SEE ALSO:
 MORE SPECIFIC (36)
 MORE GENERAL (23)
 MULTI-WORD (3)

emerge

727 **rise**
 1304 emerge

RISE v (RANK 727, FREQ 58380)

	SPOKEN	FICTION	MAGAZINE	NEWSPAPER	ACADEMIC
CLICK BAR TO LIMIT					
STORED	19	59	50	54	20
MORE	5946	18695	13348	13067	7324

DEFINITIONS (WORDNET) (BAD ENTRY?)

(5/18) **1.** move upward **2.** increase in value or to a higher point **3.** rise to one's feet **4.** rise up **5.** become more extreme

COLLOCATES (click to see with RISE)

noun price, rate, year, sun, cost, voice, water, stock, temperature, air **misc** from, above, fall, high, continue, slowly **up** from, against, again, smoke, voice, protest, anger, overthrow, revolt, rebellion

Appendix D
Tests of Normality

Tests of Normality for Paired-Samples T-Tests

First, Tables D1 and D2 below display the distribution of the difference in the scores for the frequency of formulaic sequences in the pretest and posttest, pretest and delayed posttest, and posttest and delayed posttest.

Table D1				
<i>Descriptive Statistics: The Mean Difference of the Frequency of the Target Formulaic Sequences</i>				
		Statistic	Std. Error	
Frequency Difference: Pretest and Posttest	Mean	-13.2500	1.90345	
	95% Confidence Interval for Mean	Lower Bound	-17.4395	
		Upper Bound	-9.0605	
	5% Trimmed Mean	-13.6111		
	Median	-13.5000		
	Variance	43.477		
	Std. Deviation	6.59373		
	Minimum	-23.00		
	Maximum	-3.00		
	Range	26.00		
	Interquartile Range	5.75		
	Skewness	1.110	.637	
	Kurtosis	2.888	1.232	
Frequency Difference: Pretest and Delayed Posttest	Mean	-12.4167	1.92456	
	95% Confidence Interval for Mean	Lower Bound	-16.6526	
		Upper Bound	-8.1807	
	5% Trimmed Mean	-12.5185		
	Median	-13.0000		
	Variance	44.447		

	Std. Deviation		6.66686	
	Minimum		-26.00	
	Maximum		3.00	
	Range		29.00	
	Interquartile Range		3.75	
	Skewness		.336	.637
	Kurtosis		3.347	1.232
Frequency Difference: Posttest and Delayed Posttest	Mean		.8333	1.26031
	95% Confidence Interval for Mean	Lower Bound	-1.9406	
		Upper Bound	3.6073	
	5% Trimmed Mean		.5370	
	Median		-.5000	
	Variance		19.061	
	Std. Deviation		4.36585	
	Minimum		-4.00	
	Maximum		11.00	
	Range		15.00	
	Interquartile Range		6.50	
	Skewness		1.165	.637
	Kurtosis		1.270	1.232

Table D2

Tests of Normality: The Mean Difference of the Frequency of Formulaic Sequences

	Shapiro-Wilk		
	Statistic	df	Sig.
Pretest and Posttest	.904	12	.179
Pretest and Delayed Posttest	.857	12	.045
Posttest and Delayed Posttest	.896	12	.142

* $p > .05$

According to Tables D1 and D2, the dependent variable representing the difference in the scores for the frequency of the target formulaic sequences in the pretest and the posttest is approximately normally distributed ($M = -13.25$, $SD = 6.59$, $W = .17$, $n = 12$). Although Table D1 demonstrates that the dependent variable is a little skewed and kurtotic, with a skewness of 1.10 ($SE = .63$, z -value = 2.33) and Kurtosis of 2.88 ($SE = 1.23$, z -value = 2.33), the null hypothesis is not rejected due to the fact that $p > .05$. The two tables above also shows that the difference in the frequency of formulaic sequences in the pretest and the delayed posttest deviates from normality ($M = -12.41$, $SD = 6.66$, $W = .045$, $n = 12$). The detected departure of this variable from normality is due to kurtosis = 3.34 ($SE = 1.23$, z -value = 2.71) rather than skewness = .33 ($SE = .63$, z -value = .52), resulting in the rejection of the null hypothesis the data are normally distributed. However, since paired-samples t-tests are robust to violations of normality in that their results are powerful in the presence of slight normality violation (Lumley et al., 2002), a paired-samples t-test was computed between the frequency of the target formulaic sequences in the pretest and those in the delayed posttest to highlight any significant differences between the two. Furthermore, the third dependent variable, representing the difference between the frequency of formulaic sequences in the posttest and the delayed posttest, is approximately normally distributed ($M = .83$, $SD = 4.36$, $W = .14$, $n = 12$) with a skewness of 1.16 ($SE = .63$, z -value = 1.82) and kurtosis of 1.27 ($SE = 1.23$, z -value = 1.03).

Second, the difference in the scores for the occurrence of the target formulaic sequences were tested statistically for their distribution.

Table D3

Descriptive Statistics: The Mean Difference of the Occurrence of the Target Formulaic Sequences

		Statistic	Std. Error	
Occurrence Difference: Pretest and Posttest	Mean	-9.1667	.86017	
	95% Confidence Interval for Mean	Lower Bound	-11.0599	
		Upper Bound	-7.2734	
	5% Trimmed Mean	-9.2963		
	Median	-8.5000		
	Variance	8.879		
	Std. Deviation	2.97973		
	Minimum	-13.00		
	Maximum	-3.00		
	Range	10.00		
	Interquartile Range	4.75		
	Skewness	.416	.637	
	Kurtosis	.078	1.232	
Occurrence Difference: Pretest and Delayed Posttest	Mean	-8.7500	.69767	
	95% Confidence Interval for Mean	Lower Bound	-10.2856	
		Upper Bound	-7.2144	
	5% Trimmed Mean	-8.8889		
	Median	-9.5000		
	Variance	5.841		
	Std. Deviation	2.41680		
	Minimum	-12.00		
	Maximum	-3.00		
	Range	9.00		
	Interquartile Range	3.00		
	Skewness	1.179	.637	
	Kurtosis	1.824	1.232	
Occurrence Difference:	Mean	.4167	.62107	

Posttest and Delayed Posttest	95% Confidence Interval for Mean	Lower Bound	-.9503	
		Upper Bound	1.7836	
	5% Trimmed Mean		.3519	
	Median		.5000	
	Variance		4.629	
	Std. Deviation		2.15146	
	Minimum		-3.00	
	Maximum		5.00	
	Range		8.00	
	Interquartile Range		2.00	
	Skewness		.597	.637
	Kurtosis		.856	1.232

Table D4

Tests of Normality: The Mean Difference of the Occurrence of the Target Formulaic Sequences

	Shapiro-Wilk		
	Statistic	df	Sig.
Pretest and Posttest	.923	12	.309
Pretest and Delayed Posttest	.890	12	.116
Posttest and Delayed Posttest	.949	12	.625

* $p > .05$

Based on Tables D3 and D4, the dependent variable representing the difference in the scores for the occurrence of the target formulaic sequences in the pretest and posttest of the study sample is approximately normally distributed ($M = -9.16$, $SD = 2.97$, $W = .30$, $n = 12$). However, the variable is a little skewed and kurtotic, skewness = .41 ($SE = .63$, z -value = .65) and kurtosis = .07 ($SE = 1.23$, z -value = .63). Similarly, the difference in the scores for the occurrence of formulaic sequences in the pretest and delayed posttest is approximately normally distributed ($M = -8.75$, $SD = 2.41$, $W = .11$, $n = 12$), with a skewness of 1.79 ($SE = .63$, z -value =

1.48) and a kurtosis of 1.82 ($SE = 1.23$, z -value = 1.48). The difference in the scores for the posttest and the delayed posttest is as well approximately normally distributed ($M = .41$, $SD = 2.15$, $W = .62$, $n = 12$) with a skewness of .59 ($SE = .63$, z -value = .93) and kurtosis of .85 ($SE = 1.23$, z -value = .69). The presence of the z -values within ± 1.96 indicate that the three above-listed variables do not differ significantly from normality; moreover, the null hypothesis is not rejected since $p > .05$.

Third, the difference in the score for the first judge's evaluation were tested statistically for their distribution.

		Statistic	Std. Error	
Evaluation Difference J1: Pretest and Posttest	Mean	-1.1667	.24873	
	95% Confidence Interval for Mean	Lower Bound	-1.7141	
		Upper Bound	-.6192	
	5% Trimmed Mean	-1.1852		
	Median	-1.2500		
	Variance	.742		
	Std. Deviation	.86164		
	Minimum	-2.50		
	Maximum	0.50		
	Range	3.00		
	Interquartile Range	1.25		
	Skewness	.493	.637	
Kurtosis	-.051	1.232		
Evaluation Difference J1: Pretest and Delayed Posttest	Mean	-1.3750	.16428	
	95% Confidence Interval for Mean	Lower Bound	-1.7366	
		Upper Bound	-1.0134	

	5% Trimmed Mean		-1.4167	
	Median		-1.5000	
	Variance		.324	
	Std. Deviation		.56909	
	Minimum		-2.00	
	Maximum		.00	
	Range		2.00	
	Interquartile Range		.88	
	Skewness		1.193	.637
	Kurtosis		2.115	1.232
Evaluation Difference J1: Posttest and Delayed Posttest	Mean		-.2083	.25716
	95% Confidence Interval for Mean	Lower Bound	-.7743	
		Upper Bound	.3577	
	5% Trimmed Mean		-.2037	
	Median		.0000	
	Variance		.794	
	Std. Deviation		.89082	
	Minimum		-1.50	
	Maximum		1.00	
	Range		2.50	
	Interquartile Range		1.50	
	Skewness		-.164	.637
	Kurtosis		-1.248	1.232

	Shapiro-Wilk		
	Statistic	df	Sig.
Pretest and Posttest	.960	12	.785
Pretest and Delayed Posttest	.851	12	.038
Posttest and Delayed Posttest	.922	12	.299

Based on Tables D5 and D6, the difference in the scores for the first judge's evaluation of the pretest and posttest is approximately normally distributed ($M = -1.16$, $SD = .86$, $W = .78$, $n = 12$), with a skewness of .49 ($SE = .63$, $z\text{-value} = .77$) and a kurtosis of -.51 ($SE = 1.23$, $z\text{-value} = -.41$). Thus, the null hypothesis, i.e., the data are normally distributed, is confirmed. The second dependent variable, which represents the difference in the scores for the evaluation of the pretest and the delayed posttest, is not normally distributed ($M = -1.37$, $SD = .56$, $W = .03$, $n = 12$) due to the fact that $p < .05$, and, thus, the null hypothesis is rejected. However, the z -values for the skewness = 1.19 ($SE = .63$, $z\text{-value} = 1.88$) and kurtosis = 2.11 ($SE = 1.23$, $z\text{-value} = 1.71$) of that variable do not depart significantly from normality as they exist within +/- 1.96. It is worth noting here that despite the deviation of the second dependent variable from normality, a paired-samples t -test was computed on this pair because, as mentioned earlier, the results of t -tests are powerful in the presence of a slight normality violation (Lumley et al., 2002). The third dependent variable, representing the difference in the evaluation scores for the posttest and the delayed posttest, is approximately normally distributed ($M = -.20$, $SD = .89$, $W = .29$, $n = 12$). Moreover, the z -values for the skewness and kurtosis of this variable are within +/- 1.96:

skewness = $-.16$ ($SE = .63$, z -value = $-.25$) and kurtosis = -1.24 ($SE = 1.23$, z -value = -1). Thus, the null hypothesis, which suggests that the data are normally distributed, is not rejected.

Fourth, the difference in the scores for the evaluation of the second judge were tested statistically to investigate their distribution.

		Statistic	Std. Error	
Evaluation Difference J2: Pretest and Posttest	Mean	-1.5833	.44735	
	95% Confidence Interval for Mean	Lower Bound	-2.5680	
		Upper Bound	-.5987	
	5% Trimmed Mean	-1.5648		
	Median	-2.0000		
	Variance	2.402		
	Std. Deviation	1.54968		
	Minimum	-4.50		
	Maximum	1.00		
	Range	5.50		
	Interquartile Range	2.38		
	Skewness	.121	.637	
	Kurtosis	-.054	1.232	
Evaluation Difference J2: Pretest and Delayed Posttest	Mean	-1.5417	.25716	
	95% Confidence Interval for Mean	Lower Bound	-2.1077	
		Upper Bound	-.9757	
	5% Trimmed Mean	-1.5185		
	Median	-1.5000		
	Variance	.794		
	Std. Deviation	.89082		
Minimum	-3.50			

	Maximum		.00	
	Range		3.50	
	Interquartile Range		1.00	
	Skewness		-.429	.637
	Kurtosis		1.442	1.232
Evaluation Difference J2: Posttest and Delayed Posttest	Mean		.0417	.34520
	95% Confidence Interval for Mean	Lower Bound	-.7181	
		Upper Bound	.8014	
	5% Trimmed Mean		.0463	
	Median		.2500	
	Variance		1.430	
	Std. Deviation		1.19579	
	Minimum		-2.00	
	Maximum		2.00	
	Range		4.00	
	Interquartile Range		2.00	
	Skewness		-.261	.637
	Kurtosis		-.742	1.232

Table D8

Tests of Normality: The Mean Difference of the Second Judge's Evaluation

	Shapiro-Wilk		
	Statistic	df	Sig.
Pretest and Posttest	.914	12	.237
Pretest and Delayed Posttest	.933	12	.418
Posttest and Delayed Posttest	.952	12	.664

* $p > .05$

According to Tables D7 and D8, the first dependent variable extracted from the differences in the scores for the second judge's evaluation of the pretest and posttest is approximately normally

distributed ($M = -1.58$, $SD = 1.54$, $W = .23$, $n = 12$), with a skewness of .12 ($SE = .63$, z -value = .19) and a kurtosis of -.54 ($SE = 1.23$, z -value = -.43). The second dependent variable, representing the difference in the scores for the pretest and delayed posttest evaluation, is as well approximately normally distributed ($M = -1.54$, $SD = .89$, $W = .41$, $n = 12$), with a skewness of -.42 ($SE = .63$, z -value = -.66) and a kurtosis of 1.44 ($SE = 1.23$, z -value = 1.17). Finally, the difference in the scores for the posttest and the delayed posttest evaluation, although a little skewed (Skewness = -.26, $SE = .63$, z -value = -.41) and kurtotic (kurtosis = -.74, $SE = 1.23$, z -value = -.60), is also approximately normally distributed ($M = .41$, $SD = 1.19$, $W = .66$, $n = 12$). In this respect, since the skewness and kurtosis z -values are within +/- 1.96 and $p > .05$, the null hypothesis (the data are normally distributed) is not rejected for the three dependent variables.

Fifth, the distribution of the difference in the scores for the third judge's evaluation was tested statistically.

		Statistic	Std. Error	
Evaluation Difference J3: Pretest and Posttest	Mean	-.7083	.09650	
	95% Confidence Interval for Mean	Lower Bound	-.9207	
		Upper Bound	-.4959	
	5% Trimmed Mean	-.6759		
	Median	-.5000		
	Variance	.112		
	Std. Deviation	.33428		
	Minimum	-1.50		
	Maximum	-.50		
	Range	1.00		
Interquartile Range	.50			

	Skewness		-1.455	.637
	Kurtosis		1.388	1.232
Evaluation Difference J3: Pretest and Delayed Posttest	Mean		-.6667	.12812
	95% Confidence Interval for Mean	Lower Bound	-.9487	
		Upper Bound	-.3847	
	5% Trimmed Mean		-.6574	
	Median		-.5000	
	Variance		.197	
	Std. Deviation		.44381	
	Minimum		-1.50	
	Maximum		.00	
	Range		1.50	
	Interquartile Range		.50	
	Skewness		-.139	.637
	Kurtosis		-.254	1.232
	Evaluation Difference J3: Posttest and Delayed Posttest	Mean		.0417
95% Confidence Interval for Mean		Lower Bound	-.2748	
		Upper Bound	.3581	
5% Trimmed Mean			.0185	
Median			.0000	
Variance			.248	
Std. Deviation			.49810	
Minimum			-.50	
Maximum			1.00	
Range			1.50	
Interquartile Range			1.00	
Skewness			.470	.637
Kurtosis			-.654	1.232

	Shapiro-Wilk		
	Statistic	df	Sig.
Pretest and Posttest	.674	12	.000
Pretest and Delayed Posttest	.900	12	.160
Posttest and Delayed Posttest	.877	12	.080

The difference in the scores for the pretest and posttest evaluation ($M = -.70$, $SD = .33$, $W = .0$, $n = 12$), as shown in Table D10, is not normally distributed since $p < .05$, and, hence, the null hypothesis, suggesting that the data are normally distributed, is rejected. This scores are as well skewed and kurtotic: skewness = -1.45 ($SE = .63$, z -value = -2.30) and kurtosis = 1.38 ($SE = 1.23$, z -value = 1.12), as demonstrated in Table D9. The second dependent variable, representing the differences in the scores for the pretest and the delayed posttest, is approximately normally distributed ($M = -.66$, $SD = .44$, $W = .16$, $n = 12$), and, hence, the null hypothesis is not rejected. Moreover, although the scores are a little skewed (skewness = $-.13$, $SE = .63$, z -value = $-.20$) and kurtotic (kurtosis = $-.25$, $SE = 1.23$, z -value = $-.20$), they do not depart significantly from normality. The third dependent variable is also approximately normally distributed ($M = .04$, $SD = .49$, $W = .08$, $n = 12$), with a skewness of $.47$ ($SE = .63$, z -value = $.74$) and a kurtosis of $-.65$ ($SE = 1.23$, z -value = $-.52$). Accordingly, the null hypothesis for the third dependent variable is not rejected. As mentioned earlier, despite the fact that the scores for the pretest and the posttest evaluation deviate from normality, a paired samples t-test was computed because the results of paired-samples t-test are robust to normality violation (Lumley et al., 2002).

Tests of Normality for Correlation Coefficients

First, the distribution of the dependent variable, i.e., the first judge's evaluation, was explored in a histogram, and descriptive statistics were run to investigate the mean, median, mode, and standard deviation of the first dependent variable.

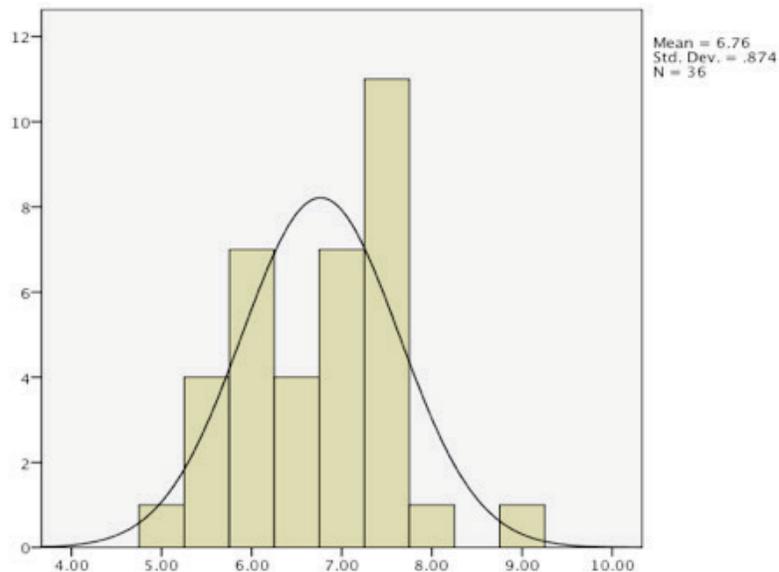


Figure D1. The distribution of the first judge's Evaluation

Table D11

Descriptive Statistics: The First Judge's Evaluation

N	Valid	36
	Missing	0
Mean		6.7639
Median		7.0000
Mode		7.50
Std. Deviation		.87412

Based on Dornyi's (2007) criteria for data that are normally distributed, i.e., "the greatest frequency of scores in the middle with smaller frequencies toward the extremes" (p. 208), the scores for the first judge's evaluation, as presented in Figure D1, are approximately normally

distributed. Moreover, although Figure D1 shows that the data are only approximately normally distributed, the mean = 6.76, the median = 7, the mode = 7.5, $SD = .87$, $n = 36$, as illustrated in Table D11, a Pearson product-moment correlation coefficient was computed since, as Dorny i (2007) points out, “normality does not have to be perfect because most procedures work well with data that is only approximately normally distributed” (p. 208).

Second, the distribution of the scores for the second judge’s evaluation was tested graphically and statistically.

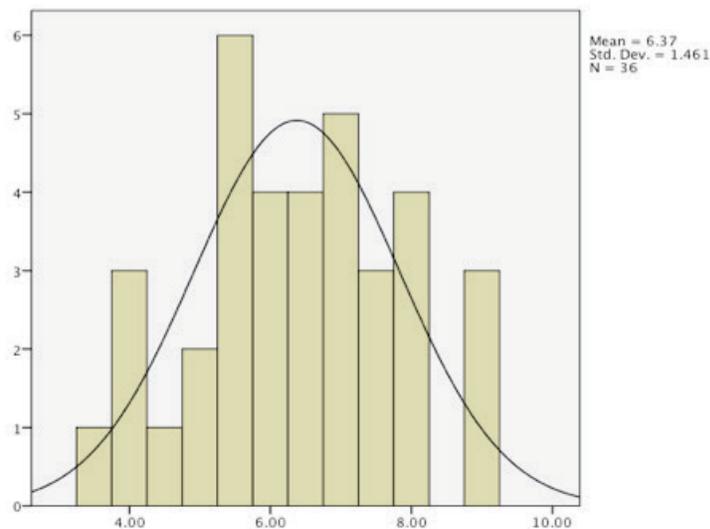


Figure D2. The distribution of the second judge's Evaluation

Table D12		
<i>Descriptive Statistics: The Second Judge's Evaluation</i>		
N	Valid	36
	Missing	0
Mean		6.3750
Median		6.5000
Mode		5.50
Std. Deviation		1.46080

A visual inspection of Figure D2 shows that the second dependent variable is approximately normally distributed the mean = 6.37, the median = 6.5, the mode = 5.5, $SD = 1.46$, $n = 36$.

Finally the distribution of the scores for the third judge's evaluation was assessed both graphically and statistically, as shown in Table D13 and Figure D3 below.

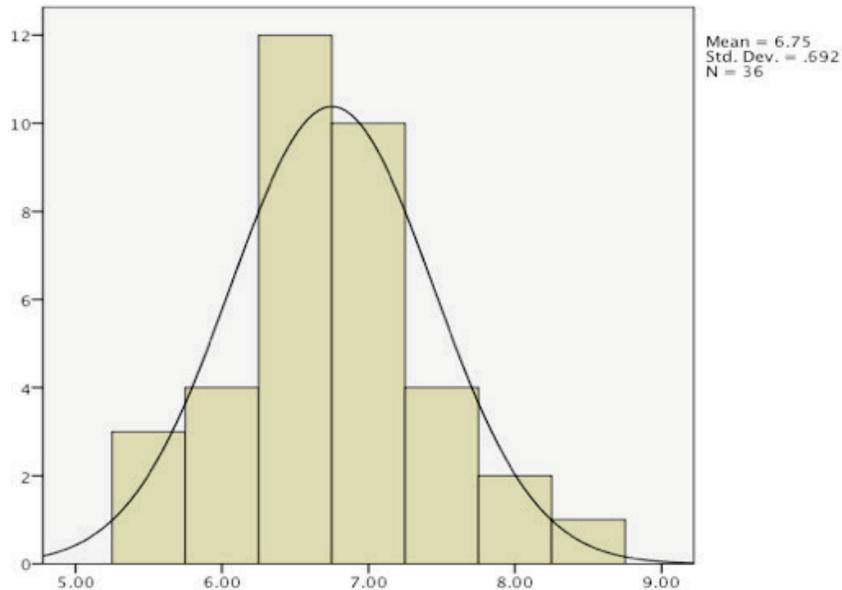


Figure D3. The distribution of the third judge's evaluation

Table D13

Descriptive Statistics: The Third Judge's Evaluation

N	Valid	36
	Missing	0
Mean		6.7500
Median		6.5000
Mode		6.50
Std. Deviation		0.69179

Based on Figure D3, the third dependent variable, which represents the scores for the third judge's evaluation, is approximately normally distributed, the mean = 6.75, the median = 6.5, the mode = 6.5, $SD = .69$, $n = 36$.

Appendix E

Quantitative Content Analysis of the Elicited Texts

Salem's Textual Data

Salem's Pretest

It show how are some materials change throw around one year. For example, sugars price was 215 in May 2010 and it encrease in December to become 400 then they decrease again in Aprial to become 300. On another hand cereals and oils were around 150-170 in May 2010 to become around 220-250 and they stay in similar price in the April 2011, Dairy dosn't have big chang price between May 2010 to April 2011 like meat also. In fact, when the price became higher thats give pad effect. For example. The poor people will be increase, and many people will do criminal thing like theft.

Table E1		
<i>Manual Coding: Salem's Pretest</i>		
Sentence #	Formulaic Sequence	Function
Sentence 2	For example	Using a discourse organizer to exemplify
Sentence 3	On the other hand	Using a discourse organizer inaccurately to contrast elements in the text
Sentence 3	Cereals and oils	Using a sentence builder to connect and compare discourse elements
Sentence 3	Between May 2010 and April 2011	Using a formula inaccurately to introduce a period of time
Sentence 4	For example	Using a discourse organizer to exemplify

Salem's Posttest

The line graph presents the rise and fall in the general price of commodity between May 2010 and April 2011. According to the graph, both dairy and meat remain the same, but sugar dramatic increase until December 2010 then becomes slightly decreased. In addition to the graph, both cereals and oils are generally increase. In conclusion, it can be seen that the commodity prices have generally increased wich leads to a decline in the purchasing power of

the consumers.

Table E2		
<i>Manual Coding: Salem's Posttest</i>		
Sentence #	Formulaic Sequence	Function
Sentence 1	The line graph presents	Using a sentence head to demonstrate that something exists
Sentence 1	the rise and fall	Using a collocation to introduce the chart trends
Sentence 1	the general price of commodity	Using a tangible framing attribute to indicate the characteristics of the following noun phrase
Sentence 1	Between May 2010 and April 2011	Using a formula to introduce a period of time
Sentence 2	According to the	Using a three-word bundle to introduce information presented by elements in the text
Sentence 2	both dairy and meat	Using a sentence builder to connect and compare discourse elements
Sentence 2	remain the same	Using a collocation to introduce the chart trends
Sentence 2	dramatic increase	Using a collocation to introduce the chart trends
Sentence 2	slightly decreased	Using a collocation to introduce the chart trends
Sentence 3	In addition to the	Using a transition signal to organize the text and add new information
Sentence 3	both cereals and oils	Using a sentence builder to connect and compare discourse elements
Sentence 4	In conclusion	Using a summarizer to indicate the end of the text.
Sentence 4	it can be seen that	Using a sentence stem to structure the text meaning and engage the reader in the writer's evaluation of the information presented in the text
Sentence 4	the purchasing power of the consumers	Using a tangible framing attribute to indicate the characteristics of the following noun phrase

Salem's Delayed Posttest

The graph shows how the commodity prices change over a one year. According to the graph, the biggest change happened in the price of sugar, which increased dramatically between May 2010 to December 2010, then it decreased slightly between January 2011 to April 2011.

In contrast, Both Dairy and meat remain the same. In addition, Both oils and cereals increased gradually. In conclusion, it is clear from the information of graph that price of commodity have generally increased between May 2010 to April 2011.

Table E3		
<i>Manual Coding: Salem's Delayed Posttest</i>		
Sentence #	Formulaic Sequence	Function
Sentence 1	The line graph shows	Using a sentence head to demonstrate that something exists
Sentence 1	over a one year	Using a tangible framing attribute to refer to the measurable attributes of the following noun phrase
Sentence 2	According to the	Using a three-word bundle to introduce information presented by elements in the text
Sentence 2	the price of sugar	Using a tangible framing attribute to indicate the characteristics of the following noun phrase
Sentence 2	Increased dramatically	Using a collocation to introduce the chart trends
Sentence 2	between May 2010 to December 2010	Using a formula to introduce a period of time
Sentence 2	decreased slightly	Using a collocation to introduce the chart trends
Sentence 2	between January 2011 to April 2011	Using a formula to introduce a period of time
Sentence 3	Both Dairy and meat	Using a sentence builder to connect and compare discourse elements
Sentence 3	remain the same	Using a collocation to introduce the chart trends
Sentence 4	In addition	Using a transition signal to organize the text and add new information
Sentence 4	Both oils and cereals	Using a sentence builder to connect and compare discourse elements
Sentence 4	increased gradually	Using a collocation to introduce the chart trends
Sentence 5	In conclusion	Using a summarizer to indicate the end of the text.
Sentence 5	it is clear	Using a sentence builder to introduce an impersonal evaluation
Sentence 5	the information of graph	Using a tangible framing attribute to indicate the characteristics of the following noun phrase
Sentence 5	price of commodity	Using a tangible framing attribute to indicate the characteristics of the following noun phrase
Sentence 5	between May 2010 to April 2011	Using a formula to introduce a period of time

Ellie's Textual Data

Ellie's Pretest

We see in the graph, there are more changes in commodity prices between 2010-2011. Dairy and meat are still the same price in most times, but in some cases they change. From May 2010 until June, oil's price was in 170, but after June it went up immediately and still increasing until December, and after that price was up and down. It reached 230 in 2011 April. Also, cereals' price was very low in 2010 May, but went up rapidly to 240 in December. Cereals price in April 2011 was 240. The most crazy increasing price was in sugar. It started from 220 in May 2010 and reached 400 at December, and that's why the people consume sugar a lot. In my opinion, price changes affect consumers and make their lives complicated and difficult. Finally, we see poor people in roads and public places due to these situations.

Table E4		
<i>Manual Coding: Ellie's Pretest</i>		
Sentence #	Formulaic Sequence	Function
Sentence 1	there are	Using a sentence stem to indicate that something exists
Sentence 1	between 2010-2011	Using a formula to introduce a period of time
Sentence 2	Dairy and meat	Using a sentence builder to connect and compare discourse elements
Sentence 3	From May 2010 until June	Using a formula to introduce a period of time
Sentence 3	up and down.	Using a sentence builder to connect and compare discourse elements
Sentence 7	complicated and difficult.	Using a sentence builder to connect and compare discourse elements

Ellie's Posttest

The graph shows fluctuation in food prices between May 2010 and April 2011. Sugar increased gradually from May 2010 to the end of 2010. It began with 220 but ends in December with 400.

After that, it decreased gradually to 310 by August 2011. Dairy stays at the same price over this period except some falls between May 2010 and September 2011. It increased in October and reached 220 again. Cereals and oils were too similar in differences by this period of time. Cereals prices began with 160 in May 2010. In the other hand, oils price began with 175. Although both reached the peak of increasing in December 2010, Oils price was 270 and cereals price was 230. In addition to the graph, meat price wasn't changed more like other foods in that period of time. In conclusion, food prices rose and fell from May 2010 to April 2011.

Table E5		
<i>Manual Coding: Ellie's Posttest</i>		
Sentence #	Formulaic Sequence	Function
Sentence 1	The graph shows	Using a sentence head to demonstrate that something exists
Sentence 1	fluctuation in food prices	Using a collocation to introduce the chart trends
Sentence 1	between May 2010 and April 2011	Using a formula to introduce a period of time
Sentence 2	increased gradually	Using a collocation to introduce the chart trends
Sentence 2	from May 2010 to the end of 2010	Using a formula to introduce a period of time
Sentence 4	decreased gradually	Using a collocation to introduce the chart trends
Sentence 5	over this period	Using a tangible framing attribute creatively to refer to the measurable attributes of the following noun phrase
Sentence 5	between May 2010 and September 2011	Using a formula to introduce a period of time
Sentence 7	Cereals and oils	Using a sentence builder to connect and compare discourse elements
Sentence 7	this period of time	Using a tangible framing attribute creatively to refer to the measurable attributes of the following noun phrase
Sentence 9	In the other hand	Using a discourse organizer inaccurately to contrast elements in the text
Sentence 10	both	Using a reduced sentence builder to connect and compare discourse elements
Sentence 11	In addition to the	Using a transition signal to organize the text and add new information
Sentence 11	that's period of time	Using a tangible framing attribute creatively to

		refer to the measurable attributes of the following noun phrase (grammatical mistake)
Sentence 12	In conclusion	Using a summarizer to indicate the end of the text
Sentence 12	rose and fell	Using a sentence builder to connect and compare discourse elements
Sentence 12	from May 2010 to April 2011	Using a formula to introduce a period of time

Ellie's Delayed Posttest

The graph shows the fluctuation in commodity prices from May 2011 to April 2011. It can be seen that all prices of foods increased over that period. First, Sugar Price increased from 220 in May 2010 to 400 in December 2010. Then it declined to 370 in February 2011; and it sharply decreased to 300 in April 2011. In spite of sugar, the price of meat increased from 140 to 150 at the end of this period. On the one hand, dairy prices rose and fell evenly. It began with 210 in May 2010, but it became 200 in April 2011. On the other hand, the price of Oils increased gradually to 270 in December, and it became 230 in April. Like Oils, cereals price increased gradually too, but it became 230 in December. To sum up, it has been generally accepted that the prices of food had many changes in that period, and sugar increased rapidly.

Table E6		
<i>Manual Coding: Ellie's Delayed Posttest</i>		
Sentence #	Formulaic Sequence	Function
Sentence 1	The graph shows	Using a sentence head to demonstrate that something exists
Sentence 1	the fluctuation in commodity prices	Using a collocation to introduce the chart trends
Sentence 1	from May 2011 to April 2011	Using a formula to introduce a period of time
Sentence 2	It can be seen	Using a sentence stem to structure the text meaning and engage the reader in the writer's evaluation of the information presented in the text
Sentence 3	from 220 in May 2010 to 400 in December 2010	Using a formula creatively to introduce price range and time period

Sentence 4	sharply decreased	Using a collocation to introduce the chart trends
Sentence 5	the price of meat	Using a tangible framing attribute to indicate the characteristics of the following noun phrase
Sentence 5	from 140 to 150	Using a formula to introduce price range
Sentence 5	the end of this period	Using a tangible framing attribute to indicate the characteristics of the following noun phrase
Sentence 6	On the one hand	Using a discourse organizer to contrast elements in the text
Sentence 6	rose and fell	Using a collocation to introduce the chart trends
Sentence 7	On the other hand	Using a discourse organizer to contrast elements in the text discourse elements
Sentence 7	the price of Oils	Using a tangible framing attribute to indicate the characteristics of the following noun phrase
Sentence 7	increased gradually	Using a collocation to introduce the chart trends
Sentence 9	increased gradually	Using a collocation to introduce the chart trends
Sentence 10	To sum up	Using a summarizer to indicate the end of the text
Sentence 10	It has been generally accepted	Using a sentence head to prime a topic of discussion but, at the same time, keep the writer detached from the argument
Sentence 10	the prices of food	Using a tangible framing attribute to indicate the characteristics of the following noun phrase
Sentence 10	increased rapidly	Using a collocation to introduce the chart trends

David's Textual Data

David's Pretest

The graph depicts the changes in commodity prices over last 12 months. It also, gives the changes in prices certain types of materials such as sugar, cereals, Oils, Dairy and meat respectively. First of all, sugar has been changing differently throughout the year.

It started lower and then it goes up until it reached the highest price among all the materials on December. However, the cereals and oils have been changing slightly over the 12 months. In the other side, Dairy and meat almost stayed in fixed price. In anutshell, Consumers might be affected by this sort of changes especially when it comes to this huge changes in short period of time when the Consumers can not organize their budget and they get confused and that will keep them worrying all the time.

Table E7		
<i>Manual Coding: David's Pretest</i>		
Sentence #	Formulaic Sequence	Function
Sentence 1	The graph depicts	Using a sentence head to demonstrate that something exists
Sentence 1	over last 12 months	Using a tangible framing attribute creatively but inaccurately to refer to the measurable attributes of the following noun phrase
Sentence 2	such as	Using an expository phrase to list examples
Sentence 4	cereals and oils	Using a sentence builder to connect and compare discourse elements
Sentence 4	over the 12 months	Using a tangible framing attribute creatively but inaccurately to refer to the measurable attributes of the following noun phrase
Sentence 5	Dairy and meat	Using a sentence builder to connect and compare discourse elements
Sentence 6	In anutshell	Using an idiomatic multiword sequence to indicate the end of the text (spelling mistake)
Sentence 6	this sort of changes	Using a tangible framing attribute creatively to indicate the characteristics of the following NP
Sentence 6	period of time	Using a tangible framing attribute to indicate the characteristics of the following noun phrase

David's Posttest

The graph **shos** the changes in commodity prices during the last year. According to the graph, the price of sugar has been dramatically increased from **May**, 2010 to December in the same year.

Then it increased slightly in 2011. As shown in the graph, in 2011, oils and cereals had fluctuation in their prices. On the other hand, meat and dairy almost remained the same price which is around 200\$. In conclusion, the prices of these materials changed differently. On the one hand, cereals and oils increased. On the other hand, meat and dairy almost remained the same. Furthermore, it can be seen from the graph that non of the prices of metarials decreased.

Table E8		
<i>Manual Coding: David's Posttest</i>		
Sentence #	Formulaic Sequence	Function
Sentence 1	The graph shos	Using a sentence head to demonstrate that something exists (spelling mistake)
Sentence 2	According to the	Using three-word bundle to introduce information presented by elements in the text
Sentence 2	the price of sugar	Using a tangible framing attribute to indicate the characteristics of the following noun phrase
Sentence 2	dramatically increased	Using a collocation to introduce the chart trends
Sentence 2	from May, 2010 to December in the same year	Using a formula creatively to introduce a period of time (comma splice)
Sentence 3	increased slightly	Using a collocation to introduce the chart trends
Sentence 4	As shown in	Using a text reflexive bundle to organize the text and make a direct reference to elements presented in the text
Sentence 4	oils and cereals	Using a sentence builder to connect and compare discourse elements
Sentence 4	fluctuation in their prices	Using a collocation to introduce the chart trends
Sentence 5	On the other hand	Using a discourse organizer to contrast elements in the text
Sentence 5	meat and dairy	Using a sentence builder to connect and compare discourse elements
Sentence 5	remained the same	Using a collocation to introduce the chart trends
Sentence 6	In conclusion	Using an idiomatic multiword sequence to indicate the end of the text (spelling mistake)

Sentence 6	the prices of these materials	Using a tangible framing attribute to indicate the characteristics of the following noun phrase
Sentence 7	On the one hand	Using a discourse organizer to contrast elements in the text
Sentence 7	cereals and oils	Using a sentence builder to connect and compare discourse elements
Sentence 8	On the other hand	Using a discourse organizer to contrast elements in the text
Sentence 8	meat and dairy	Using a sentence builder to connect and compare discourse elements
Sentence 8	remained the same	Using a collocation to introduce the chart trends
Sentence 9	it can be seen	Using a sentence stem to structure the text meaning and engage the reader in the writer's evaluation of the information presented in the text
Sentence 9	the prices of metarials	Using a tangible framing attribute to indicate the characteristics of the following noun phrase (spelling mistake)

David's Delayed Posttest

The graph shows the changes in food prices over a period of one year 2010-2011. According to the graph, there are five different types of food. First, sugar which increased dramatically from 200\$ in May 2010 and ended up to 310\$ in April 2011. It reached its peak which was 400\$ in December 2010. On the other hand, cereals had a slight increase from 155\$ in May 2010 and ended up to around 250\$ in April 2011. In addition to cereals, oils increased gradually from 150\$ to 250\$. However, the price of dairy remained stable in the range of 200\$. Lastly, Meat also remained the same as well in around 150\$. In a nut shell, it is quite obvious that non of the goods declined. Furthermore, it summarized that sugar, cereals, and oils increased in different levels. In contrast, oils and meat almost remained In the same Price during the year.

Table E9		
<i>Manual Coding: David's Delayed Posttest</i>		
Sentence #	Formulaic Sequence	Function
Sentence 1	The graph shows	Using a sentence head to demonstrate that something exists

Sentence 1	over a period of one year	Using a tangible framing attribute to refer to the measurable attributes of the following noun phrase
Sentence 2	According to the	Using three-word bundle to introduce information presented by elements in the text
Sentence 2	there are	Using a sentence stem to indicate that something exists
Sentence 2	different types of food	Using a tangible framing attribute to indicate the characteristics of the following noun phrase
Sentence 3	increased dramatically	Using a collocation to introduce the chart trends
Sentence 5	On the other hand	Using a discourse organizer to contrast elements in the text
Sentence 5	slight increase	Using a collocation to introduce the chart trends
Sentence 6	In addition to	Using a transition signal to organize the text and add new information
Sentence 6	increased gradually	Using a collocation to introduce the chart trends
Sentence 6	from 150\$ to 250\$	Using a formula to introduce price range
Sentence 7	the price of dairy	Using a tangible framing attribute to indicate the characteristics of the following noun phrase
Sentence 7	remained the same	Using a collocation to introduce the chart trends
Sentence 7	the range of 200\$	Using a tangible framing attribute to indicate the characteristics of the following noun phrase
Sentence 8	remained the same	Using a collocation to introduce the chart trends
Sentence 9	In a nut shell	Using a summarizer to indicate the end of the text (spelling mistake)
Sentence 9	it is quite obvious	Using a sentence builder creatively to introduce an impersonal evaluation
Sentence 11	oils and meat	Using a sentence builder to connect and compare discourse elements
Sentence 11	remained In the same Price	Using a collocation creatively to introduce trends presented in the chart

Mary's Textual Data

Mary's Pretest

The commodity prices of five aliments have been changing during May 2010 to April 2010.

The meat's price have not had a huge increment so it would not affect consumers. Diferent case in the sugar behavior. it increased dramatically, almost two fold in December 2010. The behavior of cereals, oils and dairies have been fluctuating since the beginning of this study. observing a constant increment of cereals and oils, but the dairies a tendency to dwindling.

For costumers buying cereals, oils, dairies, and meat are not a huge difference in price, but For sugars the increment during april 2010 to april 2011 is very important.

Table E10		
<i>Manual Coding: Mary's Pretest</i>		
Sentence #	Formulaic Sequence	Function
Sentence 1	The commodity prices of five aliments	Using a tangible framing attribute to indicate the characteristics of the following noun phrase
Sentence 1	during May 2010 to April 2010	Using a formula inaccurately to introduce a period of time
Sentence 2	increased dramatically	Using a collocation to introduce the chart trends
Sentence 3	The behavior of cereals, oils and dairies	Using a tangible framing attribute creatively to indicate the characteristics of the following noun phrase
Sentence 3	the beginning of this study	Using a tangible framing attribute to indicate the characteristics of the following noun phrase
Sentence 4	cereals and oils	Using a sentence builder to connect and compare discourse elements
Sentence 5	during april 2010 to april 2011	Using a formula inaccurately to introduce a period of time

Mary's Posttest

The line graph represents the fluctuation in the prices of food from May 2010 to April 2011. according to the graph the price of meat has slightly increased during the studied period.

When it comes about cereals and oils, it can be seen that both have experienced a gradual

increase from May to December 2010, after that the prices of cereals and oil have slightly decreased. the sugar's price has the same behavior then oils and cereals in two-fold proportion. Finally, the price of dairies has been stable during the period. as **it has be seen** in the graph for consumers the price of sugar, cereals and oils affect their budget, especially the price of sugar that increase around 100 index. In conclusion, the graph represents the fluctuation of sugar, cereals, oils, dairies and meat during 12 months.

Table E11		
<i>Manual Coding: Mary's Posttest</i>		
Sentence #	Formulaic Sequence	Function
Sentence 1	The line graph represents	Using a sentence head creatively to demonstrate that something exists
Sentence 1	the fluctuation in the prices	Using a collocation to introduce the chart trends
Sentence 1	the prices of food	Using a tangible framing attribute to indicate the characteristics of the following noun phrase
Sentence 1	from May 2010 to April 2011	Using a formula to introduce a period of time
Sentence 2	according to the	Using a three-word bundle to introduce information presented by elements in the text
Sentence 2	the price of meat	Using a tangible framing attribute to indicate the characteristics of the following noun phrase
Sentence 2	slightly increased	Using a collocation to introduce the chart trends
Sentence 3	cereals and oils	Using a sentence builder to connect and compare discourse elements
Sentence 3	it can be seen	Using a sentence stem to structure the text meaning and engage the reader in the writer's evaluation of the information presented in the text
Sentence 3	both	Using a formula creatively as an anaphoric reference to a prior noun phrase
Sentence 3	gradual increase	Using a collocation to introduce the chart trends
Sentence 3	from May to December 2010	Using a formula to introduce a period of time
Sentence 3	the prices of cereals and oil	Using a tangible framing attribute to indicate the characteristics of the following noun phrase
Sentence 3	slightly decreased	Using a collocation to introduce the chart trends
Sentence 4	oils and cereals	Using a sentence builder to connect and compare discourse elements
Sentence 5	the price of dairies	Using a tangible framing attribute to indicate the

		characteristics of the following noun phrase
Sentence 6	it has be seen	Using a sentence stem inaccurately to structure the text meaning and engage the reader in the writer's evaluation of the information presented in the text
Sentence 6	he price of sugar, cereals and oils	Using a tangible framing attribute to indicate the characteristics of the following noun phrase
Sentence 6	the price of sugar	Using a tangible framing attribute to indicate the characteristics of the following noun phrase
Sentence 7	In conclusion	Using a summarizer to indicate the end of the text
Sentence 7	the graph represents	Using a sentence head creatively to demonstrate that something exists
Sentence 7	the fluctuation of sugar, cereals, oils, dairies and meat	Using a tangible framing attribute to indicate the characteristics of the following noun phrase

Mary's Delayed Posttest

The graph represents the fluctuation in the price of cereals, oils, dairies and meat from May 2010 to April 2011. The price of meat has remained the same during that period as well as dairy.

When it comes to the price of cereals, it has increased slightly. It can be seen that sugar **an** oils has increased exponentially until December 2010, however they has decreased after that date. It is noticed that consumers have been affected by the prices of sugar, cereals, and oils. The strong impact was the double increase of sugar prices from May 2010 to December 2010. In conclusion, the graph represents the prices' fluctuation of sugar, cereals, oils, dairies, and meat during one year.

Table E12		
<i>Manual Coding: Mary's Delayed Posttest</i>		
Sentence #	Formulaic Sequence	Function
Sentence 1	The line graph represents	Using a sentence head creatively to demonstrate that something exists
Sentence 1	the fluctuation in the price	Using a collocation to introduce the chart trends
Sentence 1	the price of cereals, oils, dairies and meat	Using a tangible framing attribute to indicate the characteristics of the following noun phrase

Sentence 1	rom May 2010 to April 2011	Using a formula to introduce a period of time
Sentence 2	The price of meat	Using a tangible framing attribute to indicate the characteristics of the following noun phrase
Sentence 2	remained the same	Using a collocation to introduce the chart trends
Sentence 2	as well as	Using a discourse marker inaccurately to connect and signal transitions between the constituents of a phrase
Sentence 3	the price of cereals	Using a tangible framing attribute to indicate the characteristics of the following noun phrase
Sentence 3	increased slightly	Using a collocation to introduce the chart trends
Sentence 4	It can be seen	Using a sentence stem to structure the text meaning and engage the reader in the writer's evaluation of the information presented in the text
Sentence 4	sugar an oils	Using a sentence builder to connect and compare discourse elements (Spelling mistake)
Sentence 4	increased exponentially	Using a collocation to introduce the chart trends
Sentence 5	the prices of sugar, cereals, and oils	Using a tangible framing attribute to indicate the characteristics of the following noun phrase
Sentence 6	the double increase of sugar prices	Using a tangible framing attribute to indicate the characteristics of the following noun phrase
Sentence 6	from May 2010 to December 2010	Using a formula to introduce a period of time
Sentence 7	In conclusion	Using a summarizer to indicate the end of the text
Sentence 7	the graph represents	Using a sentence head creatively to demonstrate that something exists
Sentence 7	the prices' fluctuation of sugar, cereals, oils, dairies, and meat	Using two formulas in one phrase; a collocation to describe trends presented in the chart embedded in a tangible framing attribute to indicate the characteristics of the following noun phrase

Eva's Textual Data

Eva's Pretest

The graph describes the changes in the price of daily food over the last 12 months. Sugar increase to the peak amount on Dec. 2010. Cereals increase gradually and oils begins 170 which is more expensive than cereals ends 225 to a little lower than cereals Dairy starts 210 ends 200 decrease/ go down on Aug 2010 increase December Meat is the most steadily increasly Overall. sugar is the most expensive food for consumers. The meat keeps almost the same for the consumers. The price in dairy go down a little bit over the last 12 months. The price of cereals and oils increase so and respectively. I believe consumers will buy more sugar since sugar is cheaper than before. People will choose oils instead of cereals.

Dairy and meat are almost the same price and will not affect consumers' decision too much.

I also believe that people will still buy the necessary good in daily life no matter how expensive it is.

Table E13		
<i>Manual Coding: Eva's Pretest</i>		
Sentence #	Formulaic Sequence	Function
Sentence 1	The graph describes	Using a sentence head creatively to demonstrate that something exists
Sentence 1	the price of daily food	Using a tangible framing attribute to indicate the characteristics of the following noun phrase
Sentence 1	over the last 12 months	Using a tangible framing attribute creatively to refer to the measurable attributes of the following noun phrase
Sentence 8	over the last 12 months	Using a tangible framing attribute creatively to refer to the measurable attributes of the following noun phrase
Sentence 8	The price of cereals and oils	Using a tangible framing attribute to indicate the characteristics of the following noun phrase
Sentence 11	Dairy and meat	Using a sentence builder to connect and compare discourse elements

Eva's Posttest

The graph shows the rise and fall in the price of food from May 2010 to April 2011. According to the graph, it is generally seen that the price of sugar increased dramatically from May 2010 to December 2010 and decreased to 300\$ from December 2010 to April 2011. At the same time, oils' price started at 170 and reached the peak at December 2011, ending up with 225. In the meanwhile, cereals' price shows a steady increase from May 2010 and fluctuated from December 2010 to April 2011. The price of meat increased a little from May 2010 to January and remains the same from February to April 2011. In conclusion, the product price increased over a year. The price of sugar, cereals and oils reached the highest at December 2010. However, Dairy's price decreased a little while the price of meat increased.

Since these products are necessary for people's lives I think people will buy them as usual.

The difference may be that consumers may buy less of food.

Table E14		
<i>Manual Coding: Eva's Posttest</i>		
Sentence #	Formulaic Sequence	Function
Sentence 1	The graph shows	Using a sentence head creatively to demonstrate that something exists
Sentence 1	the rise and fall	Using a collocation to introduce the chart trends
Sentence 1	the price of food	Using a tangible framing attribute to indicate the characteristics of the following noun phrase
Sentence 1	from May 2010 to April 2011	Using a formula to introduce a period of time
Sentence 2	According to the	Using a three-word bundle to introduce information presented by elements in the text
Sentence 2	it is generally seen that	Using a sentence stem creatively to structure the text meaning and engage the reader in the writer's evaluation of the information presented in the text
Sentence 2	the price of sugar	Using a tangible framing attribute to indicate the characteristics of the following noun phrase
Sentence 2	increased dramatically	Using a collocation to introduce the chart trends
Sentence 2	from May 2010 to	Using a formula to introduce a period of time

	December 2010	
Sentence 2	from December 2010 to April 2011	Using a formula to introduce a period of time
Sentence 3	At the same time	Using a discourse marker to connect constituents and signal transition
Sentence 4	steady increase	Using a collocation to introduce the chart trends
Sentence 4	from December 2010 to April 2011	Using a formula to introduce a period of time
Sentence 5	The price of meat	Using a tangible framing attribute to indicate the characteristics of the following noun phrase
Sentence 5	from May 2010 to January	Using a formula to introduce a period of time
Sentence 5	remains the same	Using a collocation to introduce the chart trends
Sentence 5	from February to April 2011	Using a formula to introduce a period of time
Sentence 6	In conclusion	Using a summarizer to indicate the end of the text
Sentence 6	over a year	Using a tangible framing attribute creatively to refer to the measurable attributes of the following noun phrase
Sentence 7	The price of sugar, cereals and oils	Using a tangible framing attribute to indicate the characteristics of the following noun phrase
Sentence 8	the price of meat	Using a tangible framing attribute to indicate the characteristics of the following noun phrase

Eva's Delayed Posttest

The line graph describes the rise and fall of food price between May 2010 and April 2011.

According to the graph, it is clearly seen that the price of sugar rise dramatically from May 2010 to December 2010, and then it fell down to around 300 in April 2011. Dairy price starts from 210 and then fluctuate over a one year, reaching the same price in April 2011.

Meanwhile, oil's price shows a steady increase until December 2010, declines to 225 in April 2011. In terms of cereals, its price keeps increase over a year, reaching the maximum in December, 2010. The price of meat increase a little bit from May 2010 to February 2011 and then keeps the same. Over all, we can see that food prices changes differently over this year.

I believe that consumers will still buy these basic food since they are necessary for daily life.

The difference may be the numbers they will buy depends on the change of price.

Table E15		
<i>Manual Coding: Eva's Delayed Posttest</i>		
Sentence #	Formulaic Sequence	Function
Sentence 1	The line graph describes	Using a sentence head creatively to demonstrate that something exists
Sentence 1	the rise and fall of food price	Using a collocation and a tangible framing attribute to form construct a noun phrase and introduce the trends presented in the chart
Sentence 1	between May 2010 and April 2011	Using a formula to introduce a period of time
Sentence 2	According to the	Using a three-word bundle to introduce information presented by elements in the text
Sentence 2	it is clearly seen	Using a sentence stem creatively to structure the text meaning and engage the reader in the writer's evaluation of the information presented in the text
Sentence 2	the price of sugar	Using a tangible framing attribute to indicate the characteristics of the following noun phrase
Sentence 2	rise dramatically	Using a collocation to introduce the chart trends
Sentence 2	from May 2010 to December 2010	Using a formula to introduce a period of time
Sentence 3	over a one year	Using a tangible framing attribute creatively but inaccurately to refer to the measurable attributes of the following noun phrase
Sentence 4	steady increase	Using a collocation to introduce the chart trends
Sentence 5	over a year	Using a tangible framing attribute to refer to the measurable attributes of the following noun phrase
Sentence 6	The price of meat	Using a tangible framing attribute to indicate the characteristics of the following noun phrase
Sentence 6	from May 2010 to February 2011	Using a formula to introduce a period of time
Sentence 6	keeps the same	Using a collocation creatively to introduce the chart trends
Sentence 7	over this year	Using a tangible framing attribute creatively to refer to the measurable attributes of the following noun phrase
Sentence 8	the change of price	Using a tangible framing attribute to indicate the characteristics of the following noun phrase

Ali's Textual Data

Ali's Pretest

The graph shows changes of sugar, cereals, oils, Dairy and meat prices over one year. It is clear that sugar has the large change during this period. In May 2010, the price of sugar was about 210 4. In November in the same year. Sugar price reached to peak by 400\$.

However meat prices still in its same level during this period with price between 140 to 200\$. In contrast, oils and cereals were fluctuating during this year between 150 to 250.

In May the price of cereals was 160 but this price keep increasing to reat to around 200\$ in December. In conclusion, the big change in the price over the five goods was in sugar, while meat constant to its level between 140 to 190. Other goods were fluctuated during the period between 160 to 260\$ per one. According to the graph we can see that the price increased during this time almost for all five goods. Undoutably, this increasing has a huge affect of people life. Many people will become unable to buy these necessary goods.

Table E16		
<i>Manual Coding: Ali's Pretest</i>		
Sentence #	Formulaic Sequence	Function
Sentence 1	The graph shows	Using a sentence head creatively to demonstrate that something exists
Sentence 1	changes of sugar, cereals, oils, Dairy and meat prices	Using a tangible framing attribute to indicate the characteristics of the following noun phrase
Sentence 1	over one year	Using a tangible framing attribute creatively to refer to the measurable attributes of the following noun phrase
Sentence 2	It is clear that	Using a sentence builder to introduce an impersonal evaluation
Sentence 3	the price of sugar	Using a tangible framing attribute to indicate the characteristics of the following noun phrase
Sentence 5	between 140 to 200\$	Using a formula inaccurately to introduce price range
Sentence 6	oils and cereals	Using a sentence builder to connect and

		compare discourse elements
Sentence 6	between 150 to 250	Using a formula inaccurately to introduce price range
Sentence 7	the price of cereals	Using a tangible framing attribute to indicate the characteristics of the following noun phrase
Sentence 8	In conclusion	Using a summarizer to indicate the end of the text
Sentence 8	between 140 to 190	Using a formula inaccurately to introduce price range
Sentence 9	between 160 to 260\$	Using a formula inaccurately to introduce price range
Sentence 10	According to the	Using a three-word bundle to introduce information presented by elements in the text

Ali's Posttest

The graph shows the rise and fall of five products over twelve months. From May 2010 to April 2011. According to the graph, the highest change was in the price of sugar. On May 2010, the price of sugar was 210, and the price has sharply increased between May and December to reach a peak of 900 on December. However, there was a dramatic fall in the price of sugar between May and April 2011. On the other hand, the price of cereals and oils fluctuated between 150 and 250 over this period of time. It can be seen that the price of both dairy and meat remain stable over these twelve months. In conclusion, it is clear that the price of sugar, cereals and oils increased while the price of meat and dairy remained at the same level over this period of time. Due to the fact that there are many products that people can not live without such as sugar and oil, the rise of price of these products adversely affect the standard of living of people.

Table E17		
<i>Manual Coding: Ali's Posttest</i>		
Sentence #	Formulaic Sequence	Function
Sentence 1	The graph shows	Using a sentence head creatively to demonstrate that something exists
Sentence 1	the rise and fall	Using a collocation to introduce the chart trends

Sentence 1	over twelve months	Using a tangible framing attribute creatively to refer to the measurable attributes of the following noun phrase
Sentence 1	From May 2010 to April 2011	Using a formula to introduce periods of time
Sentence 2	According to the	Using a three-word bundle to introduce information presented by elements in the text
Sentence 2	the price of sugar	Using a tangible framing attribute to indicate the characteristics of the following noun phrase
Sentence 3	the price of sugar	Using a tangible framing attribute to indicate the characteristics of the following noun phrase
Sentence 3	sharply increased	Using a collocation to introduce the chart trends
Sentence 3	between May and December	Using a formula to introduce periods of time
Sentence 4	there was	Using a sentence stem to indicate that something exists
Sentence 4	dramatic fall	Using a collocation to introduce the chart trends
Sentence 4	the price of sugar	Using a tangible framing attribute to indicate the characteristics of the following noun phrase
Sentence 4	between May and April 2011	Using a formula to introduce periods of time
Sentence 5	On the other hand	Using a discourse organizer to contrast elements in the text
Sentence 5	the price of cereals and oils	Using a tangible framing attribute to indicate the characteristics of the following noun phrase
Sentence 5	the price of cereals and oils fluctuated	Using a collocation to introduce the chart trends
Sentence 5	between 150 and 250	Using a formula creatively to introduce price change
Sentence 5	over this period of time	Using a tangible framing attribute creatively to refer to the measurable attributes of the following noun phrase
Sentence 6	It can be seen that	Using a sentence stem to structure the text meaning and engage the reader
Sentence 6	the price of both dairy and meat	Using a tangible framing attribute to indicate the characteristics of the following noun phrase
Sentence 6	both dairy and meat	Using a sentence builder to connect and compare discourse elements
Sentence 6	remain stable	Using a collocation to introduce trends presented in the chart
Sentence 6	over these twelve months	Using a tangible framing attribute creatively to refer to the measurable attributes of the following noun phrase
Sentence 7	In conclusion	Using a summarizer to indicate the end of the text
Sentence 7	it is clear that	Using a sentence builder to introduce impersonal evaluation

Sentence 7	the price of sugar, cereals and oils	Using a tangible framing attribute to indicate the characteristics of the following noun phrase
Sentence 7	the price of meat and dairy	Using a tangible framing attribute to indicate the characteristics of the following noun phrase
Sentence 7	over this period of time	Using a tangible framing attribute creatively to refer to the measurable attributes of the following noun phrase
Sentence 8	Due to the fact that	Using a stance bundle to introduce a degree of certainty
Sentence 8	there are	Using a sentence stem to indicate that something exists
Sentence 8	such as	Using an expository phrase to list examples
Sentence 8	sugar and oil	Using a sentence builder to connect and compare discourse elements
Sentence 8	the rise of price	Using a tangible framing attribute to indicate the characteristics of the following noun phrase
Sentence 8	the standard of living of people	Using a tangible framing attribute to indicate the characteristics of the following noun phrase

Ali's Delayed Posttest

The graph shows the rise and fall in the price of four products over twelve months. According to the graph, the price of sugar increased dramatically to reach a peak of 400\$ in December 2010 before dropping to 300\$ in April 2011. In addition, the price of meat increased slightly over these months. However, the price of both cereal and oils fluctuated during this period of time. On the other hand, the price of dairy decreased slightly from 210\$ in May 2010 to 200\$ in April 2011. In conclusion, it is clear that the price of sugar, cereal, oils and meat increased with there was a slight decrease in the price of dairy during this period of time.

Table E18		
<i>Manual Coding: Ali's Delayed Posttest</i>		
Sentence #	Formulaic Sequence	Function
Sentence 1	The graph shows	Using a sentence head creatively to demonstrate that something exists
Sentence 1	the rise and fall	Using a collocation to introduce the chart trends
Sentence 1	the price of four products	Using a tangible framing attribute to indicate the characteristics of the following noun phrase

Sentence 1	over twelve months	Using a tangible framing attribute creatively to refer to the measurable attributes of the following noun phrase
Sentence 2	According to the	Using a three-word bundle to introduce information presented by elements in the text
Sentence 2	the price of sugar	Using a tangible framing attribute to indicate the characteristics of the following noun phrase
Sentence 2	increased dramatically	Using a collocation to introduce the chart trends
Sentence 3	In addition	Using a transition signal to organize the text and add new information
Sentence 3	the price of meat	Using a tangible framing attribute to indicate the characteristics of the following noun phrase
Sentence 3	increased slightly	Using a collocation to introduce the chart trends
Sentence 3	over these months	Using a tangible framing attribute creatively to refer to the measurable attributes of the following noun phrase
Sentence 4	the price of both cereal and oils	Using a tangible framing attribute to indicate the characteristics of the following noun phrase
Sentence 4	both cereal and oils	Using a sentence builder to connect and compare discourse element
Sentence 4	the price of both cereal and oils fluctuated	Using a collocation to introduce the chart trends
Sentence 5	On the other hand	Using a discourse organizer to contrast elements in the text
Sentence 5	the price of dairy	Using a tangible framing attribute to indicate the characteristics of the following noun phrase
Sentence 5	from 210\$ in May 2010 to 200\$	Using a formula to introduce price range
Sentence 6	In conclusion	Using a summarizer to indicate the end of the text
Sentence 6	it is clear that	Using a sentence builder to introduce impersonal evaluation
Sentence 6	the price of sugar, cereal, oils and meat	Using a tangible framing attribute to indicate the characteristics of the following noun phrase
Sentence 6	there was	Using a sentence stem to indicate that something exists
Sentence 6	slight decrease	Using a collocation to introduce chart trends
Sentence 6	the price of dairy	Using a tangible framing attribute to indicate the characteristics of the following noun phrase

MATA's Textual Data

MATA's Pretest

The graph shows the changes in commodity prices between 2010 and 2011.

As it has been represented in the graph, Y axis is the differences in the prices during the year and X axis is represented the months. According to the graph there is a quite changes in the prices of sugar where it reached its peak in the end of 2010 by 400 and in May 2010 was in its lowest prices by 230. The graph also show the differences of prices of cereals, Oils, and Dairy where is all of them reached their peak in Dec, 2010. On the other hand it seems that there is no changes in prices. For my perspective, changes in the prices has really affectes in the consumers.

Table E19		
<i>Manual Coding: MATA's Pretest</i>		
Sentence #	Formulaic Sequence	Function
Sentence 1	The graph shows	Using a sentence head creatively to demonstrate that something exists
Sentence 1	between 2010 and 2011	Using a formula to introduce a period of time
Sentence 2	According to the	Using a three-word bundle to introduce information presented by elements in the text
Sentence 2	there is	Using a sentence stem to indicate that something exists
Sentence 2	the prices of sugar	Using a tangible framing attribute to indicate the characteristics of the following noun phrase
Sentence 2	the end of 2010	Using a tangible framing attribute to indicate the characteristics of the following noun phrase
Sentence 3	The graph also show	Using a sentence head creatively but with a grammatical mistake to demonstrate that something exists
Sentence 3	the differences of prices	Using a tangible framing attribute to indicate the characteristics of the following noun phrase
Sentence 4	On the other hand	Using a discourse organizer to contrast elements in the text
Sentence 4	there is	Using a sentence stem to indicate that something exists

MATA's Posttest

The graph shows the difference of food's prices between May 2010 and April 2011. According to the graph, the prices of sugar increased dramatically from May 2010 to Dec 2010 where it reached its peak by 400 \$. In contrast it decreased between Dec 2010 and April 2011. Similarly, the prices of cereals, oils and dairy had dramatic increase from May 2010 to the end of the same year. On the other hand, meat's prices remained the same in the first seven months with just a slight increase from Dec 2010 to April 2011. In conclusion, as it has been shown in the graph that the commodity prices had rise and fall over a period of one year. It has been generally believed that changes in prices have positive and negative effect on people. For example, the increases of the prices effect negatively in people's life. Due to the fact that most people in such poor countries can not afford the money to buy these products which means that the economy income will decrease. In contrast, the decreases of prices have positive impact on people's life as well as the economy income.

Table E20		
<i>Manual Coding: MATA's Posttest</i>		
Sentence #	Formulaic Sequence	Function
Sentence 1	The graph shows	Using a sentence head creatively to demonstrate that something exists
Sentence 1	the difference of food's prices	Using a tangible framing attribute to indicate the characteristics of the following noun phrase
Sentence 1	between May 2010 and April 2011	Using a formula to introduce a period of time
Sentence 2	According to the	Using a three-word bundle to introduce information presented by elements in the text
Sentence 2	the prices of sugar	Using a tangible framing attribute to indicate the characteristics of the following noun phrase
Sentence 2	increased dramatically	Using a collocation to introduce the chart trends
Sentence 2	from May 2010 to Dec 2010	Using a formula to introduce a period of time
Sentence 3	between Dec 2010 and April 2011	Using a formula to introduce a period of time

Sentence 4	the prices of cereals, oils and dairy	Using a tangible framing attribute to indicate the characteristics of the following noun phrase
Sentence 4	dramatic increase	Using a collocation to introduce the chart trends
Sentence 4	May 2010 to the end of the same year	Using a formula to introduce a period of time
Sentence 5	On the other hand	Using a discourse organizer to contrast elements in the text
Sentence 5	remained the same	Using a collocation to introduce the chart trends
Sentence 5	a slight increase	Using a collocation to introduce the chart trends
Sentence 5	from Dec 2010 to April 2011	Using a formula to introduce a period of time
Sentence 6	In conclusion	Using a summarizer to indicate the end of the text
Sentence 6	as it has been shown in the graph	Using a text-reflexive bundle creatively to organize the text and make a direct reference to elements presented in the text (spelling mistake)
Sentence 6	rise and fall	Using a collocation to introduce the chart trends
Sentence 6	over a period of one year	Using a tangible framing attribute to refer to the measurable attributes of the following noun phrase
Sentence 7	It has been generally believed	Using a sentence head to prime a topic of discussion but, at the same time, keep the writer detached from the argument
Sentence 7	positive and negative	Using a sentence builder to connect and compare discourse elements
Sentence 8	For example	Using a discourse organizer to exemplify
Sentence 8	the increases of the prices	Using a tangible framing attribute to indicate the characteristics of the following noun phrase
Sentence 9	Due to the fact that	Using a stance bundle to introduce a degree of certainty
Sentence 10	the decreases of prices	Using a tangible framing attribute to indicate the characteristics of the following noun phrase
Sentence 10	as well as the	Using a discourse marker to connect and signal transitions between the constituents of a phrase

MATA's Delayed Posttest

The graph presents the fluctuation of food prices from May 2010 to April 2011. According to the chart, the prices of sugar increased gradually from May 2010 and the end of the same year, reaching its peak by 400\$. In contrast, it decreased starting from the beginning of 2011 to the fourth month of that year. Similarly, the prices of cereals and oils had a gradual increase in the

same time that the prices of sugar increased with a slight decrease between December 2010 and April 2011. On the other hand, prices of Dairy and meat remained the same in the first five months, with slight increase between Sep 2010 and April 2011. In conclusion, as it is clear from the data given, the prices of commodity had rise and fall over a period of one year. It has been believed that changes in the food prices have negative and positive effects on the consumption of any product. For example, if the prices increases will affect the ability of people to buy these kind of food; therefore it will impact the economy income. On the other hand, if they decrease, it will effects the consumers as well as the economy income.

Table E21		
<i>Manual Coding: MATA's Delayed Posttest</i>		
Sentence #	Formulaic Sequence	Function
Sentence 1	The graph presents	Using a sentence head creatively to demonstrate that something exists
Sentence 1	the fluctuation of food prices	Using two formulas at a time; a collocation to describe trends presented in the chart embedded in a tangible framing attribute to indicate the characteristics of the following noun
Sentence 1	from May 2010 to April 2011	Using a formula to introduce a period of time
Sentence 2	According to the	Using a three-word bundle to introduce information presented by elements in the text
Sentence 2	the prices of sugar	Using a tangible framing attribute to indicate the characteristics of the following noun phrase
Sentence 2	increased gradually	Using a collocation to introduce the chart trends
Sentence 3	from the beginning of 2011 to the fourth month of that year	Using a formula creatively to introduce a period of time
Sentence 4	the prices of cereals and oils	Using a tangible framing attribute to indicate the characteristics of the following noun phrase
Sentence 4	gradual increase	Using a collocation to introduce the chart trends
Sentence 4	the prices of sugar	Using a tangible framing attribute to indicate the characteristics of the following noun phrase
Sentence 4	slight decrease	Using a collocation to introduce the chart trends
Sentence 4	between December 2010 and April 2011	Using a formula to introduce a period of time
Sentence 5	On the other hand	Using a discourse organizer to contrast elements

		in the text
Sentence 5	Dairy and meat	Using a sentence builder to connect and compare discourse elements
Sentence 5	remained the same	Using a collocation to introduce the chart trends
Sentence 5	slight increase	Using a collocation to introduce the chart trends
Sentence 5	between Sep 2010 and April 2011	Using a formula to introduce a period of time
Sentence 6	In conclusion	Using a summarizer to indicate the end of the text
Sentence 6	it is clear	Using a sentence builder to introduce an impersonal evaluation
Sentence 6	the prices of commodity	Using a tangible framing attribute to indicate the characteristics of the following noun phrase
Sentence 6	rise and fall	Using a collocation to introduce the chart trends
Sentence 6	over a period of one year	Using a tangible framing attribute to refer to the measurable attributes of the following noun phrase
Sentence 7	It has been believed	Using a sentence head to prime a topic of discussion but, at the same time, keep the writer detached from the argument
Sentence 7	negative and positive	Using a sentence builder to connect and compare discourse elements
Sentence 7	the consumption of any product	Using a tangible framing attribute to refer to the measurable attributes of the following noun phrase
Sentence 8	For example	Using a discourse organizer to exemplify
Sentence 8	the ability of people	Using a tangible framing attribute to refer to the measurable attributes of the following noun phrase
Sentence 9	On the other hand	Using a discourse organizer to contrast elements in the text
Sentence 9	as well as the	Using a discourse marker to connect and signal transitions between the constituents of a phrase

Jacob's Textual Data

Jacob's Pretest

Commodity prices change from month to another, so we saw difference in prices. I think there is a relationship between change in price and consumers and the clearest way to explain that is when the price at sugar increase in december and it decrease in April. In the same way, Oils goe up in December after it was down in May 2010. What I think in this situation is that if prices of commodity goes up, consumers won't use them very much. The price of the meet keep up from May 2010 to April 2011. In addition, cereals start growing in May from 160 Tan to 260 tan in April 2013 but it suffers from some hesited in June to February. Dairy has different direction up and down between April to February but overall, dairy goes up like other commodity. There are many affect for consumers If prices change. Fore example many people avoid to use high priced commodity.

Table E22		
<i>Manual Coding: Jacob's Pretest</i>		
Sentence #	Formulaic Sequence	Function
Sentence 1	from month to another	Using a formula to introduce a period of time
Sentence 2	between change in price and consumers	Using a formula creatively to construct a noun phrase
Sentence 3	prices of commodity	Using a tangible framing attribute to indicate the characteristics of the following noun phrase
Sentence 4	The price of the meet	Using a tangible framing attribute to indicate the characteristics of the following noun phrase (spelling mistake)
Sentence 4	from May 2010 to April 2011	Using a formula to introduce a period of time
Sentence 5	In addition	Using a transition signal to organize the text and add new information
Sentence 5	from 160 Tan to 260 tan	Using a formula to introduce price range
Sentence 5	in June to February	Using a formula inaccurately to introduce a period of time
Sentence 6	up and down	Using a sentence builder to connect and

		compare discourse elements and introduce trends presented in the chart
Sentence 6	between April to February	Using a formula inaccurately to introduce a period of time
Sentence 7	There are	Using a sentence stem to indicate that something exists
Sentence 7	Fore example	Using a discourse organizer to exemplify (spelling mistake)

Jacob's Posttest

The graph shows the fluctuation in the prices of food between May 2010 and April 2013.

According to the graph, the price of meat dramatically increased over one year. In addition, the price of oil dramatically increased between May 2010 and September 2010 then it has a sharp increase until December 2010. On the other hand, it was decreasing dramatically from December 2010 to April 2011. Cereal has a dramatically increased between May 2010, but on the other hand, it was decreasing from December 2010 to April 2011. In addition to cereal, sugar has dramatical increase between 2010 May and April 2011 but it decreased from December 2010 to April 2011. Based on the graph, the dairy has decreased sharply from May 2010 to August 2010 but it increased dramatically between August 2010 and December 2010 but it decreased again dramatically. In conclusion, this graph has seen the rise and down in the price of food over a period of 12 months. As it has been generally noted the change of price effect of consumers. For instance, the comodities that has high prices, people avoid to buy them but the foods that has lower prices, people go for them so the owners of shopping mall should make resonable prices for their commodity.

Table E23		
<i>Manual Coding: Jacob's Posttest</i>		
Sentence #	Formulaic Sequence	Function
Sentence 1	The graph shows	Using a sentence head creatively to demonstrate

		that something exists
Sentence 1	the fluctuation in the prices	Using a collocation to introduce the chart trends
Sentence 1	the prices of food	Using a tangible framing attribute to indicate the characteristics of the following noun phrase
Sentence 1	between May 2010 and April 2013	Using a formula to introduce a period of time
Sentence 2	According to the	Using a three-word bundle to introduce information presented by elements in the text
Sentence 2	the price of meat	Using a tangible framing attribute to indicate the characteristics of the following noun phrase
Sentence 2	dramatically increased	Using a collocation to introduce the chart trends
Sentence 3	In addition	Using a transition signal to organize the text and add new information
Sentence 3	the price of oil	Using a tangible framing attribute to indicate the characteristics of the following noun phrase
Sentence 3	dramatically increased	Using a collocation to introduce the chart trends
Sentence 3	between May 2010 and September 2010	Using a formula creatively to introduce a period of time
Sentence 3	sharp increase	Using a collocation to introduce the chart trends
Sentence 4	On the other hand	Using a discourse organizer to contrast elements in the text
Sentence 4	decreasing dramatically	Using a collocation to introduce the chart trends
Sentence 4	from December 2010 to April 2011	Using a formula to introduce a period of time
Sentence 5	dramatically increased	Using a collocation to introduce the chart trends
Sentence 5	on the other hand	Using a discourse organizer to contrast elements in the text
Sentence 5	from December 2010 to April 2011	Using a formula to introduce a period of time
Sentence 6	In addition to	Using a transition signal to organize the text and add new information
Sentence 6	dramatical increase	Using a collocation to introduce the chart trends
Sentence 6	between 2010 May and April 2011	Using a formula to introduce a period of time
Sentence 6	from December 2010 to April 2011	Using a formula to introduce a period of time
Sentence 7	Based on the	Using a formula to frame an attribute of the following noun phrase
Sentence 7	decreased sharply	Using a collocation to introduce the chart trends
Sentence 7	from May 2010 to August 2010	Using a formula to introduce a period of time
Sentence 7	increased dramatically	Using a collocation to introduce the chart trends
Sentence 7	between August 2010 and December 2010	Using a formula to introduce a period of time

Sentence 7	decreased again dramatically	Using a collocation to introduce the chart trends
Sentence 8	In conclusion	Using a summarizer to indicate the end of the text
Sentence 8	the rise and down	Using a collocation inaccurately to describe trends presented in the chart (lexical error)
Sentence 8	the price of food	Using a tangible framing attribute to indicate the characteristics of the following noun phrase
Sentence 8	over a period of 12 months	Using a tangible framing attribute to refer to the measurable attributes of the following noun phrase
Sentence 9	it has been generally noted	Using a sentence head to prime a topic of discussion but, at the same time, keep the writer detached from the argument
Sentence 9	the change of price	Using a tangible framing attribute to indicate the characteristics of the following noun phrase
Sentence 9	For instance	Using a formula to list examples

Jacob's Delayed Posttest

The graph shows the fluctuation in the prices of food from May 2010 to April 2011. According to the graph, the prices of meat increased dramatically from May 2010 to April 2011. In addition, the price of dairy decreased sharply between May 2010 and August 2010, but it had a dramatically increase from August 2010 to December 2010. Although the price of dairy decreased over a month, but it increased dramatically again between January and April 2011. In addition to dairy, the price of oils increased slightly from May 2010 to December 2010. However, it had a slightly decrease between December 2010 and April 2011. The price of cereals increased from May 2010 to December 2010. As the price of meat, cereals decrease over a month, but it increased again between January 2011 and April 2011.

The price of sugar had a slightly increase between May 2010 and December 2010 then it decrease from December 2010 to April 2011. In conclusion, it can be seen that, the prices of food have fluctuated from May 2010 to April 2011. It can be noticed, the more the price increase, the less people buy them.

Table E24		
<i>Manual Coding: Jacob's Delayed Posttest</i>		
Sentence #	Formulaic Sequence	Function
Sentence 1	The graph shows	Using a sentence head to demonstrate that something exists
Sentence 1	the fluctuation in the prices	Using a collocation to introduce the chart trends
Sentence 1	the prices of food	Using a tangible framing attribute to indicate the characteristics of the following noun phrase
Sentence 1	from May 2010 to April 2011	Using a formula to introduce a period of time
Sentence 2	According to the	Using a three-word bundle to introduce information presented by elements in the text
Sentence 2	the prices of meat	Using a tangible framing attribute to indicate the characteristics of the following noun phrase
Sentence 2	increased dramatically	Using a collocation to introduce the chart trends
Sentence 2	from May 2010 to April 2011	Using a formula to introduce a period of time
Sentence 3	In addition	Using a transition signal to organize the text and add new information
Sentence 3	the price of dairy	Using a tangible framing attribute to indicate the characteristics of the following noun phrase
Sentence 3	decreased sharply	Using a collocation to introduce the chart trends
Sentence 3	between May 2010 and August 2010	Using a formula to introduce a period of time
Sentence 3	dramatically increase	Using a collocation to introduce the chart trends
Sentence 3	from August 2010 to December 2010	Using a formula to introduce a period of time
Sentence 4	the price of dairy	Using a tangible framing attribute to indicate the characteristics of the following noun phrase
Sentence 4	increased dramatically	Using a collocation to introduce the chart trends
Sentence 4	between January and April 2011	Using a formula to introduce a period of time
Sentence 5	In addition to	Using a transition signal to organize the text and add new information
Sentence 5	the price of oils	Using a tangible framing attribute to indicate the characteristics of the following noun phrase
Sentence 5	increased slightly	Using a collocation to introduce the chart trends
Sentence 5	from May 2010 to December 2010	Using a formula to introduce a period of time
Sentence 6	slightly decrease	Using a collocation to introduce the chart trends

Sentence 6	between December 2010 and April 2011	Using a formula to introduce a period of time
Sentence 6	The price of cereals	Using a tangible framing attribute to indicate the characteristics of the following noun phrase
Sentence 6	from May 2010 to December 2010	Using a formula to introduce a period of time
Sentence 7	the price of meat	Using a tangible framing attribute to indicate the characteristics of the following noun phrase
Sentence 7	over a month	Using a tangible framing attribute to refer to the measurable attributes of the following noun phrase
Sentence 7	between January 2011 and April 2011	Using a formula to introduce a period of time
Sentence 8	The price of sugar	Using a tangible framing attribute to indicate the characteristics of the following noun phrase
Sentence 8	slightly increase	Using a collocation to introduce the chart trends
Sentence 8	between May 2010 and December 2010	Using a formula to introduce a period of time
Sentence 8	from December 2010 to April 2011	Using a formula to introduce a period of time
Sentence 9	In conclusion	Using a summarizer to indicate the end of the text
Sentence 9	it can be seen	Using a sentence stem to structure the text meaning and engage the reader in the writer's evaluation of the information presented in the text
Sentence 9	the prices of food	Using a tangible framing attribute to indicate the characteristics of the following noun phrase
Sentence 9	the prices of food have fluctuated	Using a collocation to introduce the chart trends
Sentence 9	from May 2010 to April 2011	Using a formula to introduce a period of time
Sentence 9	It can be noticed	Using a sentence stem to structure the text meaning and engage the reader in the writer's evaluation of the information presented in the text

Moo's Textual Data

Moo's Pretest

In the last twelve months commodity prices go up. Especially sugar its price changes from 200,000 per ton into 400,000 per ton in Dec 2010. Overall when an item price goes up in an expected way the consumers will try to find a different way or resource of it. In my opinion the main reason that makes all of these prices go up is that the consumers want more of one product sugar for example when they buy sugar more than they want and the company doesn't have any more until the next season the price will go up immediately.

Table E25		
<i>Manual Coding: Moo's Pretest</i>		
Sentence #	Formulaic Sequence	Function
Sentence 2	from 200,000 per ton into 400,000 per ton	Using a formula to introduce price range
Sentence 3	for example	Using a formula to list examples

Moo's Posttest

The graph shows fluctuation in goods prices from May, 2013 to April 2011. According to the graph sugar has increased dramatically between May and Dec 2010, after Dec, 2010 the price of sugar has slightly decreased into 200 US dollar. Both cereals and oil have gradual increase over a period of eight months in 2010. In 2011 their prices have fluctuated. Dairy and meat almost remain the same between May 2010 and Apr 2011. To sum up, According to the graph most of the food prices increased over a period of two years; the increase in prices will affect consumers lives in one way or another due to the limited income they have.

Table E26		
<i>Manual Coding: Moo's Posttest</i>		
Sentence #	Formulaic Sequence	Function
Sentence 1	The graph shows	Using a sentence head to demonstrate that something exists
Sentence 1	fluctuation in goods prices	Using a collocation to introduce the chart trends
Sentence 1	from May, 2013 to April 2011	Using a formula to introduce a period of time (comma splice)
Sentence 2	According to the	Using a three-word bundle to introduce information presented by elements in the text
Sentence 2	increased dramatically	Using a collocation to introduce the chart trends
Sentence 2	between May and Dec 2010	Using a formula to introduce a period of time
Sentence 2	the price of sugar	Using a tangible framing attribute to indicate the characteristics of the following noun phrase
Sentence 2	slightly decreased	Using a collocation to introduce the chart trends
Sentence 3	Both cereals and oil	Using a sentence builder to connect and compare discourse elements
Sentence 3	gradual increase	Using a collocation to introduce the chart trends
Sentence 3	over a period of eight months	Using a tangible framing attribute to refer to the measurable attributes of the following noun phrase
Sentence 4	prices have fluctuated	Using a collocation to introduce the chart trends
Sentence 5	Dairy and meat	Using a sentence builder to connect and compare discourse elements
Sentence 5	remain the same	Using a collocation to introduce the chart trends
Sentence 5	between May 2010 and Apr 2011	Using a formula to introduce a period of time
Sentence 6	To sum up	Using a summarizer to indicate the end of the text
Sentence 6	According to the	Using a three-word bundle to introduce information presented by elements in the text
Sentence 6	over a period of two years	Using a tangible framing attribute to refer to the measurable attributes of the following noun phrase
Sentence 6	due to the	Using a stance bundle to introduce a degree of certainty

Moo's Delayed Posttest

The graph shows the changes in the food prices over a period of one year. The price of sugar increased gradually from May 2010 to Dec 2010; and then the price decreased gradually between

Dec 2010 and April 2011. Both cereals and oils rise and fall evenly over a one year period, However their prices are higher than before. Dairy price fluctuated from May 2010 to Apr 2011. Finally the meat price slightly increased between Jan 2011 and Apr 2011. All of these increases in goods prices will affect consumers in some way. To sum up, all of the goods prices generally increased from May 2010 to April 2011.

Table E27		
<i>Manual Coding: Moo's Delayed Posttest</i>		
Sentence #	Formulaic Sequence	Function
Sentence 1	The graph shows	Using a sentence head to demonstrate that something exists
Sentence 1	over a period of one year	Using a tangible framing attribute to refer to the measurable attributes of the following noun phrase
Sentence 2	The price of sugar	Using a tangible framing attribute to indicate the characteristics of the following noun phrase
Sentence 2	increased gradually	Using a collocation to introduce the chart trends
Sentence 2	from May 2010 to Dec 2010	Using a formula to introduce a period of time
Sentence 2	decreased gradually	Using a collocation to introduce the chart trends
Sentence 2	between Dec 2010 and April 2011	Using a formula to introduce a period of time
Sentence 3	Both cereals and oil	Using a sentence builder to connect and compare discourse elements
Sentence 3	rise and fall evenly	Using a collocation to introduce the chart trends
Sentence 3	over a one year period	Using a tangible framing attribute to refer to the measurable attributes of the following noun NP
Sentence 4	price have fluctuated	Using a collocation to introduce the chart trends
Sentence 4	from May 2010 to Apr 2011	Using a formula to introduce a period of time
Sentence 5	To sum up	Using a summarizer to indicate the end of the text
Sentence 5	from May 2010 to April 2011	Using a formula to introduce a period of time

Gabriella's Textual Data

Gabriella's Pretest

The graph shows things which are buying and selling throughout one years period. There are five different products whose price change over the last 12 month. In 2010, between May and January, the prices of products are between 130 an 220. In Autumn there is an increase in sugar. When we look the price of meat we can see that there is no so much differences throughtout the year. There are upwards and downwards in cereals and dairies, December to February. Oil and sugar price stayed stable throughtout one mont, there is a sharp increase, 260 to 320, in sugar. The cheapest product is meat because its price is the lowest through the period. The consumers may buy always the meat because it is the cheapest.

However, if they want to make cookies, cakes or sugar stuffs they need to think it two time because the prices are so high to afford. To sum up, the prices changed through twelve months. The most remarkable thing is the sugers high price. Customers can buy cereals, oils or dairy because their price are normal to afford.

Table E28		
<i>Manual Coding: Gabriella's Pretest</i>		
Sentence #	Formulaic Sequence	Function
Sentence 1	The graph shows	Using a sentence head to demonstrate that something exists
Sentence 1	buying and selling	Using a sentence builder to connect and compare discourse elements
Sentence 2	There are	Using a sentence stem to indicate that something exists
Sentence 2	over the last 12 month	Using a tangible framing attribute to refer to the measurable attributes of the following noun phrase
Sentence 3	between May and January	Using a formula to introduce a period of time
Sentence 3	the prices of products	Using a tangible framing attribute to indicate the characteristics of the following noun phrase

Sentence 3	between 130 an 220	Using a formula to introduce a period of time
Sentence 4	there is	Using a sentence stem to indicate that something exists
Sentence 5	the price of meat	Using a tangible framing attribute to indicate the characteristics of the following noun phrase
Sentence 5	there is	Using a sentence stem to indicate that something exists
Sentence 6	There are	Using a sentence stem to indicate that something exists
Sentence 6	upwards and downwards	Using a sentence builder to connect and compare discourse elements
Sentence 6	cereals and dairies	Using a sentence builder to connect and compare discourse elements
Sentence 7	Oil and sugar	Using a sentence builder to connect and compare discourse elements
Sentence 7	stayed stable	Using a collocation creatively to introduce the chart trends
Sentence 7	there is	Using a sentence stem to indicate that something exists
Sentence 7	sharp increase	Using a collocation to introduce the chart trends
Sentence 11	To sum up	Using a summarizer to indicate the end of the text.

Gabriella's Posttest

The graph shows different kind of goods prices over a period of one year. According to the graph, the increases and decreases in these commodities will affect people who are in cooking sector. It has been clear from the chart that the sugar has the highest price even if the price decrease to 300\$. Cereals oils and dairy decreased 20 dolar between January and February. The price of meat remains stable throughout the year. The sugar's and the oils prices did not change from January to February. In conclusion, It can be seen that there was an increase, a decrease and a stability in the goods between the years 2010 and 2011.

More obvious thing from the chart is that the sugar protects its high price during one year period.

Table E29		
<i>Manual Coding: Gabriella's Posttest</i>		
Sentence #	Formulaic Sequence	Function

Sentence 1	The graph shows	Using a sentence head to demonstrate that something exists
Sentence 1	different kind of goods prices	Using a tangible framing attribute to indicate the characteristics of the following noun phrase
Sentence 1	over a period of one year	Using a tangible framing attribute to refer to the measurable attributes of the following noun phrase
Sentence 2	According to the	Using a three-word bundle to introduce information presented by elements in the text
Sentence 2	the increases and decreases	Using a sentence builder to connect and compare discourse elements
Sentence 3	It has been clear	Using a sentence builder to introduce an impersonal evaluation
Sentence 4	between January and February	Using a formula to introduce a period of time
Sentence 4	The price of meat	Using a tangible framing attribute to indicate the characteristics of the following noun phrase
Sentence 4	remains stable	Using a collocation to introduce the chart trends
Sentence 5	The sugar's and the oils	Using a sentence builder to connect and compare discourse elements
Sentence 5	from January to February	Using a formula to introduce a period of time
Sentence 6	In conclusion	Using a summarizer to indicate the end of the text
Sentence 6	It can be seen	Using a sentence stem to structure the text meaning and engage the reader in the writer's evaluation of the information presented in the text
Sentence 6	there was	Using a sentence stem to indicate that something exists
Sentence 6	between the years 2010 and 2011	Using a formula to introduce a period of time

Gabriella's Delayed Posttest

The graph shows the changes of the goods prices between the years 2010 and 2011. According to the graph, the sugar had the highest price. Moreover it protects its highest price when we compare it with other goods. According to the graph cereals and oils were fluctuating over a one year period. But their price increased at the end of the year.

Meat stayed stable over a one year period. However the price changed 100\$ to 150\$.

The commodity price of dairy started with high price however it closed with low price.

There was a fall in Dairy price 220 to 200. The sugar's price rose rapidly until december then it decreased 400\$ to 300\$. However it had still the highest price. In conclusion, it can be clearly seen that sugar always had the highest price and meat always had the lowest price.

The other commodities prices were fluctuating during one year.

Table E30		
<i>Manual Coding: Gabriella's Delayed Posttest</i>		
Sentence #	Formulaic Sequence	Function
Sentence 1	The graph shows	Using a sentence head to demonstrate that something exists
Sentence 1	the changes of the goods prices	Using a tangible framing attribute to indicate the characteristics of the following noun phrase
Sentence 1	between the years 2010 and 2011	Using a formula to introduce a period of time
Sentence 2	According to the	Using a three-word bundle to introduce information presented by elements in the text
Sentence 3	According to the	Using a three-word bundle to introduce information presented by elements in the text
Sentence 3	cereals and oils	Using a sentence builder to connect and compare discourse elements
Sentence 3	over a one year period	Using a tangible framing attribute to refer to the measurable attributes of the following noun phrase
Sentence 4	stayed stable	Using a collocation to introduce the chart trends
Sentence 4	over a one year period	Using a tangible framing attribute to refer to the measurable attributes of the following noun phrase
Sentence 5	The commodity price of dairy	Using a tangible framing attribute to indicate the characteristics of the following noun phrase
Sentence 6	There was	Using a sentence stem to indicate that something exists
Sentence 6	rose rapidly	Using a collocation to introduce the chart trends
Sentence 7	In conclusion	Using a summarizer to indicate the end of the text
Sentence 7	it can be clearly seen	Using a sentence stem creatively to structure the text meaning and engage the reader in the writer's evaluation of the information presented in the text
Sentence 8	prices were fluctuating	Using a collocation to introduce trends presented in the chart

Shireen's Textual Data

Shireen's Pretest

Between 2010 and 2011 the food price changed dramatically which has a direct effect on consumption level for families and individuals. In 2010 sugar price jumped from 250 per unit to 400 per unit. Other essential products such as cereals, oil, Dairy and meat faced the same phenomenon. That continuous increase in essential products' prices left the consumer with no choice but cut the consumption level, which results in negative consequences on health and performance of each member in the family. Children in the family suffer the most as that cut in their daily nutrition affects their development. For example, Iron deficiency causes dizziness which affects children's ability to concentrate during school time. Moreover, the constant nagging hunger impacts their cognitive abilities.

Table E31		
<i>Manual Coding: Shireen's Pretest</i>		
Sentence #	Formulaic Sequence	Function
Sentence 1	Between 2010 and 2011	Using a formula to introduce a period of time
Sentence 1	families and individuals	Using a sentence builder to connect and compare discourse elements
Sentence 2	from 250 per unit to 400 per unit	Using a formula to introduce price range
Sentence 3	such as	Using an expository phrase to list examples
Sentence 3	health and performance	Using a sentence builder to connect and compare discourse elements
Sentence 4	For example	Using a discourse organizer to exemplify

Shireen's Posttest

This line graph presents the fluctuating in the food prices between 2010 and 2011. On one hand, sugar and oil has dramatic increases between May and December. On the other hand products such as Dairy and meat almost remained the same. This fluctuating in food prices has a Direct effect

on the consumers consumptions. For example, a large family that used to spend around \$ 200 weekly on grocery products affter the increas they either have to increase their budget or decrease the quantity that their buying per week. As result of that the families may have to chang their nutritions happits to be able to meet the demand of healthy life style. To sum up, the increase in food prices has a direct effect on our health and consumption hapits.

Table E32		
<i>Manual Coding: Shireen's Posttest</i>		
Sentence #	Formulaic Sequence	Function
Sentence 1	This line graph present	Using a sentence head but inaccurately to demonstrate that something exists
Sentence 1	the fluctuating in the food prise	Using a collocation to introduce the chart trends
Sentence 1	between 2010 and 2011	Using a formula to introduce a period of time
Sentence 2	On one hand	Using a discourse organizer to contrast elements in the text
Sentence 2	sugar and oil	Using a sentence builder to connect and compare discourse elements
Sentence 2	dramatic increas	Using a collocation to introduce the chart trends
Sentence 2	between May and December	Using a formula to introduce a period of time
Sentence 3	On the other hand	Using a discourse organizer to contrast elements in the text
Sentence 3	such as	Using an expository phrase to list examples
Sentence 3	Dairy and meat	Using a sentence builder to connect and compare discourse elements
Sentence 3	remained the same	Using a collocation to introduce the chart trends
Sentence 4	This fluctuating in food prices	Using a collocation inaccurately to introduce trends presented in the chart
Sentence 5	For example	Using a discourse organizer to exemplify
Sentence 6	As result of	Using a resultative signal inaccurately to make an explicit reference to cause and effect relations in the text
Sentence 6	the demand of healthy life style	Using a tangible framing attribute to indicate the characteristics of the following noun phrase
Sentence 7	To sum up	Using a summarizer to indicate the end of the text

Shireen's Delayed Posttest

This line graph presents the fluctuation in food prices over one year period. Between May 2010 and December 2010 some food products had dramatic increases such as sugar and oil.

On the other hand some food products remained the same. For example, cereals, dairy products and meat prices did not change between May 2010 and December 2010. As a result of these changes families and individuals found themselves in a situation where they have to respond to these changes. For instance, the consumption level for both the individuals and families have change. When the prices increased they have to either adjust their consumption to a less quantity or they have to increase their monthly spending on groceries to be able to purchase the same amount that they used to buy before. To conclude, the fluctuation in the food prices has a direct effect on the consumption for both families and individuals.

Table E33		
<i>Manual Coding: Shireen's Delayed Posttest</i>		
Sentence #	Formulaic Sequence	Function
Sentence 1	This line graph presents	Using a sentence head but inaccurately to demonstrate that something exists
Sentence 1	the fluctuation in food prices	Using a collocation to introduce the chart trends
Sentence 1	over one year period	Using a formula to introduce a period of time
Sentence 2	Between May 2010 and December 2010	Using a formula to introduce a period of time
Sentence 2	dramatic increases	Using a collocation to introduce the chart trends
Sentence 2	such as	Using an expository phrase to list examples
Sentence 2	sugar and oil	Using a sentence builder to connect and compare discourse elements
Sentence 3	On the other hand	Using a discourse organizer to contrast elements in the text
Sentence 3	remained the same	Using a collocation to introduce the chart trends
Sentence 4	For example	Using a discourse organizer to exemplify
Sentence 4	dairy products and meat prices	Using a sentence builder creatively to connect and compare discourse elements
Sentence 4	between May 2010 and December 2010	Using a formula to introduce a period of time

Sentence 5	For example	Using a discourse organizer to exemplify
Sentence 6	As a results of	Using a resultative signal inaccurately to make an explicit reference to cause and effect relations in the text
Sentence 6	families and individuals	Using a sentence builder to connect and compare discourse elements
Sentence 7	For instance	Using a formula to list examples
Sentence 7	both the individuals and families	Using a sentence builder to connect and compare discourse elements
Sentence 8	To conclude	Using a summarizer to indicate the end of the text
Sentence 8	the fluctuation in the food prices	Using a collocation to introduce the chart trends
Sentence 8	both families and individuals	Using a sentence builder to connect and compare discourse elements

Abedeer's Textual Data

Abedeer's Pretest

In this graph we can notice that one of the product like sugar keep growing during the year and the consumers still using it or in need to it on the other hand we can see that the dairy product stayed at the same range of price and the number of the consumers decrease.

Also from the graphs we can compare the cereals at December got the highest during the year and growing down at the beginning of 2011 which means that increase the price of the cereals effect indirect way on the consumer. However, the consumer be awared of the side effects of the sugar but it is still the highest product in the graph that means we have to teach humans about the positivity and negativitey of some products. So they can decide is it useful to their daily life or not. At the end we can see that changes in prices doesn't affect much the consumer decesions but it's their way to deal with the product.

Table E34		
<i>Manual Coding: Abedeer's Pretest</i>		
Sentence #	Formulaic Sequence	Function
Sentence 1	one of the product	Using a tangible framing attribute to indicate the characteristics of the following noun phrase
Sentence 1	on the other hand	Using a discourse organizer inaccurately to contrast elements in the text
Sentence 1	the same range of price	Using a tangible framing attribute to indicate the characteristics of the following noun phrase
Sentence 1	the number of the consumers	Using a tangible framing attribute to indicate the characteristics of the following noun phrase
Sentence 2	the price of the cereals	Using a tangible framing attribute to indicate the characteristics of the following noun phrase
Sentence 3	the side effects of the sugar	Using a tangible framing attribute to indicate the characteristics of the following noun phrase
Sentence 3	the positivity and negativitey of some products	Using a sentence builder to connect and compare discourse elements embedded in a tangible framing attribute to indicate the characteristics of the following noun phrase

Abedeer's Posttest

This graph shows the rise and fall in prices of food between May 2010 and April 2011 and the behaviour issue of the buyer. According to the graph, it is clear that the sugar prices have increased gradually from May 2010 and peaking on December 2010 before it started to drop gradually in 2011. This indicates that the consumers' demand more sugar in winter than they do in summer due to the weather climate and holidays. Both cereals and oils are correlated. They've shown a gradual increase in their prices from May 2010 and peaking in Dec, 2010 and then started to fluctuate. However, dropping to a price that is higher than their lowest price in 2010. Dairy and meat have shown a very minimal change in price and demand which we could consider unchanged for the rest of the year. In conclusion, it is clear from the data given that storable commodities like sugar, cereals and oils change in price during a full year. On the other hand, non-storable commodities like meat and dairy do not change in price during a year cycle. In this respect, consumers should plan their home finance according to their consumption behaviour.

Table E35		
<i>Manual Coding: Abedeer's Posttest</i>		
Sentence #	Formulaic Sequence	Function
Sentence 1	This graph shows	Using a sentence head to demonstrate that something exists
Sentence 1	the rise and fall	Using a collocation to introduce trends presented in the chart
Sentence 1	between May 2010 and April 2011	Using a formula to introduce a period of time
Sentence 1	the behaviour issue of the buyer	Using a tangible framing attribute to indicate the characteristics of the following noun phrase
Sentence 2	According to the	Using a three-word bundle to introduce information presented by elements in the text
Sentence 2	it is clear	Using a sentence builder to introduce an impersonal evaluation
Sentence 2	increased gradually	Using a collocation to introduce trends

		presented in the chart
Sentence 2	drop gradually	Using a collocation to introduce the chart trends
Sentence 3	due to the	Using a stance bundle to introduce cause and effect
Sentence 4	Both cereals and oils	Using a sentence builder to connect and compare discourse elements
Sentence 5	gradual increas	Using a collocation to introduce the chart trends
Sentence 6	Dairy and meat	Using a sentence builder to connect and compare discourse elements
Sentence 6	price and demand	Using a sentence builder to connect and compare discourse elements
Sentence 6	the rest of the year	Using a tangible framing attribute to indicate the characteristics of the following noun phrase
Sentence 7	In conclusion	Using a summarizer to indicate the end of the text
Sentence 7	it is clear from	Using a sentence builder to introduce an impersonal evaluation
Sentence 8	On the other hand	Using a discourse organizer to contrast elements in the text
Sentence 8	meat and dairy	Using a sentence builder to connect and compare discourse elements
Sentence 9	according to their	Using a three-word bundle creatively to establish links between elements in the text

Abedeer's Delayed Posttest

The graph shows the rise and fall in food prices for the period of May 2010 to April 2011.

According to the graph, sugar, cereals and oils are increased gradually in 2010 and have a slight decrease in the begging of year 2011. Then from January tell April 2011 there is a big decrease in both sugar and oils. On the other hand, cereals has a slight increase. In addition, dairy and meat have a slight decrease at the beginning of 2010 and stayed at the same rate for the rest of the year and tell April 2011. In conclusion, we can notice the changes in commodity prices depend on the needs of the customerce which depend on the weather.

On the other hand the non commodity food stayed the same with a little changes in prices and consuming. In conclusion, food prices rose and fell from May 2010 to April 2011.

Table E36		
<i>Manual Coding: Abedeer's Delayed Posttest</i>		
Sentence #	Formulaic Sequence	Function
Sentence 1	This graph shows	Using a sentence head to demonstrate that something exists
Sentence 1	the rise and fall	Using a collocation to introduce trends presented in the chart
Sentence 1	for the period of May 2010 to April 2011	Using a formula creatively to introduce a period of time
Sentence 2	According to the	Using a three-word bundle to introduce information presented by elements in the text
Sentence 2	increased gradually	Using a collocation to introduce trends presented in the chart
Sentence 2	slight decrease	Using a collocation to introduce the chart trends
Sentence 2	the beginning of year 2011	Using a tangible framing attribute to indicate the characteristics of the following noun phrase
Sentence 3	from January till April	Using a formula creatively to introduce a period of time
Sentence 3	both sugar and oils	Using a sentence builder to connect and compare discourse elements
Sentence 4	On the other hand	Using a discourse organizer to contrast elements in the text
Sentence 4	slight increase	Using a collocation to introduce the chart trends
Sentence 5	In addition	Using a transition signal to organize the text and add new information
Sentence 5	dairy and meat	Using a sentence builder to connect and compare discourse elements
Sentence 5	the beginning of 2010	Using a tangible framing attribute to indicate the characteristics of the following noun phrase
Sentence 5	the rest of the year	Using a tangible framing attribute to indicate the characteristics of the following noun phrase
Sentence 6	In conclusion	Using a summarizer to indicate the end of the text
Sentence 6	the needs of the customer	Using a tangible framing attribute to indicate the characteristics of the following noun phrase
Sentence 7	On the other hand	Using a discourse organizer to contrast elements in the text
Sentence 7	stayed the same	Using a collocation to introduce the chart trends
Sentence 7	prices and consuming	Using a sentence builder to connect and compare discourse elements