

Newcastle
University

**Acquisition of Lexical Collocations: A corpus-assisted
contrastive analysis and translation approach**

Rezan Mohammed Alharbi

**Thesis Submitted in Partial Fulfilment of the Requirements for the
Degree of Doctor of Philosophy in Applied Linguistics**

Newcastle University

School of Education, Communication and Language Sciences

Jan 2017

Abstract

Research from the past 20 years has indicated that much of natural language consists of formulaic sequences or chunks. It has been suggested that learning vocabulary as discrete items does not necessarily help L2 learners become successful communicators or fluent and accurate language users. Collocations, i.e. words that usually go together as one form of formulaic sequences, constitute an inherent problem for ESL/ EFL learners. Researchers have submitted that non-congruent collocations, i.e. collocations that do not have corresponding L1 equivalents, are especially difficult to acquire by ESL/ EFL learners. This study examines the effect of three Focus-on-Forms instructional approaches on the passive and active acquisition of non-congruent collocations: 1) the non-corpus-assisted contrastive analysis and translation (CAT) approach, 2) the corpus-assisted CAT approach, and 3) the corpus-assisted non-CAT approach. To fully assess the proposed combined condition (i.e. the corpus-assisted CAT) and its learning outcomes, a control group under no-condition was included for a baseline comparison. Thirty collocations non-congruent with the learners' L1 (Arabic) were chosen for this study.

129 undergraduate EFL learners in a Saudi University participated in the study. The participants were assigned to the three experimental groups and to the control group following a cluster random sampling method. The corpus-assisted CAT group performed (L1/ L2 and L2/ L1) translation tasks with the help of bilingual English/ Arabic corpus data. The non-corpus CAT group was assigned text-based translation tasks and received contrastive analysis of the target collocations and their L1 translation options from the teacher. The non-contrastive group performed multiple-choice/ gap-filling tasks with the help of monolingual corpus data, focusing on the target items. Immediately after the intervention stage, the three groups were tested on the retention of the target collocations by two tests: active recall and passive recall. The same tests were administered to the participants three weeks later. The corpus-assisted CAT group significantly outperformed the other two groups on all the tests. These results were discussed in light of the 'noticing', 'task-induced involvement load', and 'pushed output' hypotheses and the influence that L1 exerts on the acquisition of L2 vocabulary. The discussion includes an evaluation of the three instructional conditions in

relation to different determinants, dimensions and functions within the hypotheses.

Dedication

*To my beloved grandparents, Zaini and Ishrat, and to my Dad
(may Allah rest their souls in peace)*

Acknowledgements

First and foremost, my thanks should be to Allah (SWT) for helping and guiding me, and for providing me with patience and strength throughout this tough journey towards a PhD.

My sincere thanks go to my supervisors Dr. Mei Lin and Dr. Dawn Knight for their insightful comments, constant encouragement, patience and kind support.

I would like to express my gratitude to the students who participated in this study, and to the University staff who gave me access to the classrooms.

In addition, I would also like to thank my good friends and colleagues in the ECLS faculty for their continuing moral support, encouragement and advice.

Thanks are also due to the Saudi Ministry of Higher Education and King Saud University who generously funded this thesis through a scholarship.

I am immensely indebted to Dr. Suhad Sonbul whose guidance, insightful comments, patience, and continuous positive attitude and encouragement helped this study to see the light of day. I am also indebted to my friends in Saudi and in Newcastle. I feel very fortunate to be surrounded by such a sincere, kind and supportive group of friends. This journey would have been much more difficult without them being around to listen, comfort, advise and encourage.

Finally, and most importantly, I have no words to express my profound gratitude to my lovely mother Sabah Kashmiri and to my dearest husband Majed Salama whose love, prayers, encouragement and support kept me going. This journey would have not been completed without them in my life. I owe a special debt of gratitude and apologies to my children, Esam, Dana and Qusai for their tolerance of my mood swings, absence and lack of support sometimes, and for being understanding and comforting most of the time. My sincere thanks go to my much-loved brothers (Razen and Rayan) and sisters (Noran and Ghofran) and to their beautiful families for supporting me wholeheartedly throughout my postgraduate studies.

Contents

Abstract.....	i
Dedication.....	iii
Acknowledgements	iv
Contents.....	v
List of Tables.....	x
List of Figures.....	xii
List of Abbreviations.....	xiii
Chapter 1: Introduction.....	1
1.1 Field of research	1
1.2 Aim.....	2
1.3 Research hypotheses.....	3
1.4 Thesis outline.....	3
Chapter 2: Collocations: Focus of the Thesis.....	6
2.1 Vocabulary and language learning	6
2.2 Vocabulary in the EFL context of Saudi Arabia	7
2.3 Vocabulary knowledge: a multidimensional construct	8
2.4 A needs analysis study: narrowing the research scope.....	12
2.4.1 Analysis and findings.....	13
2.4.2 Discussion.....	16
2.5 Collocational knowledge of EFL Learners.....	19
2.5.1 Corpus-based research	20
2.5.2 Research involving paper-and-pencil elicitation tests	23
2.5.3 Research involving psycholinguistic measures	25
2.6 The role of formulaic sequences in language learning	26
2.7 What is a collocation?.....	28
2.7.1 The frequency-based approach	29
2.7.1.1 Statistical measurements of collocations.....	32
2.7.2 The phraseological approach	35
2.7.3 A working definition of collocation: a complementary approach	38
2.8 Summary.....	39
Chapter 3: Methodological rationale: corpus- assisted contrastive analysis and translation	41
3.1 Second/ foreign language acquisition: to instruct or not to instruct	42
3.2 Instructed second language acquisition	43
3.2.1 Instructed vocabulary acquisition	46

4.3.2.1 Worksheets for experimental group 1 (-DDL +CAT) and experimental group 2 (+DDL +CAT)	103
4.3.2.2 Worksheets for experimental groups 3 (+DDL -CAT)	104
4.3.3 Designing the corpus data sheets	105
4.4 Procedure (experimental groups)	110
4.4.1 Experimental group 1 (-DDL +CAT)	111
4.4.2 Experimental group 2 (+DDL +CAT)	113
4.4.3 Experimental group 3 (+DDL -CAT)	114
4.5 Collecting data on collocational knowledge: measures.....	114
4.5.1 Pre, post and delayed post-tests of passive collocational knowledge.....	115
4.5.2 Pre, post and delayed post-tests of active collocational knowledge	116
4.6 Marking the tests and analysing the data.....	116
4.6.1 Marking (English ↔ Arabic translation tests).....	116
4.6.2 Analysing the data	118
4.7 Validity and reliability of the research	118
4.7.1 Reliability.....	118
4.7.2 Validity	120
4.7.2.1 Multiple facets of validity	121
4.7.2.2 Internal validity	123
4.7.2.3 External validity	124
4.8 Ethical considerations.....	125
4.9 Summary.....	126
Chapter 5: Data Analysis and Results	127
5.1 Parametric versus non-parametric statistical tests	127
5.1.1 Checking assumptions	128
5.2 Effect of treatments: within group comparisons.....	130
5.2.1 Effect of (-DDL +CAT/ group 1) treatment on collocational knowledge	132
5.2.1.1 Effect of (-DDL +CAT/ group 1) treatment on passive knowledge	133
5.2.1.2 Effect of (-DDL +CAT/ group 1) treatment on active knowledge.....	135
5.2.2 Effect of (+DDL + CAT/ group 2) treatment on collocational knowledge	137
5.2.2.1 Effect of (+DDL +CAT/ group 2) treatment on passive knowledge	138
5.2.2.2 Effect of (+DDL + CAT/ group 2) treatment on active knowledge.....	139
5.2.3 Effect of (+DDL -CAT/ group 3) treatment on collocational knowledge	141
5.2.3.1 Effect of (+DDL -CAT/ group 3) treatment on passive knowledge	142
5.2.3.2 Effect of (+DDL -CAT/ group 3) treatment on active knowledge.....	143

5.3 Effect of No-treatment: control group.....	145
5.3.1 The control group performance on passive recall tests	146
5.3.2 The control group performance on active recall tests.....	147
5.4 Summary	148
5.5 Effect of the treatments: between- group comparisons.....	150
5.5.1 Entry level	150
5.5.2 Participants performance on post-tests: All groups.....	152
5.5.2.1 Participants performance on post-tests of passive collocational knowledge.....	152
5.5.2.2 Participants' performance on post-tests of active collocational knowledge.....	154
5.5.3 Effect of CAT.....	155
5.5.3.1 Effect of (-DDL +CAT VS. +DDL +CAT) on passive knowledge of collocations	156
5.5.3.2 Effect of (-DDL +CAT VS. +DDL -CAT) treatments on collocational active knowledge.....	159
5.5.4 Effect of DDL.....	161
5.5.4.1 Effect of (+DDL +CAT VS. +DDL -CAT) on passive collocational knowledge	162
5.5.4.2 Effect of (+DDL +CAT VS. +DDL -CAT) on active collocational knowledge	164
5.6 Summary	166
Chapter 6: Discussion	170
6.1 Section one: participants' knowledge of collocations.....	170
6.1.1 Pre-treatments knowledge of collocations	170
6.1.2 Post-treatments knowledge of collocations.....	174
6.2 Discussion of the research hypotheses and sub-hypotheses	176
6.2.1 Hypothesis one and sub-hypotheses.....	176
6.2.2 Hypothesis two and sub-hypotheses.....	178
Noticeability	180
Involvement load	184
Pushed output.....	186
6.2.3 Why medium effect size?	190
6.3 Summary and final remarks on the proposed CAT condition	190
Chapter 7: Conclusion.....	192
7.1 Brief summary of the study.....	192
7.2 contribution of the study	193
7.3 Limitations and directions for future research	194

7.3.1 The target non-congruent collocations	194
7.3.2 Corpus resources	195
7.3.3 Measurements of collocational knowledge.....	195
7.3.4 Further reflections	196
7.4 Pedagogical implications	197
7.4.1 Developing collocational knowledge.....	197
7.4.2 Scaffolding contrastive FFI and raising cross-linguistic collocational awareness	197
7.4.3 Scaffolding corpus-assisted vocabulary learning	199
References	200
Appendix A: Vocabulary level tests (K2 & K3)	221
Appendix B: Target collocations, frequency and MI scores	224
Appendix C: Clued-recall pilot test	225
Appendix D: MC distracters.....	229
Appendix E: Reading passages and comprehension questions (samples).....	229
Appendix F: Experimental groups' worksheets (samples).....	234
Appendix G: Bilingual corpus data (samples).....	240
Appendix H: Monolingual corpus data (samples).....	242
Appendix I: Pre, post and delayed post-test (active recall)	245
Appendix J: Pre, post and delayed post-test (passive recall).....	247
Appendix K: Research information sheet.....	249
Appendix L: Consent form	250

List of Tables

Table 2. 1: What is involved in knowing a word	11
Table 3. 1: Summary of R. Ellis' (2001a) FFI types.....	46
Table 4. 1: Descriptive statistics and normality test (QOPT).....	95
Table 4. 2: QOPT (Kruskal-Wallis test for between groups comparison).....	96
Table 4. 3: VLT Descriptive Statistics and Normality test	97
Table 4. 4: VLT (between-groups comparison).....	97
Table 4. 5: Sets of target non-congruent collocations.....	103
Table 4. 6: intra-rater reliability	120
Table 4. 7: Intraclass Correlation Coefficient	120
Table 4. 8: Summary of research methods.....	126
Table 5. 1: Shapiro-Wilk test of normality of pre-tests.....	129
Table 5. 2: Shapiro-Wilk test of normality of post-tests.....	129
Table 5. 3: Shapiro-Wilk test of normality of delayed post-tests.....	130
Table 5. 4: Descriptive statistics (-DDL +CAT/ group 1)	132
Table 5. 5: All passive recall tests (-DDL +CAT/ group 1).....	133
Table 5. 6: Pre/ post-tests of passive recall (-DDL +CAT/ group 1).....	134
Table 5. 7: Pre /delayed post-tests of passive recall (-DDL +CAT/ group 1).....	134
Table 5. 8: Post/ delayed post-tests of passive recall (-DDL +CAT/ group 1).....	135
Table 5. 9: All active recall tests (-DDL +CAT/ group 1).....	135
Table 5. 10: Pre/ post-tests of active recall (-DDL +CAT/ group 1)	136
Table 5. 11: Pre/ delayed post-tests of active recall (-DDL +CAT/ group 1).....	136
Table 5. 12: Post/ delayed post-test of active recall (-DDL +CAT/ group 1).....	136
Table 5. 13: Descriptive statistics (+DDL +CAT/ group 2).....	137
Table 5. 14: All passive recall tests (+DDL +CAT/ group 2).....	138
Table 5. 15: Pre/ post-tests of passive recall (+DDL +CAT/ group 2)	138
Table 5. 16: Pre/ delayed post-tests of passive recall (+DDL +CAT/ group 2).....	139
Table 5. 17: Post/ delayed post-tests of passive recall (+DDL +CAT/ group 2)	139
Table 5. 18: All active recall tests (+DDL +CAT/ group 2)	139
Table 5. 19: Pre/ post-tests of active recall (+DDL +CAT/ group 2)	140
Table 5. 20: Pre/ delayed post-tests of active recall (+DDL +CAT/ group 2).....	140
Table 5. 21: Post/ delayed post-tests of active recall (+DDL +CAT/ group 2)	141
Table 5. 22: Descriptive statistics (+DDL -CAT/ group 3)	141
Table 5. 23: All passive recall tests (+DDL -CAT/ group 3).....	142
Table 5. 24: Pre/ post-tests of passive recall (+DDL -CAT/ group 3)	142
Table 5. 25: Pre/ delayed post-tests of passive recall (+DDL -CAT/ group 3).....	143
Table 5. 26: Post/ delayed post-tests of passive recall (+DDL -CAT/ group 3).....	143
Table 5. 27: All active recall tests (+DDL -CAT/ group 3)	144
Table 5. 28: Pre/ post-tests of active recall (+DDL -CAT/ group 3)	144
Table 5. 29: Pre/ delayed post-tests of active recall (+DDL -CAT/ group 3).....	144
Table 5. 30: Post/ delayed post-tests of active recall (+DDL -CAT/ group 3)	145
Table 5. 31: Descriptive statistics of the control group	145
Table 5. 32: All passive recall tests of the control group.....	146

Table 5. 33: Paired-sample t-tests of passive recall tests (control group).....	147
Table 5. 34: All active recall tests of the control group	147
Table 5. 35: Wilcoxon signed-rank tests of active recall tests (control group).....	148
Table 5. 36: Descriptive statistics of the pre-test scores of all groups	150
Table 5. 37: ANOVA’s test results of pre-tests of passive recall between all groups .	151
Table 5. 38: Kruskal-Wallis test of pre-tests of the active recall between all groups..	151
Table 5. 39: Descriptive statistics of the post- tests of passive recall of all groups.....	152
Table 5. 40: Post-tests of the passive collocational knowledge between all groups....	153
Table 5. 41: Descriptive statistics of the post-tests of active recall of all groups	154
Table 5. 42: Post-tests of the active collocational knowledge between all groups	155
Table 5. 43: Pre-tests of passive collocational knowledge (group 1 & 2)	156
Table 5. 44: Immediate post-tests of passive collocational knowledge (group 1&2)..	156
Table 5. 45: Immediate post-tests of passive collocational knowledge (group 1& control group/ 2& control group)	157
Table 5. 46: Delayed post-tests of passive collocational knowledge (group 1 & 2)....	158
Table 5. 47: Delayed post-tests of passive collocational knowledge (group 1& control group/ 2& control group)	158
Table 5. 48: Pre-tests of active collocational knowledge (group 1 & 2).....	159
Table 5. 49: Immediate post-tests of active collocational knowledge (group 1 & 2) ..	159
Table 5. 50: Immediate post-tests of active collocational knowledge (group 1& control group/ 2& control group)	160
Table 5. 51: Delayed post-tests of active collocational knowledge (group 1 & 2).....	161
Table 5. 52: Delayed post-tests of active collocational knowledge (group 1& control group/ 2& control group)	161
Table 5. 53: Pre-tests of passive collocational knowledge (group 2 & 3)	162
Table 5. 54: Immediate post-tests of passive collocational knowledge (group 2 & 3)	162
Table 5. 55: Immediate post-tests of passive collocational knowledge (group 2& control group/ 3& control group)	163
Table 5. 56: Delayed post-tests of passive collocational knowledge (group 2 & 3)....	163
Table 5. 57: Delayed post-tests of passive collocational knowledge (group 2&control group/ 3&control group)	164
Table 5. 58: Pre-tests of active collocational knowledge (group 2 & 3).....	164
Table 5. 59: Immediate post-tests of active collocational knowledge (group 2 & 3) ..	165
Table 5. 60: Immediate post-tests of active collocational knowledge (group 2& control group/ 3& control group)	165
Table 5. 61: Delayed post-tests of active collocational knowledge (group 1, 2& control group)	166

List of Figures

Figure 2. 1: Illustration of surface co-occurrence for the word pair (<i>hat, roll</i>)	30
Figure 2. 2: Illustration of textual co-occurrence for the word pair (<i>hat, over</i>).....	31
Figure 2. 3: Illustration of syntactic co-occurrence	32
Figure 2. 4: Phraseological categories	36
Figure 2. 5: Collocational continuum	37
Figure 2. 7: Types of lexical collocations	37
Figure 3. 1: Concordance of 'majority' from the British National Corpus (BNC) in..... KWIC format (taken from the Brigham Young University, BYU, website).....	62
Figure 4. 1: The overall research design and data collection procedure.....	93
Figure 4. 2: Placement test categorization (Chi-Square test).....	96
Figure 4. 3: The concordance of 'violence' in WordSmith concordance tool.....	108
Figure 4. 4: The concordance of 'violence' in Examine32 Text Search tool	109
Figure 4. 5: An example of Kuwait University E/A Parallel Corpus layout	110
Figure 4. 6: Collocation sets occurrences in reading passages	111

List of Abbreviations

Adj.	Adjective
BIA	Bilingual Interactive Activation (Model)
BNC	British National Corpus
CAH	Contrastive Analysis Hypothesis
CAT	Contrastive Analysis and Translation
COCA	Corpus of Contemporary American English
DDL	Data- Driven Learning
ECLS	School of Education, Communication & Language Sciences
EFL	English as a Foreign Language
ESL	English as a Second Language
FFI	Form-Focused Instruction
FL	Foreign Language
FonF	Focus on Form
FonFs	Focus on Forms
ISLA	Instructed Second Language Acquisition
KWIC	Key Word In Context
L1	First Language
L2	Second Language
LSP	Language for Specific Purposes
<i>M</i>	Mean
<i>MD</i>	Median
MC	Multiple choice
MFI	Meaning-Focused Instruction
MI	Mutual Information
MT	Mother Tongue
<i>N</i>	Number (sample size)
NNS	Non-Native Speakers
NS	Native Speakers
OQPT	Oxford Quick Placement Test
PLI	Planned Lexical Instruction
<i>r</i>	Effect Size

RHM	Revised Hierarchal Model
SD	Standard Deviation
TOEFL	Test of English as Foreign Language
USE	Uppsala student English Corpus
SLA	Second Language Acquisition
VLT	Vocabulary Level Test

Chapter 1: Introduction

1.1 Field of research

After a long history of neglect in second language teaching and learning pedagogy, it has now been suggested that vocabulary learning is a vital component and a central part of language learning (Meara, 1980; Nation, 2001). Researchers (e.g. Milton, 2009, 2013; Nation, 2001; Nation & Webb, 2011; Read, 2000) have asserted that words constitute the building blocks of language, and that language will not exist without them. For many years, research and practice on L2 vocabulary has been primarily concerned with single discrete words (Schmitt & Carter, 2004). The view that having a large repertoire of words is advantageous for all language learners is not in question, but it is not enough. Wray (2002) confirms this notion as she states: “To know a language you must know not only its individual words, but also how they fit together” (p. 143). As suggested by Gyllstad (2007), certain learner categories, such as university-level students, translators and students trained to be teachers, need to attain a native-like command of an L2. To achieve that, learners need to attend to formulaic language/ prefabricated chunks. With regard to formulaic language, Wray (2002) summarizes three observations made in the literature: (1) native speakers seem to use formulaic language as an easy option in their communication and processing; (2) learners in the early stages of L1 and L2 acquisition rely heavily on formulaic language; (3) formulaic language has, strikingly, been found to be the biggest stumbling block to sounding native-like for L2 learners of intermediate and advanced proficiency.

During the last few decades, collocation as a sub-category of formulaic language has received substantial attention in the field of second language learning. There is extensive discussion in the theoretical literature (e.g. Henriksen, 1999; McCarthy, 1990; Nation, 2001; Richards, 1976) on the advantages of developing collocational knowledge in language learning. It is broadly acknowledged that collocation is a crucial aspect of lexical knowledge. A well-developed collocational knowledge is necessary to transfer receptive word knowledge into productive use (Beheydt, 1987; Lin, 2002; Liu, 2000). Nevertheless, a body of empirical research has demonstrated that EFL learners, even those at advanced levels, have been plagued by underdeveloped collocational knowledge. Additionally, their knowledge of collocations is strongly and negatively influenced by their L1. Collocations that have no literal equivalents in the learners’ L1 are harder to produce and to process (see e.g. Nesselhauf, 2003, 2005; Yamashita & Jiang, 2010).

Given these arguments and evidence, the theoretical and empirical literature both call for pedagogical actions to develop learners' collocational knowledge and to raise their awareness of this linguistic phenomenon and the interlingual difficulties it may constitute.

1.2 Aim

During the last two decades, there have been conflicting views among linguists on how collocational knowledge could be developed, and on the best way to acquire it in instructional settings. Some researchers (e.g. Marton, 1977; Krashen 1989) have argued that single words and collocations are best acquired incidentally through exposure to language input. Researchers in other recent studies have asserted that though such acquisition through exposure is possible, explicit instruction (i.e. Form-Focused Instruction/ FFI) is the best way to learn collocations (e.g. Laufer & Girsai, 2008a, b; Webb & Kagimoto, 2009, 2011; Sonbul, 2012).

Experimental studies exploring EFL learners' acquisition of collocations under different FFI conditions are relatively scarce, despite the long-standing interest and increased attention to the notion of collocation in the literature. Among the studies that do exist, few have addressed the acquisition of non-congruent collocations, i.e. collocations with no word-for-word equivalents in the learners' L1 (Laufer & Girsai, 2008a, b; Chan & Liou, 2005). These studies employed different Form-Focused instructional approaches to the teaching and learning of non-congruent collocations and argued for their efficacy. While Laufer and Girsai's studies called for a contrastive FFI of vocabulary that entails interlingual comparisons with learners' L1 and translation (CAT), Chan and Liou's study called for a pedagogical implementation of Data-Driven Learning (DDL) and corpus resources such as bilingual concordancers for learners to acquire non-congruent collocations. Both studies claimed a raised collocational awareness as a result of the employed approach. To the best of the current researcher's knowledge, no empirical study has attempted to investigate the efficacy of both instructional approaches on the acquisition of non-congruent collocations. This methodological gap needs to be addressed (for a detailed discussion, see chapter 3 below). Accordingly, the aim of this research is to investigate the efficacy of a corpus-assisted contrastive analysis and translation approach for learning lexical non-congruent collocations and for raising learners' awareness. The DDL and CAT approaches have been included to validate the comparison and to establish theoretical grounds for the superiority of my proposed approach/condition. This research targeted adjective/ noun non-congruent collocations. Since most of the reviewed literature has focused on verb/ noun collocations, it was intriguing to

investigate the effect of the three FFI conditions on the acquisition of a different type of lexical collocations.

1.3 Research hypotheses

This research tests two hypotheses with their respective sub-hypotheses. They are as follows:

H1. The corpus-assisted CAT condition will lead to the learning of a significantly larger number (if any) of adj. /noun collocations than the non-corpus-assisted CAT condition.

- a) The corpus-assisted CAT condition will lead to the passive recall of a significantly larger number (if any) of adj. /noun collocations than non-corpus assisted CAT condition.
- b) The corpus-assisted CAT condition will lead to the active recall of a significantly larger number (if any) of adj. /noun collocations than the non-corpus assisted CAT condition.
- c) The differences between the conditions in active and passive recall (if any) will be retained in a delayed post-test.

H2. The corpus-assisted CAT condition will lead to the learning of a significantly larger number (if any) of adj. /noun collocations than the corpus-assisted non-CAT condition.

- a) The contrastive analysis and translation conditions (both) will lead to the passive recall of a significantly larger number (if any) of adj. /noun collocations than the non-contrastive and translation tasks.
- b) The corpus-assisted CAT condition will lead to the active recall of a significantly larger number (if any) of adj. /noun collocations than the corpus-assisted non-CAT condition.
- c) The differences between the conditions in active and passive recall (if any) will be retained in a delayed post-test.

1.4 Thesis outline

This thesis consists of seven chapters, each of which is briefly introduced below.

Chapter 2 justifies in comprehensive detail the need for this research and the focus of the thesis based on three grounds: (1) the results of a small-scale need analysis study

undertaken in a Saudi EFL context; (2) the importance of formulaic sequences in language learning; (3) EFL learners' underdeveloped knowledge of collocations. It then presents two approaches to defining collocations in the area of lexicology, i.e. the frequency-based approach and the phraseological approach, with a critique of both approaches. The chapter concludes with a definition of collocation from a complementary definitional perspective for the purpose of this research.

Chapter 3 reviews the literature on collocations in relation to second-language acquisition and instruction, and develops a methodological rationale for the purpose of this research. In other words, it attempts to answer the question of 'why employ a corpus-assisted contrastive analysis and translation approach?' After reviewing the literature on instructed SLA and the empirical research on vocabulary and collocation learning, the researcher defines a methodological gap, and proposes a corpus-assisted contrastive analysis and translation approach to learning collocations. The literature on SLA is reviewed with the aim of providing the theoretical underpinnings as to how learning of collocations occurs with corpus-assisted CAT, with a consideration of the features of corpus resources and the affordance of cross-linguistic/ contrastive analysis for FL vocabulary learning.

Chapter 4 provides a detailed account of and justifications for the methodology employed in this study to investigate the learning outcomes of the corpus-assisted CAT group as well as the two comparative ones. The data elicitation methods for the three experimental groups and instruments include extraction of the target non-congruent collocations as determined by the complementary approach. The instruments section also comprises a detailed account on designing the bilingual corpus-data sheets, intervention worksheets, and tests for collocational passive and active knowledge. Additionally, the chapter includes a brief section on the quantitative methods of data analysis followed by a critical discussion of the validity, reliability and ethical issues of the present research.

Chapter 5 goes into elaborate detail of the analysis process, providing justifications for every utilised statistical procedure, and presenting the findings in relation to each experimental condition and to the research hypotheses.

Chapter 6 discusses the findings, outlining the quantitative changes that occurred in learners' collocational knowledge after receiving one of the three experimental interventions. It shows how these findings relate to the existing literature. Most significantly, it provides an evaluation of the three instructional conditions in relation to

different aspects within the research's theoretical framework, thus providing a justification and explanation for the superiority of the results attained by corpus-assisted CAT.

Chapter 7 summarizes the main findings of the study in relation to the research aims. The strengths and limitations of the study are also reflected upon in this chapter, and directions and suggestions for future research are provided. Most prominently, in response to what provided the impetus for the research, pedagogical implications are provided.

Chapter 2: Collocations: Focus of the Thesis

This chapter pinpoints the rationale behind the focus of this thesis i.e. teaching and learning of non-congruent collocations. It also aims to define collocations as used in this research. The current researcher's motivation for examining the teaching and learning of non-congruent collocations is driven by: (1) the results of a small-scale needs analysis study undertaken in a Saudi EFL context; (2) the importance of formulaic sequences in language learning; (3) problems in EFL students' collocational knowledge. Accordingly, sections 2.1, 2.2 and 2.3 present the importance of vocabulary knowledge for language learning, a reflection on this knowledge in the Saudi EFL context, and the taxonomy and the notion of multidimensionality of vocabulary knowledge. Section 2.4 then presents the exploratory needs analysis study. The study was undertaken to narrow the scope of the research to specific vocabulary knowledge construct, using the previous taxonomy as a guideline. Section 2.5 presents the notion of formulaic language as an umbrella term for collocations and the importance this entails for language learning. Section 2.6 and sub-sections 2.6.1, 2.6.2, and 2.6.3 detail issues regarding EFL learners' collocational knowledge as perceived in research involving different elicitation methods. Section 2.7 and sub-sections 2.7.1 and 2.7.2 are presentations of approaches to defining collocations, while sub-section 2.7.3 includes a definition of collocation as employed in this research. This chapter concludes with a summary of the chapters' main points and issues in section 2.8.

2.1 Vocabulary and language learning

Vocabulary learning is only one sub-goal of several important language learning goals in the classroom as observed by Nation (2001). Nation provides the mnemonic LIST to refer to these goals: L= language, which comprises vocabulary; I = ideas, which includes cultural knowledge as well as content and subject matter knowledge; S = skills, which involves accuracy, fluency, strategies and the process of language learning; T = text or discourse, which refers to the way sentences fit together to make larger units. The acquisition of large numbers of words has typically been perceived by second language learners as a vital element of learning that language (Laufer & Hulstijn, 2001). In fact, many learners see second language learning basically as a matter of learning vocabulary (Read, 2000). They also see acquisition of vocabulary as their greatest challenge (Meara, 1980).

A vast knowledge of vocabulary (words) has empirically been proven to be crucial for the mastery of other language skills such as reading comprehension (Haynes & Baker,

1993; Huckin & Bloch, 1993), writing (Laufer, 1998) and listening and speaking (Joe, 1995). Consequently, lack of such knowledge might result in EFL/ ESL learners being incompetent in these receptive and productive language skills. Lack of vocabulary knowledge is also believed to be responsible for communication failure outside the language classroom. Read (2000) points out that, even at advanced levels, second-language learners are aware of the limitations in their knowledge of second-language vocabulary and that these limitations obstruct their ability to communicate effectively in the target language. In other words, they constantly experience 'lexical gaps'. The EFL context in Saudi Arabia is no exception.

2.2 Vocabulary in the EFL context of Saudi Arabia

Research on English vocabulary in Saudi Arabia's EFL context has confirmed the importance of vocabulary to EFL learners as well as the difficulties they encounter in attaining both fluency and an overall English proficiency. This is evident in the literature where many investigatory studies on Saudi students' English proficiency, conducted between 1978 and 1980, showed startling results. Al-Guayyed (1997) commented on the overall average TOEFL results of these students and noted that out of the 474,000 candidates from 143 different countries who applied for the TOEFL in that period, the Saudi students attained the fifth rank from the bottom. The weakness of the Saudi students was apparent in all four language skills covered in the test. Al-Guayyed (1997) partially attributed this weakness in the Saudi test takers to the lack of adequate vocabulary knowledge. Moreover, researchers on different English language skills (e.g. Alfallaj, 1998; Alhammadi, 1998; Almazroou, 1988) claimed that a larger repertoire of vocabulary would have resulted in a better comprehension of test questions and a better performance in all language skills.

Until recently, the problem of vocabulary has continued to be evident in research in the Saudi EFL context. For example, Alqahtani (2009) considered the lack of English vocabulary knowledge as a serious problem for EFL learners in the Saudi context, and emphasised the importance of learning vocabulary for the students' academic achievements in English courses. Additionally, Al-Sugayyer (2006) and Alhawsawi (2013) suggest that EFL learners in high schools and undergraduates in preparatory programmes simply memorise some vocabulary items and explicit grammatical rules. The researchers suggest that this is insufficient to attain reasonable communicative competence, let alone attaining adequate fluency. Al-Sugayyer (2006) emphasised the

probable defects in the learning and teaching processes in the Saudi EFL context. In relation to vocabulary, Albousaif (2011) stressed that the defects in the Saudi students' mastery of vocabulary could be attributed to the mismatch between what language teachers think are the best vocabulary learning and teaching strategies for their students, and those actually used and perceived by the learners to be good. Albousaif (2011) suggested that the Saudi learners are very much teacher-dependent when it comes to learning vocabulary. According to the researcher, this results from a lack of effective vocabulary-teaching methods that would foster autonomy by teachers, and lack of awareness of the importance of autonomous vocabulary learning by students.

The studies reviewed above tend to be quite general regarding their definition of what constitutes vocabulary knowledge and what aspects of vocabulary knowledge would seem to be of greater importance to or more challenging for the learners in this EFL context. They mostly addressed vocabulary knowledge in terms of the size and word repertoire of learners' vocabulary. However, attaining sufficient vocabulary knowledge is more complex than merely learning words. This is because words are not discrete units of language. Rather, there are intertwining systems and levels, and there are many aspects to know about a particular word, with varying degrees of knowing (Nation, 2001). Hence, the following sections will address the multidimensional construct of vocabulary knowledge. Then, a small-scale needs analysis study will be presented in order to narrow the scope of this research and address the vocabulary aspect that might be most useful to attend to.

2.3 Vocabulary knowledge: a multidimensional construct

Vocabulary knowledge, also referred to as word knowledge (Laufer, 1990a; Milton, 2013), lexical knowledge (Laufer & Goldstein, 2004) and lexical competence (Henriksen, 1999), is a complex and multifaceted construct (Daller et al., 2007). What is involved in knowing a word has many interpretations in the literature on foreign language vocabulary teaching, learning and assessment. One very common way of addressing the construct of word knowledge is by dividing it into receptive knowledge and productive knowledge. According to Nation (2001), receptive vocabulary use "involves perceiving the form of a word while listening or reading and retrieving its meaning" (p. 24). Productive vocabulary use on the other hand involves the learner's desire to express a word's meaning through writing or speaking and retrieving and producing its appropriate written or spoken form. Researchers (e.g. Corson, 1995; Laufer & Goldstein, 2004; Laufer & Girsai, 2008a, b;

Meara, 1990) employ the terms passive vocabulary (for reading and listening) and active vocabulary (for writing and speaking) in a synonymous manner to refer to receptive and productive vocabulary.¹ The distinction between receptive/ passive and productive/ active vocabulary knowledge is perceived by some researchers (e.g. Faerch, Haastrup & Phillipson, 1984; Palmberg, 1987; Teichroew, 1982) as being on a continuum. Vocabulary knowledge in a foreign language, in that sense, is defined as "a continuum between ability to make sense of a word and ability to activate the word automatically for productive purposes" (Faerch, Haastrup, & Phillipson, 1984, p. 100). At one end of the continuum, the learners would start with words that they have not come across before, but which they can nevertheless understand when first encountered. Berman et al. (1968, cited in Palmberg, 1987) referred to these words as potential vocabulary. The researchers suggested that as learners move along the continuum, they enter the area of real vocabulary, which comprises those words that the learners have learned at some point in the learning process, and that they can either only understand (passive real vocabulary) or both understand and use (active real vocabulary). One criticism of this continuum-based approach is that in the passive-active word knowledge distinction, the threshold at which receptive knowledge becomes productive, is not clear (Laufer & Goldstein, 2004; Schmitt, 2010).

A second common definition of knowing a word is by making a distinction between breadth of word knowledge and depth of word knowledge (Milton, 2009, 2013). Put simply, breadth of knowledge, sometimes called vocabulary size, refers to the number of words a learner knows (Daller et al., 2007). On the other hand, depth of knowledge refers to the multi-aspect nature of word knowledge and covers a word's relations with other words, i.e. syntagmatic and paradigmatic associations (Henriksen, 1999).² Vermeer (2001) argued against the clear cut distinction between breadth and width of vocabulary knowledge, suggesting that they are interdependent i.e. developing depth in vocabulary knowledge is conditional upon developing vocabulary breadth. Milton (2009, 2013) stresses that simple binary divisions such as breadth and depth, or receptive and productive do not really do justice to the intricacy of word knowledge. Many researchers (e.g. Laufer, 1990a; McCarthy, 1990; Schmitt, 2000) have discussed the notion of word

¹ The terms receptive/passive and productive/active will be used synonymously in this thesis.

² Syntagmatic associations are "associations that complete a phrase (syntagm)" such as hold/ hands (Meara, 2009, p. 6). Paradigmatic associations are "ones in which the stimulus word and the response that it evokes both belong to the same part of speech, nouns evoking nouns, verbs evoking verbs, and so on" such as boy/ girl (Meara, 2009, p. 6).

knowledge, and attempted to create an all-inclusive description of vocabulary knowledge. However, Nation's (2001) proposed description of word knowledge is the most comprehensive (Daller et al., 2007), and the nearest existing definitive list of what is involved in knowing a word (Milton, 2013).

Nation (2001) introduced the notion of word knowledge as the receptive and productive knowledge of a word's form, meaning and use. Each area of knowledge was divided into three sub-divisions (see table 2.1). Each of the sub-divisions in Nation's list is further subdivided into receptive knowledge and productive. Milton (2009, 2013) submits that the receptive and productive distinction fits in well with this model, and it maintains the notion that there is a measurable distinction between these two types of knowledge. On the other hand, the breadth and width distinction is less clearly outlined. Vocabulary breadth would involve the 'form' area, but may also include the form and meaning sub-division from the 'meaning' area (Daller et al., 2007; Milton, 2009, 2013). Vocabulary depth would, by implication, include all the left categories and sub-categories in Nation's table (ibid).

Daller et al. (2007) summarised these aspects of knowledge in a hypothetical three-dimensional 'lexical space'. The researchers added a third dimension to breadth and depth by characterising vocabulary knowledge in terms of automaticity. They called this dimension 'fluency', with which learners would be able to use the words they know and the information at their disposal on the use of these words. This dimension of fluency may involve the speed and accuracy with which a word can be recognised or called to mind in speech or writing. Regarding this theoretical model, Milton (2009) suggests that it lacks detail, but one way of operationalising it is to presume that breadth and depth refer to passive word knowledge, while fluency is an aspect of productive word knowledge a learner has.

Table 2. 1: What is involved in knowing a word

(adapted from Nation, 2001, p. 27) R= receptive knowledge/ P= productive knowledge

Form	Spoken	R	What does the word sound like?
		P	How is the word pronounced?
	Written	R	What does the word look like?
		P	How is the word written and spelled?
	Word parts	R	What parts are recognised in this word?
		P	What word parts are needed to express the meaning?
Meaning	Form and meaning	R	What meaning does this word form signal?
		P	What word form can be used to express this meaning?
	Concept and referents	R	What is included in the concept?
		P	What items can the concept refer to?
	Associations	R	What other words does this make us think of?
		P	What other words or types of words must we use with this one?
Use	Grammatical function	R	In what patterns does the word occur?
		P	In what patterns must we use this word?
	Collocations	R	What words or types of words occur with this one?
		P	What words or types of words must we use with this one?
	Constraints on use (register, frequency, etc.)	R	Where, when, and how often would we expect to meet this word?
P		Where, when, and how often can we use this word?	

Unfortunately, in the EFL learning and teaching context, some of these aspects of vocabulary or word knowledge, such as knowledge of a word's form and meaning on different levels of reception and production, have received great attention, while other important aspects of knowledge of words use such as collocations are rarely mentioned (Souza Hodne, 2009). As Milton (2009) puts it:

“[t]he first sub-division, form and meaning, is the part most of us will think of in terms of knowing a word. It involves being able to link the form, however it occurs, to a meaning, and often in a foreign language this involves forming a link between a foreign language word and its translation in the native language” (p. 14).

Brown (2010) also suggests that this single aspect of vocabulary knowledge (form and meaning) receives by far the most attention in the textbooks, while the other aspects receive little or no attention.

The next sections will examine the aforementioned essentiality of vocabulary learning in higher education in the Saudi EFL context. They will explore the aspects of vocabulary

knowledge being focused on in the teaching context and the challenges faced by the learners.

2.4 A needs analysis study: narrowing the research scope

Language teachers do not always identify the precise learning problems encountered by learners or the learners' needs in a given teaching context. When it comes to vocabulary learning, the famous question by Allwright (1984) "why don't learners learn what teachers teach?" has always been at the back of the current researcher's mind. An abundance of research has been conducted to address different matters in relation to the aforementioned question, suggesting the mismatch between the teachers' agenda and the learners' needs as a fundamental problem in teaching and learning. For example, Lewis (2000) suggests that learners learn what they are ready for and in ways that may or may not match what teachers do. Teachers might be focussing on and addressing aspects of language that might not be problematic for their students, or neglecting aspects that are worth addressing. Hence, in designing their lesson plans, teachers should target those aspects that would meet the students' learning needs. Failure to achieve this goal might result in dissatisfaction, frustration and discouragement for both teachers and learners.

The EFL context in Saudi is no exception. Being a former teaching assistant who taught vocabulary courses (as well as other courses of English language skills) at a higher education institution in Saudi for three consecutive terms, the current researcher has always been frustrated that the students do not seem to learn the taught vocabulary. In this context, 'learn' means the students' ability to both understand the meaning of a particular word, and to use it accurately in speaking and writing. This lack of learning became clear from the unsatisfactory results of the students' vocabulary achievement tests throughout the course. It is quite confusing and misleading to point out and highlight the learners' problems and needs in terms of vocabulary learning without having an insight into both the teaching and learning contexts.

To investigate the present research context, a small-scale exploratory study was conducted. The study aimed at outlining the issues around vocabulary learning by investigating teachers' and learners' views on the following topics:

- Difficulties and problems with vocabulary observed by teachers
- Strategies used in teaching vocabulary
- Difficulties and problems with vocabulary experienced by learners

- Strategies used in learning vocabulary

For this study, semi-structured interviews were conducted with five English language teachers and fifteen students in a university in Saudi Arabia. The learners were first and second year undergraduates majoring in English. A thematic analysis approach (Braun and Clarke, 2006) was adopted in analysing the interview data.

2.4.1 Analysis and findings

a. Teachers

The data obtained from the teachers' responses about students' vocabulary problems show that the teachers were conscious of and concerned about their students' apparent inability to employ the taught words in meaningful sentences or in the appropriate semantic context. For example, a teacher who had taught a vocabulary course for five years reported that students tended to store a lot of the taught vocabulary items in their minds as part of their receptive knowledge simply because they did not know how to use it. She believes that the students may recognise the word forms and understand their meanings when they read, but they might be unsure about how to use the words in speaking or in written work. On that matter, a senior lecturer and language teacher stated: T1. "What is the point of learning words without knowing how to use them!"

The teachers also reported that students are probably unaware about the possible restrictions of using particular words in certain contexts or in combination with other words. For example, T3 stated:

T3. "Students don't stop and think about the appropriateness of using a vocabulary item in the context. They may use the first word that comes to their minds or the first entry in a dictionary."

This implies that the problem also involves word associations or collocation problems. Interestingly, each of the five teachers reported one or two types of collocations such as preposition/ verb, verb/ adjective and verb/ noun collocations that they believe are problematic for students.

According to the interviewed English language teachers, the teaching methods employed to teach vocabulary (words) can be categorized as: a) explaining meanings and synonyms, b) giving examples and c) providing or eliciting translations of words. Presenting the words in different contexts and checking the students' ability to use them is not particularly emphasised during the teaching process. In fact, only two of the five teachers

reported engaging the students in the learning process, which assumedly occurs partially during vocabulary classes. They stated:

T1: “If you give them the meanings of words voluntarily you will have a class of thirty students sitting there without knowing how involved they are, so I ask them to look up words in a dictionary in class to ‘observe’ their use in different contexts and [I] engage them in thinking and communicative activities.”

T4: “I urge them to ask questions about the words and discuss the contexts with them.”

Interestingly, none of the vocabulary teachers has referred explicitly to the teaching of word associates such as collocations as part of their teaching agenda although they were allegedly cognizant about the formerly highlighted problem of vocabulary use in their students’ language production.

Other teachers used words like ‘present’, ‘give’ and ‘tell’ to describe their teaching and ‘ask’ to describe their roles in facilitating the learning. For example:

T8: “I present the vocabulary, explain meanings, give examples then ask for other examples. I also tell the students what preposition goes with what verb etc.”

T4: “We ask them to use flash cards and to keep learning diaries.”

T5: “I ask them to read more and use dictionaries.”

It is worth mentioning that the teachers identified other vocabulary problems encountered by the learners and observed by the teachers such as word derivations and spelling. However, this is considered by four of the five teachers as mainly lexical *mistakes* rather than *errors*, i.e. students are sometimes able to self-correct the mistakes when revising their work.

When responding to a discursive question about the potential reasons for vocabulary problems, teachers mainly reported that students are very dependent on teachers and textbooks as key resources of information and vocabulary knowledge. T2 stated “*they [the students] idealise their teachers, so they [the teachers] become their only source of language and knowledge*”. They also commented that students tend to memorise words rather than learn different aspects of it, and that they tend to learn words in isolation or in only limited contexts.

b. *Students*

Students' responses regarding vocabulary learning difficulties were quite consistent with the teachers' answers. Although some students (5 out of 15) reported spelling as a major problem, the majority of them (10 out of 15) reported that they encounter difficulties using words correctly in contexts, despite their abilities to sometimes recognise their meanings when they encounter them. Recalling memorised vocabulary suitable for a given context was identified as another problem by most of the interviewed students. Examples of some of the elicited responses include:

S1: "I have many vocabularies [sic], but I don't know how to use it."

S2: "I can understand the native speakers, but I can't talk like them."

S3: "Sometimes I don't know if it is suitable to use the word in this sentence or not."

The students' reflection upon their own learning strategies showed that most of them use translation to help them remember and memorise the meanings of words. Students also reported that they use mnemonics and repetition to memorise word spelling and pronunciation. On the teachers' role in facilitating vocabulary learning and their teaching techniques, some of the responses were very spontaneous and extremely interesting. Eleven students summarized the teaching techniques used by teachers as explaining vocabulary meanings and providing translations in Arabic with one example or two.

S8: "The teacher asked us to memorise the vocabulary every week...is there any other way other than memorisation... I don't think so... If there is any other way, I will do it without the help from my teacher."

S9: "We don't need a teacher. It is all about memorising a word."

S7: "She reads the sentence and explains and translates. We don't even have activities."

S11: "The teacher suggested flash cards. It simply does not work."

The four remaining students reported that their teacher gives them a lot of activities, makes them use a dictionary in class and compares meanings of words in English and Arabic.

S14: "Miss X is really good. She makes us use dictionaries. She gives us a lot of homework activities and compares words' meanings in English and Arabic."

When students were asked what they believe is needed to help them overcome the difficulties they reported with vocabulary, only three of them gave some suggestions, including having more vocabulary courses and quizzes to enable them to memorise more

vocabulary and relying more on resources other than the textbooks. The rest of the students were unsure about what to say in response to the researcher's question, as they are apparently unaware of any other ways of learning and developing their vocabulary knowledge.

2.4.2 Discussion

The findings of the interviews conducted with teachers and students regarding vocabulary difficulties and teaching and learning techniques showed a clear mismatch between the learners' needs and the teachers teaching agenda and teaching focus. Considering Nation's (2001) taxonomy of word knowledge and the receptive/ productive distinction, the students in this context seem to be mainly struggling with the productive aspect of a word's use, which was evident from their reported difficulty with vocabulary. In fact, this finding is consistent with the literature on vocabulary learning difficulties in the wider context of EFL. In most models of L2 vocabulary acquisition, receptive knowledge precedes the more complex productive knowledge and use of vocabulary (Laufer, 1998; Meara, 1996; Nation, 1990). A longitudinal study conducted by Laufer (1998) showed that learners' L2 receptive vocabulary developed to a greater extent than their productive vocabulary. The difference in development between receptive and productive vocabulary has been attributed to the lack of production tasks that provides opportunities for using both known and new vocabulary. In the specific context of EFL in Saudi, Al-Jarf (2006) asserts that vocabulary learning and teaching constitutes a major problem for EFL learners and teachers. In her study, Al-Jarf reported that freshman students have difficulties in different aspects of vocabulary knowledge including associating, and using English words. This clearly indicates a struggle in the learners' production of vocabulary meaning and use according to Nation's taxonomy (see table 2.1).

Despite the students' struggle with vocabulary production and use, most of the interviewed teachers did not report much (if anything) about changing their teaching approach to meet the learners' needs. As indicated by the interview data, most of the teachers employed a grammar translation approach to teaching vocabulary. They mainly focussed on form-meaning links in teaching discrete words, while mostly neglecting other aspects of vocabulary knowledge, thus resulting in erroneous language use and production. Zimmerman (1997) affirms that the students' failure in oral and written language usage has one of the worst impacts on the learners' motivation. Despite the attempts made by a few teachers (only two in my research) to encourage vocabulary

production through discussion and communication, these attempts do not seem to be systematic in their objectives and do not seem to encourage profound, progressive and contextualized vocabulary production, let alone raising any collocational awareness or developing any autonomous vocabulary learning skills.

Many researchers (e.g. Henriksen, 1999; Lin, 2002; Liu, 2000) emphasise the importance of converting learners' receptive vocabulary into productive vocabulary. Different suggestions have been made for attaining this shift. For example, in the longitudinal study of Danish learners' acquisition of English adjectives, Haastrup and Henriksen (1998) attempted to trace the participants' L2 vocabulary development along three lexical competence dimensions by collecting a range of receptive and productive performances. By comparing the results on the three dimensions,³ they hypothesised that depth of knowledge of a lexical item is important for precise understanding. They also suggested that rich meaning representation is an important factor for a word to become productive. Thus, they emphasise the strong interrelationships among the three vocabulary-learning continua with an emphasis on the importance of semantic network building. Moreover, Beheydt (1987, p. 57) points out that "the learner has not really semantized a new word until he knows its morphological, syntactic, and collocational profile as well as its meaning potential."

Supporting Beheydt's (1987) observations, Liu (2000) confirms that the more often students are taught English collocations, the more correctly they can make use of vocabulary. Lin (2002) came to the same conclusion while investigating the effects of collocation instruction on students' English vocabulary developments. Lin (2002) found that students made progress in producing vocabulary after receiving explicit instruction on collocations. According to Cowie (1992), English collocations are important in receptive as well as productive language competence. A similar assertion was made by Nattinger (1988). Both researchers suggested that English collocations are useful not only for English comprehension but for English production as well.

Nattinger (1980) states that "language production consists of piecing together the ready-made units appropriate for particular situations, and that comprehension relies on knowing which of these patterns to predict in these situations" (p. 341). Moreover,

³ The partial-practice knowledge dimension, depth-of-knowledge dimension and receptive-productive dimension.

Hussein (1990) states that “without the appropriate use of vocabulary, vocabulary learning is meaningless” (p. 129). According to Hussein, students should observe the restriction on the co-occurrence of words and items within a sentence and heed lexical restrictions. Brown (1974) pointed out that learning collocations enables learners to gradually recognise language chunks used by native speakers in speech and writing and to get a feel for using words in natural combinations with other words as well.

Despite this significance of collocations in converting receptive/ passive knowledge of vocabulary into productive/ active knowledge, the needs analysis data show that this construct of vocabulary knowledge has been neglected. The interviewed teachers did not indicate any emphasis on teaching collocations or raising collocational awareness. Hence, it is most likely that learners in this context, as in other EFL contexts, are lacking the required collocational competence for attaining native-like accuracy (Ellis, 1996) or near-native competency (McCarthy, 1990).

The approach to vocabulary learning used by the students who participated in the interviews mirrors the teachers’ focus in the sense that the students translate discrete words into their L1 and memorise the equivalent meanings. They also use verbal and written repetition to memorise words and their spellings. Although such strategies are reportedly helpful, Schmitt & Schmitt (1993) reported that they seem to fall at the ‘superficial’ end of the processing continuum, thus leading to shallow learning. They suggest that such strategies by themselves are unlikely to result in permanent learning. They state that “some 'deeper' processing is likely to be necessary to stabilize the knowledge and make it available for use in real time” (Schmitt & Schmitt, 1993 p: 32). This brings back the notion of use in Nation’s (2001) taxonomy of word knowledge, collocations in particular, which are indications of word semantization and depth of knowledge as discussed above.

According to Nattinger (1988), collocations can aid learners in committing these words to memory and defining the semantic area of a word (i.e. words with related meanings), and they can permit learners to know and to predict what kinds of words would be found together. He suggests several reasons for teaching lexical phrases. The most important reason is that teaching lexical phrases (collocations with pragmatic functions particularly) will lead to fluency in speaking and writing, primarily because they shift learners’

concentration from individual words to larger structures of discourse and to the social aspects of interaction.

To conclude, the current researcher proposes that the teaching and learning of collocations can establish a connection of form and meaning, and can provide a feasible recipe to facilitate another aspect of vocabulary knowledge, namely word use. In other words, as Nation (1990) states, “teaching vocabulary in collocations is in some ways a reaction against teaching words in lists and is an attempt to learn words in context while keeping the flexibility of list running” (p. 38). Other researchers (e.g. Fan, 2009; Farghal & Obiedat, 1995; Nattinger, 1988) stress that instead of teaching vocabulary as discrete lexical items, which could result in lexical incompetence, learners must be made aware of the necessity of learning collocations.

Taylor (1983) depicts the following reasons for learning words in collocations: (1) words which are naturally associated in text are more easily learnt than those that are not; (2) vocabulary is learned best in context; (3) context alone is insufficient without careful association. In a study by Özgül and Abdülkadir (2012), the researchers compared an experimental group (30 Turkish students), which was taught new words using collocation, to a control group (29 Turkish students) which was taught the same words using traditional techniques such as synonym, antonym, definition and mother-tongue translation. The results showed a significant increase in the experimental group’s learning and retention of the taught vocabulary items as indicated by their performance in a receptive test (fill-in-the-blanks) and a productive test (gap-filling). The researchers concluded that teaching vocabulary through collocations may enhance the receptive and productive retention of new vocabulary items in EFL classes.

The following section addresses the current researcher’s second motivation for examining the teaching and learning of collocations: EFL collocational knowledge.

2.5 Collocational knowledge of EFL Learners

Research examining EFL learners’ knowledge of collocation can be classified into three main categories: (1) corpus-based research; (2) research that used paper-and-pencil elicitation tests; and (3) research that involves the use of psycholinguistic measures. Some of the aforementioned types of research have been used to investigate the use of formulaic language in advanced non-native spoken discourse (e.g. Adolphs & Durow, 2004; Foster,

2001; Oppenheim, 2000). Others have looked at the use of formulaic language in writing (e.g. Granger, 1998; Hasselgren, 1994; Nesselhauf, 2003, 2005). This section will discuss the research of EFL learners' collocational knowledge in relation to each of these classifications.

2.5.1 Corpus-based research

Corpus-based research (also called research based on production data, Nesselhauf, 2005) analyses EFL learners' written output to evaluate the appropriateness of the collocations used. One of the first influential studies under this category is Chi Man-Lai, Wong Pui-Yiu and Wong Chau-ping's (1994) study. The researchers' analysis of collocational inappropriateness of de-lexical verbs (e.g. get, make, do, etc.) was based on a million-word extract from the HKUST (Hong Kong University of Science and Technology) Learner Corpus. The Learners were of intermediate to advanced level of English proficiency with Mandarin as their L1. After a concordance of all forms of each verb was automatically generated, all faulty combinations were identified. This list was then checked against the BBI and other dictionaries, as well as with several native speakers (NS) for more verification, though the researchers did not specify on what basis the collocations were initially classified as faulty. The study concluded that learners often used de-lexical verbs interchangeably; hence they are frequently misused. The researchers also stressed the role of L1 in the production of collocation. Despite the interesting results of this study, the lack of a rigorous comparison between the extracted collocations produced by non-native speakers of English (NNS) to those of native speakers is an evident limitation of this study. Similarly, Hasselgren (1994) only employed native speakers' intuitions as an external norm for identifying errors in the word choices of a group of Norwegian university EFL learners. It was found that EFL learners recurrently use a specific type of lexical item, for which the term "lexical teddy bears" was coined. However, unlike Chi Man-Lai et al.'s study, Hasselgren's study attributed the source of most errors (42%) to the use of wrong synonyms.

Nesselhauf (2003) used native speakers' intuitions as well as idiomatic dictionaries to classify the 213 verb/ noun combinations that were extracted from the German ICLE sub-corpus. Results showed that collocation production is extremely challenging for NNS since 24% of the combinations extracted were not typical according to the classification criteria. The study concluded that, even at an advanced level, the L1 turns out to have a degree of influence on the production of collocations that goes far beyond what previous

small-scale studies have predicted. A common downside of these studies is the lack of a native speakers' corpus as a baseline for comparison.

The native/ non-native baseline comparison is evident in several other studies. For example, Granger (1998) selected one category of intensifying adverb (amplifiers ending in -ly and functioning as modifiers such as in "closely linked" etc.) in order to explore the collocational behaviour of French EFL learners. The collocations were then retrieved from a NNS sub-corpus (International Corpus of Learner English). This data was compared to the same intensifying adverbs in a synthesis of three NS corpora and a similar corpus of writing by advanced French-speaking learners of English. In this study, a similar trend emerged where the overuse of particular word combinations was statistically significant compared to other salient combinations which Granger describes as "safe bets". Additionally, the study concluded that NNS underuse native-like collocations. The possible explanation for this observation as provided by Granger is similar to that of Nesselhauf (2003) and Chi Man-Lai et al. (1994), namely L1 influence. For example, compared to NS, NNS used *completely* and *totally* correctly far more often in their writings than *highly*, due to their direct translational equivalents. In that respect Granger (1998, p. 151) states: "there is evidence that the collocations used by the learners are for the most part congruent and may thus results from transfer from L1." Another possible reason for the overuse of certain combinations is believed to be the salient and frequent use of these combinations in English.

Nesselhauf (2005) investigated the production of verb/ noun collocations by advanced German EFL learners. Nesselhauf based her comprehensive and wide-scale analysis of argumentative essays on the ICLE (International corpus of Learner English) of which 150,000 words were analysed. The extracted 2000 instances of verb/ noun combinations were then checked against dictionaries, the BNC and native intuition for combinability and acceptability. Nesselhauf reached the conclusion that the influence of the learners' L1 is far greater than what earlier small-scale studies had predicted. Durrant and Schmitt (2009) noted shortcomings in this research. They argued that since the analysis comprised the writing of large numbers of learners, it is not clear to what extent the results mask the variability of distribution of collocational categories between different learners. They also claim that the adopted analytical approach does not account for the identification and definition of collocations according to the neo-Firthian tradition, i.e. collocations as

defined according to the frequency-based approach.⁴ Likewise, Laufer and Waldman (2011) compared the use of English verb/ noun collocations in the writing of NS of Hebrew at three proficiency levels with those used by NS. They accumulated a learner corpus that consists of about 300,000 words to be compared with Louvain Corpus of Native English Essays (LOCNESS), a corpus of young adult native speakers of English. The data showed that: (1) NNS at all three proficiency levels produced far fewer collocations than NS; (2) the number of collocations improved only at the advanced level; and (3) errors, mainly those attributed to L1 influence, continued to exist even at advanced levels of proficiency. A shortcoming of this study also seems to be the employed criterion of collocational typicality (i.e. dictionaries) which could comprise limited numbers of phraseologically interesting collocations.

Another series of influential studies following the neo-Firthian tradition were conducted by Siyanova and Schmitt (2008) and Durrant and Schmitt (2009). Siyanova and Schmitt's study 1 (2008) aimed at exploring learners' use of adjective/noun collocations applying frequency/association strength criteria. They compared NNS data (from the Russian ICLE sub-corpus and a small native corpus) with NS data (from the BNC) and found that about 50% of the adjective/noun combinations produced by the NS university students were relatively frequent, strongly associated collocations. The other half of the combinations were creative in nature i.e. not typical collocations (according to the BNC). The usage of collocations by Russian university students did not differ from that by the NS in their frequencies of produced collocations. Accordingly, the researchers concluded that there were no significant discrepancies between NS and NNS in the production of frequent and strongly associated collocations. These results contradict Laufer and Waldman's (2011) finding that natives and non-natives significantly differ in the amount of typical collocations they produce. It is worth noting, however, that the difference in significance of the results of the two studies may be attributed to the criterion of collocational typicality used in each study (dictionaries in Laufer and Waldman's study versus corpus evidence in Siyanova and Schmitt's study).

Durrant and Schmitt (2009) studied the use of collocations by English native and non-native writers, focusing on modifier-noun combinations as they have been defined in the 'frequency-based' tradition. A total of 96 texts were analysed: 24 long NS texts, 24 long

⁴ Discussed in detail later in this chapter, section 2.7.1.

NNS texts, 24 short NS texts and 24 short NNS texts. The study concluded that non-native writers rely heavily on high-frequency collocations, but that they underuse less frequent, strongly associated collocations. It is also consistent with the previous research accounting for the notion that non-native writing lacks idiomatic phraseology, and tends to repeat favoured items.

Despite the diverse approaches in analysing the data and identifying collocations, the majority of these studies have mainly addressed deviations in the use of these collocations between NS and NNS. Other research however, provides a different approach and different insight into collocational knowledge.

2.5.2 Research involving paper-and-pencil elicitation tests

The second type of research on collocations in the EFL context involves the utilisation of paper-and-pencil tests to assess explicit knowledge of collocations. Granger's (1998) second study concluded that the underuse of native-like collocations and the use of atypical word combinations might be attributed to an underdeveloped sense of salience and what constitutes significant collocations. The study involved administering a collocation test to 112 participants, 56 French learners of English and 56 NS of English. Participants were asked to judge the acceptability of 15 adjectives to collocate with 11 amplifiers. Hasselgren (1994, in the second part of his study) reached similar conclusions to Granger in the sense that EFL learners show little variation in using collocations when compared to native speakers. In a third significant study, Bahns and Eldaw (1993) investigated German advanced EFL students' productive knowledge of English verb/noun collocations in a contextualised translation task and a cloze task. In the translation task, it was found that despite the collocations constituting less than a quarter of the total number of lexical words, more than half of the unacceptably translated lexical words were collocates. Thus, the researchers concluded that collocations present a major problem in the production of correct English even for advanced EFL learners, and that their collocational knowledge lags far behind their general vocabulary knowledge.

In different set of studies which involved Arab learners, collocational knowledge was also shown to be rather weak in explicit paper-and-pencil tests. Hussein (1990) assessed 200 Jordanian English majors' knowledge of 40 common collocations, using a contextualised MC test. The study's results showed unsatisfactory performance when it comes to collocational recognition (48% correct answers). Hussein (1998) replicated his previous

study with 50 students majoring in English at the Applied Sciences University, Amman. The findings of the 30-items test revealed that of the total number of collocations, only 39% were rendered correctly. In both studies, Hussein attributed the lack of collocational competence by EFL learners to different factors, but primarily to L1 influence and negative transfer. Farghal and Obiedat's study (1995) also aimed at assessing the collocational knowledge of Jordanian English majors. They administered a cloze test to group 1 and an L1-L2 translation test to group 2. Both groups showed weak collocational knowledge (18% answers in group 1 vs. 5% in group 2). However, these studies suffer from a number of serious limitations and problems. As Gyllstad (2007), Durrant (2008) and Sonbul (2012) accurately pointed out, these studies did not control either for the frequency of the selected collocations or for the adequacy of clues in context. Additionally and most notably, these studies did not consider proper native baseline data for comparison.

Another recent study with the same drawback was conducted by Brashi (2009). The study aimed at investigating the receptive/ productive verb/ noun collocational knowledge of 20 senior undergraduates majoring in English. The administered tests were a 'fill-in-the-blanks test' and a 'multiple-choice test'. The results showed that the participants performed better at the receptive level (MC) than at the productive level (fill in the blanks). The researcher ascribed these findings to a lack of native-like knowledge of English collocations, L1 influence and to the congruence of English/ Arabic collocations. In addition to suffering from the same problems as the previous studies, this study is rather small-scale. It is therefore not clear whether the percentages attained in these studies really represent Arab EFL learners' weak collocational knowledge or whether this is just a result of improper item selection.

In a more recent study, Noor and Adubaib (2011) have elicited the productive knowledge of English lexical collocations of 88 Saudi English-major students at Taibah University using a fill-in-the-blank test and a contextualised translation (Arabic/ English) test. Specialized dictionaries of collocations, native speakers' intuitions and corpus consultations were used to judge the acceptability of collocations. It is worth noting that investigating the learners' collocational knowledge was not the researchers' primary aim in this study. Rather, they intended to investigate their collocation production strategies. However, the elicitation instruments still showed results that are consistent with other research on EFL collocational knowledge in the sense that both high and low proficiency

students encountered difficulties in the production of acceptable English lexical collocations in general. The study also argues that although L1 influence and negative transfer were responsible for learners' collocational problems, there are other important intralingual factors at play.

2.5.3 Research involving psycholinguistic measures

The final, and most recent, line of research examining EFL learners' collocational knowledge entails the use of psycholinguistic measures. For example, Yamashita and Jiang (2010), employed a recognition, whole-collocation acceptability-judgment task to assess the processing of congruent (L1= L2) versus non-congruent (L1≠L2) English collocations among 28 advanced Japanese ESL speakers and 20 native speakers of English. Native speakers did not show any significant difference between the two collocation categories either in response time or error rate. The NNS made more errors with non-congruent collocations than they did with congruent collocations, but their response time was not different between the two categories. Yamashita and Jiang concluded that L2 learners are dependent on the L1 mediation process at first and that

“...it takes longer for incongruent collocations to be accepted as legitimate in the L2 mental lexicon compared with congruent collocations, but once accepted, incongruent collocations (at least short ones) may construct holistic units and may be processed as wholes without going through word by- word L1 mediation” (p. 130).

This result regarding advanced ESL learners, although it may sound plausible, is not conclusive when it comes to less advanced EFL/ ESL learners' processing of non-congruent collocations.

A similar study by Wolter and Gyllstad (2013) also employed an acceptability judgment task to investigate the influence of frequency effects on the processing of congruent (collocations that have equivalents in the learners L1) and non-congruent collocations in a second language. The task was administered to native and advanced non-native English speakers (L1 Swedish) to assess response times and error rates for 80 collocations along with a matched set of 80 non-collocational items. The results of the study suggest that advanced learners are highly sensitive to frequency effects for L2 collocations. It also plausibly suggests that the L1 may have a substantial impact on how rapidly collocations are processed in an L2. In this regard, the researcher stated that “[a] The only significant difference in RTs [response times] between the NS and NNS groups was for the *non-*

congruent items, (b) only the NNS group responded significantly faster to the congruent items over the incongruent items, (c) only the NNS group produced significantly more errors on the incongruent items when compared to the congruent items” (p. 22).

While the previous studies measured the explicit knowledge of collocations through acceptability judgment tasks, Wolter and Gyllstad (2011) utilised the collocational priming⁵ paradigm to assess implicit knowledge of congruent collocations, incongruent collocations, and control non-collocational items, and utilised a test of receptive collocational knowledge to assess the explicit knowledge of the same sets of collocations. The study involved two groups, native English speakers and EFL students (L1 Swedish). Similar to the previous studies’ results, native speakers’ performance suggested that there was a clear processing advantage for both types of collocations over control pairs, and that they did not show any differences between congruent and incongruent collocations in both tests. Non-native speakers’ performance, on the other hand, showed that there was an advantage for congruent collocations over non-congruent collocations and control pairs. Thus, the researchers reached a tentative conclusion that the L1 seems to have an influence on EFL learners’ processing of collocations.

The next section examines the role of formulaic sequences in language learning as the third reason behind this researcher’s motivation for examining the teaching and learning of collocations.

2.6 The role of formulaic sequences in language learning

“One important component of successful language learning is the mastery of idiomatic forms of expression, including idioms, collocations, and sentence frames (collectively referred to here as formulaic sequences).”

(Wray, 2000, p. 463)

Since the shift from Chomsky’s (1965) generative theory,⁶ a large body of research has directed its attention to lexical studies. Phraseology in particular has emerged as a promising area of research. Bolinger (1979) was among the pioneer linguists who questioned the generativists’ views of language learning, which as he points out fails to

⁵“The tendency for an activated word to accelerate subsequent recognition of a collocate” (Wolter and Gyllstad, 2011, p. 431).

⁶“The workings of a language can be explained by a system of rules of general acceptability” (Cowie, 1994, p. 1).

account for a significant part of observable language data. Likewise, Pawley and Syder (1983) affirmed that sounding native is not only related to knowledge of grammatical rules, but also entails knowledge of acceptable sequences. With the development of studies in corpus linguistics, data from such studies revealed that formulaicity is a pervasive phenomenon in language use (Foster, 2001). According to Erman and Warren (2000), formulaic sequences of different types constitute more than half of the written discourse they analysed, suggesting that in a text of 100 words on average only 45 single-word choices would be made. They also suggest that 58.6% of spoken discourse consists of formulaic sequences. This assertion fits well with Sinclair's (1991) proposed two principles to explain how meaning is conveyed in texts: the open-choice principle and the idiom principle. The open-choice principle views a text as resulting from a very large number of complex choices in which a series of slots have to be filled from the lexicon while satisfying grammatical restraints. In the idiom principle, on the other hand, Sinclair stresses the idea that language users have available to them a large number of semi-preconstructed phrases that constitute single choices, even though they seem to be analysable into segments. Similarly, Moon (1997) contrasted the traditional syntactic model which observes well-formedness, and which is generally built on grammatical principles, with what she called the "collocationist model" which takes into account considerations such as the predictability of the co-occurrence of words in the slots that comprises the underlying structural frame.

Accordingly, several important roles have been identified for formulaic sequences in language learning. First, formulaic sequences are believed to be the basis for the development of creative language in the first language (Peters, 1983) and childhood second language acquisition (Wray, 1999). In addition, it is now widely acknowledged that in order to attain native-like fluency, second language learners need to be in control of formulaic sequences in the L2 (Ellis, 1997). In fact, Moon (1997) suggests that the appropriate use and interpretation of formulaic sequences, or what she calls "multi-word items", by L2 speakers is a sign of their proficiency. On the contrary, lacking the appropriate knowledge of formulaic sequences might put the learners in a situation where they sound arrogant or disrespectful (Wray, 2002) as the appropriate native-like sequences follow conventions of politeness (Moon, 1997). More importantly, formulaic sequences serve two key functions in language, saving processing effort and achieving communicational and interactional functions (Moon, 1997; Schmitt & Carter, 2004; Wray, 1999; Wray, 2000).

Eventually, one would ask “but what are formulaic sequences?” Wray (2000) observes that a full understanding of what formulaic language is requires researchers to recognise that they are not dealing with a single phenomenon, but with a set of more or less closely related ones across research of different principles and types of data. Formulaic language as observed in such studies has been defined in different ways, resulting in a huge set of definitional and descriptive terms. Wray (2000, 2002) listed fifty different terms used in the literature to refer to the formulaic language phenomenon (e.g. composites, chunks, collocations, formulae, fixed expressions, multiword items/ units, lexical(ised) phrases and ready-made expressions). Ultimately, she presented the term ‘formulaic sequence’ as an umbrella term to include the wide range of phenomena variously labelled in the literature. Wray (2000) defined formulaic sequence as “a sequence, continuous or discontinuous, of words or other 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. 465).

Since the focus of the present thesis is on one category of formulaic sequence i.e. collocations, the following section and subsections address the notion of collocation and how it is defined in the literature. It concludes with a working definition of collocation as used in this research.

2.7 What is a collocation?

The word ‘collocation’ comes from Latin *collocatio* (n-), from *collocare* which means in a technical sense ‘to place together’, or ‘the action of placing things side by side or in position: the collocation of the two pieces’ (Oxford English Dictionary, Online). Linguistically, collocation is defined by the Longman Dictionary of Contemporary English (Online) as “the way in which some words are often used together, or a particular combination of words used in this way: ‘Commit a crime’ is a typical collocation in English.” Whereas the Oxford English Dictionary (online) comprises a slightly extended definition of collocation:

- “The habitual juxtaposition of a particular word with another word or words with a frequency greater than chance: the words have a similar range of collocation.
- A pair or group of words that are habitually juxtaposed: ‘strong tea’ and ‘heavy drinker’ are typical English collocations.”

The definitions of collocation in both dictionaries provide a broad sense of what collocation as a linguistic phenomenon is, and present parts of its characteristics i.e. the habitual and frequent co-occurrence. However, there is far more to defining and characterizing collocation than what a dictionary definition constitutes. In fact, there are two distinct approaches to defining collocations, the frequency-based tradition and the phraseological tradition (Barfield & Gyllstad, 2009; Nesselhauf, 2003, 2005). The next section will discuss the two approaches in more detail and clarify the differences between them in the identification of collocations.

2.7.1 The frequency-based approach

Collocation as a term was first used in its linguistic sense by British linguist J.R. Firth (1890-1960), who famously observed, “You shall know a word by the company it keeps.” Collocations are defined by Firth (1957, p. 4) as “actual words in habitual company” with reference to the significant role of collocation not only to applied linguistic research but also to that of grammar, phonetic and phonology. Collocation in the Firthian sense could be interpreted as empirical statements about the predictability of word combinations (Evert, 2008). The rather vague notion of collocation by Firth has later been significantly developed by a group of British linguists (e.g. Halliday, 1966 and Sinclair, 1991), often referred to as the Neo-Firthian school. According to Sinclair (1991, p. 170) collocations are “the occurrences of two or more words within a short space of each other in a text.” This space or span is usually, but not exclusively, defined as a distance of four words to the left and right of the ‘node’. Nesselhauf (2005) explains Sinclair’s node collocations principle by stating:

“If, for example, in a given amount of text, the word *house* is analysed, and the word occurred in an environment such as *He went back to the house. When he opened the door, the dog barked*, the words *went, back, to, the, when, he, opened, the*, are all considered to form collocations with the node *house*; these words are then called collocates” (p. 12).

Sinclair distinguishes two types of collocations, causal collocations and significant collocations. In reference to the previous example, the words *dog* and *barked* are considered significant collocations as they co-occur more often than their respective frequencies and the length of text they appear in would predict. The concept of co-occurrence of words has varied across studies and been approached differently by researchers. While some researchers adopting the frequency-based approach to collocations consider co-occurrences of all frequencies as collocations (e.g. Moon, 1998), others reserve the concept for ‘frequent’ co-occurrences (e.g. Carter, 1988; Stubbs, 1995).

For example, Carter (1988) defines collocations as “an aspect of lexical cohesion which embraces a ‘relationship’ between lexical items that regularly co-occur” (p. 163). Hoey (1991) on the other hand refers to textual co-occurrence in his definition of collocations: “the relationship a lexical item has with items that appear with greater than random probability in its (textual) context” (p. 7). This variety of identifications of collocations under the umbrella of habitual co-occurrence of words seems to add to the confusion of what constitutes a collocation. Hence, Evert (2008) who belongs to the Neo-Firthian school of defining collocations, introduced what seems to be a comprehensive and precise definition of the co-occurrences or “nearness” of word tokens for the purpose of operationalising the notion of collocation. Evert identified three types of co-occurrences: surface, textual and syntactic co-occurrences.

Surface co-occurrence as identified by Evert (2008) primarily means looking for collocates within the collocational span around the instances of a given node word, though not always combined with a node-collocate view. Span size is the most crucial choice that a researcher has to make. Many span sizes can be found in the literature, however Sinclair’s (1991) suggestion of three to five words is the most common (Evert, 2008). The following figure shows surface co-occurrences of the words *hat* and the collocate *roll* in a span size of 4 words, limited by sentence boundaries and excluding punctuation.

A vast deal of coolness and a peculiar degree of judgement, are requisite in catching a hat. A man must not be precipitate, or he runs over it; he must not rush into the opposite extreme, or he loses it altogether. [...] There was a fine gentle wind, and Mr. Pickwick’s hat rolled sportively before it. The wind puffed, and Mr. Pickwick puffed, and the hat rolled over and over, as merrily as a lively porpoise in a strong tide; and on it might have *rolled*, far beyond Mr. Pickwick’s reach, had not its course been providentially stopped, just as that gentleman was on the point of resigning it to its fate.

Figure 2. 1: Illustration of surface co-occurrence for the word pair (*hat*, *roll*) in Evert (2008, p. 13).

The arbitrary choice of the span size is one criticism against surface co-occurrence. For a span size of 3, *throw*, *party* would be accepted as co-occurrence in a sentence like *throw a birthday party*, but would not in a sentence like *throw a huge birthday party*. This in Evert’s (2008) view is particularly counterintuitive for languages with somewhat free word order where closely associated words can be found far apart.

Textual co-occurrence is a second approach which considers words to co-occur if they appear in the same textual units such as utterances or sentences (Evert, 2008). Textual co-

occurrence is easier to implement than surface co-occurrence and particularly useful in applications such as term clustering in entire documents. One limitation of textual co-occurrences is that it captures weaker dependencies, especially those resulting from paradigmatic semantic relations. For instance, in a sentence that comprises the word *bucket*, it is very likely that the word *mop* would exist too. Although the connection between *bucket* and *water* or *spade* is far stronger than *mop*, they might not necessarily be near each other in the sentence. This type of co-occurrence also tends to generate huge data sets of recurrent word pairs that could be challenging even for advanced computers (Evert, 2008).

A vast deal of coolness and a peculiar degree of judgement, are requisite in catching a <u>hat</u> .	hat —
A man must not be precipitate, or he runs <i>over</i> it ;	— over
he must not rush into the opposite extreme, or he loses it altogether.	— —
There was a fine gentle wind, and Mr. Pickwick's <u>hat</u> rolled sportively before it.	hat —
The wind puffed, and Mr. Pickwick puffed, and the <u>hat</u> rolled <i>over</i> and <i>over</i> as merrily as a lively porpoise in a strong tide ;	hat over

Figure 2. 2: Illustration of textual co-occurrence for the word pair (*hat*, *over*) Evert (2008, p. 14).

The frequency-based approach was criticised for being quite negligent of the syntactic relationship between words and whether or not they form collocations (Nesselhauf, 2005). However, many researchers have actually adopted an approach to defining collocations which to a great extent is bound by syntactical relations between word pairs (e.g. Bartsch, 2004; Evert 2004, 2008). This more restrictive approach to defining word co-occurrence is called syntactic co-occurrence, in which words with a direct (e.g. a verb + its subject or object nouns) or sometimes indirect (e.g. a verb + adjectival modifier of its noun) syntactic relation occur near each other (Evert, 2008). Unlike surface co-occurrence, syntactic co-occurrence does not set an arbitrary distance limit and is particularly appropriate if there is a long-distance dependency between collocates. In addition, syntactic co-occurrence is often used for multi-word extraction, since many types of lexicalised multiword expressions tend to appear in particular syntactic patterns (Bartsch, 2004). It discards many accidental and indirect word occurrences and thus it becomes easier to find suitable association measures to quantify the collocability of word pairs (Evert, 2008). It is worth noting that the notion of frequency of syntactic co-occurrence actually approaches the phraseological view of collocations, however, lexical restrictions between word pairs do not count in this approach.

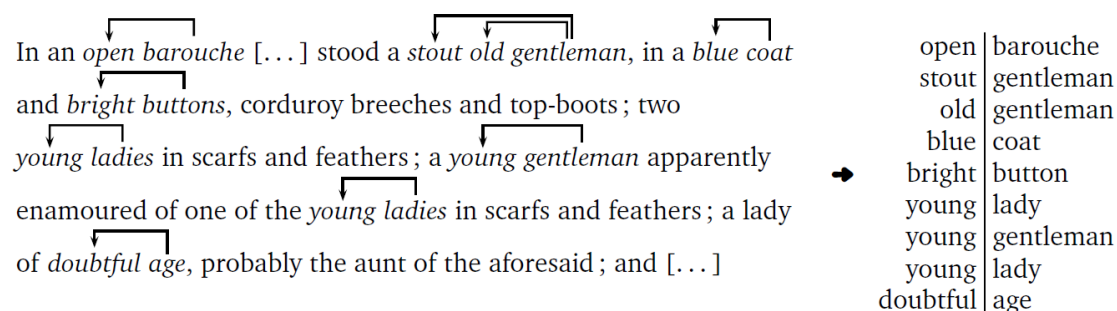


Figure 2. 3: Illustration of syntactic co-occurrence (nouns modified by prenominal adjectives) in Evert (2008, p. 15).

No matter what type of co-occurrence is used to operationalise collocability of words, collocability still needs to be quantified by mathematical association measures (Evert, 2008). Likewise, Stubbs (1995) observed that frequency of co-occurrence is not enough in identifying collocations and hence other measures of association strength are needed. In addition, Hunston (2002, p. 68) states that “collocation may be observed informally in any instance of language, but it is more reliable to measure it statistically, and for this a corpus is essential.”⁷ Thus, a brief discussion about different measures and the importance of these will be introduced in the next section.

2.7.1.1 Statistical measurements of collocations

Any program which calculates collocation takes a node word and counts the instances of all words occurring within a particular span, as noted in the previous section. This is called a list of raw frequencies, which can be displayed in order of frequency, in the order of the first occurrence of the type in the corpus, or in alphabetical order (Barnbrook, 1996). The problem with a list of raw frequency is that it does not give information on other aspects of word co-occurrence patterns (Stubbs, 1995), and it is thus not possible to attach a degree of significance to any of the figures in it (Hunston, 2002). According to Stubbs (1995), many statistical calculations compare the frequency of observed occurrences (*O*)⁸ to the expected frequency (*E*)⁹ (merely by chance) of a given pair of words in a hypothetical corpus consisting of the same words in random order.¹⁰ The pair is only considered to be a collocation if the observed co-occurrence frequency is higher than the

⁷A corpus, according to Sinclair (1996), is “a collection of pieces of language that are selected and ordered according to explicit linguistic criteria in order to be used as a sample of the language.”

⁸ The observed frequency of occurrence is the actual frequency of occurrence of a given combination of words.

⁹ The expected frequency of occurrence is based on the null hypothesis that there is no relationship between the words (Schmitt, 2010).

¹⁰ The concept of randomness is considered by many researchers (e.g. Evert, 2008; Stubbs, 1995) as somewhat bizarre when applied to language, as words do not occur randomly.

expected frequency (Evert, 2008). While the standard formula $E = f1 f2 / N$ ¹¹ can be used directly to calculate the expected frequency for textual and syntactic co-occurrences, an additional factor k representing the span size is used in the expected frequency for surface co-occurrence following the formula $E = k f1 f2 / N$ (Evert, 2008, p. 18).

Many types of statistical measurements have been introduced in the literature (e.g. Clear, 1996; Stubbs, 1995, Evert, 2008) to quantify the attraction degree between a pair of words based on the comparison between observed co-occurrence frequency and the expected frequency. Two of the most commonly used measures of significance and strength of word association are: Mutual Information (MI) score and t-score. According to Schmitt (2010), both measures compute the likelihood of two words occurring together as opposed to the likelihood of their occurring separately. However, they belong to two conceptually different approaches to making these calculations. Mutual Information (MI) comes from work in information theory, where ‘information’ is restrictedly used to mean an event which occurs in contrary proportion to its probability (Stubbs, 1995). It compares the actual co-occurrence of the two items with their expected co-occurrence if the words in the corpus were to occur in totally random order. As Stubbs (1995) observed, MI is a simple variant of O/E . It employs the following formula: $MI = \text{Log} - \text{likelihood}_2 \frac{O}{E}$.¹² With a span size of 2:2 or 3:3, an MI score of 3 or higher can be taken to be significant or “linguistically interesting” as Clear (1994) puts it. Stubbs (1995) argues that there is no strong theoretical reason for determining this value for MI, however, in empirical analysis of corpus data, this value has been shown to generate sets of semantically related words such as *ballpoint pen*, *hardly surprising* etc.¹³ Stubbs adds that although the term “linguistically interesting” is admittedly undefined, it still represents an empirical claim. Moreover, the value of an MI score is not predominantly dependent on the size of the corpus. Thus, MI scores can be compared across corpora, even if the corpora are of different sizes (Hunston, 2002; Evert, 2008).

On the contrary, t-score is “a measure of certainty of collocation i.e. how certain we can be that the collocation is not merely the result of the vagaries of a particular corpus” (Hunston, 2002, p. 73). It belongs to a set of ‘hypothesis testing’ strength of association measures (e.g. z-score, chi-squared and log-likelihood tests) which measure the utterance

¹¹ $f1$ stands for the frequency of the first word component in the corpus, $f2$ for the frequency of the second word, and N for the corpus size.

¹² cf. Schmitt (2010) for detailed information on log-likelihood.

¹³ Examples from Hunston (2002).

frequency of collocations. The t-score picks out many joint occurrences, thus it provides confidence that the association between node (n) and collocate (c) is genuine i.e. the combination of words appears together no more frequently than we would expect by chance alone (Stubbs, 1995, Schmitt, 2010). It is calculated as follows: $t\text{-score} = \frac{O-E}{\sqrt{O}}$. For a t-score to be linguistically significant, it normally needs to be 2 or higher (Evert, 2008). Unlike for MI score, corpus size is important for the t-score, because of the amount of evidence that is being taken into consideration. This means that the larger the corpus is, the more significant a large number of co-occurrences, and that an absolute t-score cannot be compared across corpora due to the potential effect of the corpus size on the t-score (Hunston, 2002).

Hunston (2002) provides more comparison between the MI score and the t-score in relation to the behavioural information and restriction of co-occurrence that both scores present. She suggests that looking at the top collocate from the point of view of the t-score has the tendency to provide information about the grammatical behaviour of a word. Conversely, observing the top collocates from the point of view of the MI score has the tendency to provide information about its lexical behaviour, particularly about more fixed or idiomatic co-occurrences such as *unflinching/ unblinking gaze*. Hunston (2002) also suggests that collocates with the highest t-scores are typically frequent words that collocate with a variety of items (e.g. *followed* collocates with *gaze* and a variety of other words). Collocates with the highest MI scores are usually less frequent words with restricted collocation such as the word *avert* which is closely associated with *gaze* and with only a limited number of other words such as *danger*. Despite the significance of the information that both measures provide, Hunston stresses that calculations of MI scores and t-scores should be carefully interpreted.

It is worth mentioning that one drawback of the frequency-based approach (especially the approaches adopting surface and textual co-occurrence) is its tendency to result in linguistically uninteresting combinations such as ‘*children toy*’ which frequently co-occur according to logical rather than any linguistic attraction (Hunston, 2002). Another disadvantage of the frequency-based approach is highlighted by Wray (1999) and Wray and Perkins (2000). Although they acknowledged that there is indeed “some sort of” relationship between frequency and formulaicity, in the sense that formulaic output is frequently called upon, and that some formulaic sequences are very frequent, they

suggested that formulaic sequences (including collocations) cannot be defined in terms of frequency alone. This is because many sequences which would be identified as formulaic for other reasons, are not at all frequent in general usage (Wray, 1999; Wray & Perkins, 2000). In that sense Howarth (1998, p. 27) has previously stated:

“The mental lexicon clearly holds more abstract entities than are identified by computational searches, and neither native speakers nor learners produce word combinations on the basis of their frequency and probability of co-occurrence.”

He also adds that a notion of significance based solely on frequency risks placing unwarranted weight on completely transparent collocations such as *have children*, which may occur frequently as a result of the topics of certain texts but are pretty unproblematic for processing. Hence, the concept of phraseological significance needs to take into consideration differences between phraseological types, and to account for the way they are processed in production by native and non-native speakers as well as by writers.

The aforementioned shortcomings of the frequency-based approach necessitate the application of the second, phraseological, qualitative approach to defining collocations.

2.7.2 The phraseological approach

In contrast with the statistically-oriented approach i.e. the frequency-based approach to defining collocations, Herbst (1996) also introduced what he referred to as the ‘significance oriented approach’ i.e. the phraseological approach. The phraseological approach has been greatly influenced by Russian phraseology, particularly in East European phraseological theory (Cowie, 1994). Advocates of this approach are interested in the analysis of what is called ‘phraseological units’ or ‘word combinations’ as well as the increasing awareness of the pervasiveness of ready-made memorised combinations in spoken and written language. Their interest was also driven by a wider acknowledgment of the significant part collocations play in first and second-language acquisition and adult language production (Cowie, 1998; Pawley & Syder, 1983). Among the main representatives of this approach are A.P. Cowie, I. Melcuk, F.J. Hausmann and R. Moon. Cowie (1994), the main advocate of this approach, considers collocations as a type of word combination, and defines collocations as “a composite unit which permits the substitutability of items for at least one of its constituent elements (the sense of the other element, or elements, remaining constant)” (Cowie, 1981, p. 224). According to Cowie, word combinations can be divided into two main types, formulae and composites. Expressions with mainly pragmatic functions such as *Good morning* or *You can say that*

again were classified as ‘formulae’. Collocations on the other hand were classified as ‘composite’ and described as primarily having syntactic functions.

Cowie’s classification of word combinations or ‘composites’ was based on two criteria, the criterion of substitutability and the criterion of transparency. Commutability or substitutability¹⁴ refers to the possibility or the degree to which the substitution of the words in the combination is restricted. Transparency refers to whether a word in the combination or the combination as a whole has a literal or non-literal meaning.

Many categorizations of word combinations have been devised following Cowie’s two criteria. However, Howarth’s (1998) classification is the most inclusive one as it draws on different works in language processing (Bolinger, 1976; Pawley & Syder, 1983), and lexicography (Cowie, 1981). His classification is as follows:

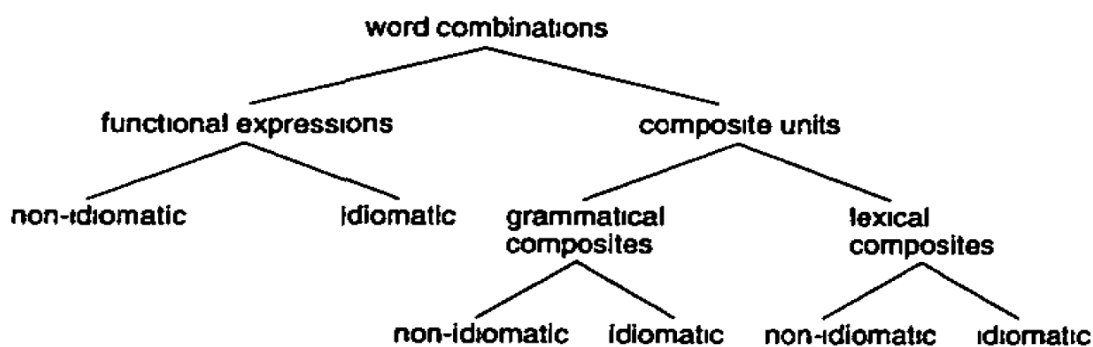


Figure 2. 4: Phraseological categories
Howarth (1998, p.27)

Howarth (1998) then distinguishes four types of composites forming a continuum from less to more restricted combinations: free combinations, restricted collocations, figurative idioms and pure idioms. Free combinations (also referred to as free collocations e.g. *blow a trumpet*) are those combinations in which words can be freely substituted and in which these words are used in their literal sense. Restricted collocations (e.g. *under attack*) are the combinations in which the substitution of words is bound to arbitrary limitations and in which one word has a literal meaning while the other is used in a non-literal sense, but the meaning of the whole combination remains transparent. Figurative idioms (e.g. *under the microscope*) refer to the combinations in which substitution is rarely allowed and which have figurative meaning that can also correspond to literal interpretation. Pure

¹⁴ Commutability, substitutability and restrictedness are used synonymously and alternatively in this section according to the literature they appear in.

idioms (e.g. *blow the gaff*) do not allow any substitution and have a purely figurative meaning.

	free combinations	restricted collocations	figurative idioms	pure idioms
Lexical composites verb +noun	blow a trumpet	blow a fuse	blow your own	blow the gaff
Grammatical composites preposition + noun	under the table	under attack	under the microscope	under the weather

Figure 2. 5: Collocational continuum
(adapted from Howarth, 1998, p.28)

While Howarth (1998) subcategorizes grammatical and lexical composites according to collocational restrictedness and semantic opacity or transparency, Benson, Benson, and Ilson (1997) subcategorize lexical and grammatical collocations on the basis of the constituents' word class. They identify seven types of lexical collocations and eight types of grammatical collocations. Lexical collocations are combinations of content words, such as verbs, nouns, adjectives or adverbs. Grammatical collocations consist of a content word and a grammatical word or structure like a preposition, infinitive or clause. Other researchers (e.g. Nesselhauf 2003, 2005) have adopted a more inclusive classification of collocations under syntactic characteristics (constituents' part of speech), semantic characteristics (sense restrictions), and commutability of elements (substitution of one or both elements).

Type	Example
Verb+ Noun	Read the book.
	Write a letter.
	Adhere to the rule.
	Unload the car.
	Dye hair: she dyed her hair red.
Adjective+ Noun	Hard book Red hair
	Noun+ Verb
Noun+ Verb	Bees buzz
Noun of noun	A pack of dogs
Adverb+ Adjective	Deeply moved
Verb+ Adverb	Argue heatedly

Figure 2. 6: Types of lexical collocations
Benson et al. (1997) adapted from Tsai (2011, p.25).

It is worth mentioning that Cowie and other researchers adopting the phraseological approach vary widely in their use of the term 'collocation', mostly in terms of

restrictedness. Thus, while some researchers use the term to refer to both free and restricted collocations, others exclusively use the term to refer to restricted collocations (Nesselhauf, 2005). For example, Hausmann (1984, cited in Van Der Meer, 1998, p. 133) defines collocations as “typical, specific and characteristic relationships between two words.” Hausmann emphasises that obviously not all combinations qualify for the term collocation. Therefore, he believes that a “banal” combination like *buy a book* is not truly a collocation, seemingly because it is not “typical, specific and characteristic” enough. Similarly, according to Benson et al., collocations are loosely fixed combinations between idioms and free combinations such as *commit murder* (Benson et al. 1986). In addition to collocations in their “loosely fixed” sense, Benson et al. also identify what they call transitional combinations/ collocations (e.g. *to catch one’s breath*) which are more “frozen” than ordinary collocations. Hence, collocations according to them are “fixed, identifiable, non-idiomatic phrases and constructions” (Benson, et al., 1997, p. xv).

Unlike the previous approach which employs frequency and statistical measurements as criteria to identify collocations in a given data set, the phraseological approach mainly uses either natives’ intuitions (Greenbaum, 1988; Hasselgren, 1994), collocational dictionaries (Laufer & Waldman, 2011), or a combination of both (Nesselhauf, 2003). These means of identification were criticised by Stubbs (1995) as being a limitation of the phraseological approach. Stubbs claims that native speakers’ intuitions, though interesting, are not a reliable source of evidence on collocational restrictions, as native speakers can provide some examples of collocations but cannot give accurate frequency estimates.

In an opposing view, Howarth (1998) acknowledges the important role of a pragmatic combination of published collocational dictionaries and (increasingly) large corpora in providing substantial amounts of data. He also emphasises the significance of recent technological developments in automatic lemmatization, tagging, and parsing, which have enabled computational processing to identify collocations at the required abstract, lexemic level. However, Howarth (1998, p. 29) asserts that “decisions about the acceptability of combinations that occur individually at very low frequencies must continue to rely heavily on human judgement.” Howarth argues that the absence of a potential combination from dictionaries and even large corpora cannot equitably exclude it from consideration. He also stresses that the collocations of most interest for studying acquisition are not usually fixed enough for automatic identification.

2.7.3 A working definition of collocation: a complementary approach

The frequency-based approach and the phraseological approach are sometimes mixed when authors who mainly adopt the phraseological approach consider frequency as an additional defining criterion (e.g. Benson et al., 1986), and vice versa (e.g. Nesselhauf, 2005). For example, Evert (2004, 2008) is a strong advocate of semantic co-occurrence as a defining criterion. Evert (2008) also stresses the close connection and the occasional overlap between the two approaches. With her working definition of collocation, Bartsch (2004) interestingly takes a middle road between the two approaches. She defines collocations as “lexically and/ or pragmatically constrained recurrent co-occurrences of at least two lexical items which are in a direct syntactic relation with each other” (Bartsch, 2004, p. 76). Thus, the two approaches to defining collocations outlined above should not be viewed in opposition but rather as complementary. An abundance of collocations identified through corpus analysis have phraseological significance on the one hand, and on the other hand, a lot of collocations with phraseological significance will stand out in corpus analysis (Sonbul, 2012). Accordingly, the present thesis will consider a fusion between the two approaches as a complementary working definition of collocation. The term ‘collocation’ is operationalised here as: “A pair of two open-class lemmas which occurs in a corpus (within a window of ± 3) above chance ($f > 5$ and $MI > 1$), and which could be combined with different degrees of usage restrictions, but which exhibit non-congruency with L1 Arabic” (adapted from Sonbul, 2012).

2.8 Summary

Research presented in this chapter has revealed that: (1) EFL learners, including those in the Saudi context, perceive vocabulary learning as an important and challenging aspect in learning English; (2) in teaching/ learning contexts some aspects of vocabulary knowledge has received greater attention (i.e. form and meaning of individual words) on different levels of reception and production, while other essential aspects of knowledge of word use (i.e. collocation) are almost neglected; (3) collocational knowledge as part of the umbrella term “formulaic sequences” is crucial for language acquisition, processing and use; (4) EFL learners, even at a very advanced level, produce fewer collocations than native speakers, and make more errors in their production; (5) EFL learners’ knowledge of L2 collocations is obviously and strongly influenced by their L1 and the collocations’ non-congruency with their mother tongue; (6) Arab EFL learners’ collocational knowledge does not appear to be any better or stronger than their European counterparts. The question is then, how can EFL learners be helped to achieve a better level of

collocational competence, especially those collocations which are non-congruent with their L1 and thus more challenging and difficult to produce? This is what my research aims to show.

Chapter 3: Methodological rationale: corpus- assisted contrastive analysis and translation

As presented and discussed in the previous chapter, collocations are only one aspect of word knowledge, but they are also part of the broader notion of formulaic sequences in a language. Research evidence suggests that EFL learners' collocational knowledge lags far behind native speakers' knowledge (Laufer & Waldman, 2011; Nesselhauf, 2003, 2005; Pawley & Syder, 1983). Additionally, EFL learners' collocational knowledge seems to be greatly affected by their L1 and by the non-congruency of collocations with their mother tongue (Nesselhauf, 2003, 2005; Laufer & Waldman, 2011).

Having established a rationale for addressing the notion of developing collocational knowledge in the EFL/ESL context (chapter 2), this chapter reviews the literature on collocations within second-language acquisition and instruction, and presents a methodological rationale for the purpose of this research. In other words, if the previous chapter attempted to answer the question of 'why focus on teaching and learning of non-congruent collocations', this chapter attempts to answer the question of 'why propose a corpus-assisted contrastive analysis and translation approach'.

Hence, the first part of this chapter will be allocated to the presentation and discussion of instructed approaches to second-language acquisition in general (sections 3.1 and 3.2) and in relevance to vocabulary acquisition (sections 3.2.1, 3.2.2). It will also comprise a critical review of the empirical research that has been conducted in the EFL classroom context to evaluate different conditions under which learners might develop collocational knowledge, thus showing the effectiveness of form-focused instruction (FFI) (section 3.2.3). Part one will be concluded with section 3.3 which pinpoints the gap in the empirical research. Part two (section 3.4) will be allocated to the justification of the proposed approach to teaching and learning collocation, i.e. the corpus-assisted contrastive analysis and translation approach. Section 3.4.1 will be allocated to the justification of the use of a data-driven approach to learning/ teaching collocations. Section 3.4.2 will present the justification for the use of the contrastive analysis and translation approach. Finally, the theoretical framework underpinning the proposed approach in this research will be detailed in section 3.5.

3.1 Second/ foreign language acquisition: to instruct or not to instruct

Most SLA researchers make a basic distinction between uninstructed (unguided, informal, naturalistic) second-language acquisition (SLA) on the one hand, and instructed second-language acquisition on the other. In uninstructed acquisition, the second language is learned through spontaneous communication in authentic natural situations, whereas instructed acquisition takes place under pedagogical guidance (Housen & Pierrard, 2005). Natural acquisition contexts are not only viewed as those contexts in which learners are exposed to the language at work or in social interaction, but they are also those classroom contexts in which the other learners are native speakers of the target language and where the instruction is directed toward native speakers rather than learners of the language (Lightbown & Spada, 2013). On the other hand, in instructional settings, the language is taught to a group of second language learners or foreign language learners, and the focus of the teachers is on the target language itself (Lightbown & Spada, 2013). The role and impact of instruction in SLA have always been controversial. Many researchers (e.g. Allwright, 1976; Corder, 1967;) have argued strongly against interfering with language learning, claiming that the best way to learn a language is by experiencing it as a medium of communication rather than treating it as an object of study. They perceived SLA merely as a result of learners' contact and interaction with the L2 environment in everyday life.

Concerning the notion of formulaic language acquisition and development, researchers, language learners, and teachers have typically perceived that the best way or perhaps the only way to develop a command of L2 formulaic sequences is through the immersion in an SL native environment, thereby maximizing the chances for repeated exposure to these combinations (Groom, 2009). The research in this area of interest shows contradictory findings. For example, Nesselhauf (2005) asserts that neither increased exposure to English in English-speaking countries, nor length of stay significantly impacts improvement in the number of collocations produced by the learners. A more recent study by Groom (2009) looked at the development of collocational knowledge through immersion of the learners in the native environment. Groom concluded his corpus-based investigation by suggesting that the claimed negative relationship between time spent in an L2 environment and the number of collocations produced by L2 learners depends on the way collocation as a concept is defined and operationalised, and on the way in which the results of a particular method of analysis are interpreted. Accordingly, the number of collocations (defined as lexical bundles of 2-3-4 and 5 words with a relatively high level

of occurrence in the USE 0 and USE 12+ corpora) produced by Groom's subjects does seem to decrease with time spent abroad. However, he argues that this decrease may entail a positive underlying trend as learners acquire and/ or introduce more variations of phraseological sequences than they already knew, hence rendering them invisible to the lexical bundle search procedure. As regards the claim that collocational accuracy is only slightly improved by a lengthy period of immersion in an L2 environment, Groom (2009) also found substantial positive correlation between collocational accuracy and L2 immersion. Nevertheless, researchers (e.g. Gass, 1989; Gass & Selinker, 2008) argue that SLA always entails the same basic processes regardless of context. As Gass (1989, p. 498) states:

“It is difficult to imagine a situation in which the fundamental processes involved in learning a non- primary language would depend on the context in which the language is learned... All learners have the capability of taking information from the input and organizing it within the framework of their current linguistic system and modifying and restructuring that system.”

Gass and Selinker (2008) affirm that the previous claim does not mean that differences in the quality and quantity of input in the different contexts do not exist, for clearly they do. For learners of another language in their own native environment, there is limited input, and a large part of the input comes from peers with typically restricted knowledge of the foreign language. Housen and Pierrard (2005) assert that SLA (and formulaic language in this case) should be regarded as a process in which the influence of instruction is an important social phenomenon. This research focusses specifically on the EFL context, and thus only classroom-based instructed SLA will be discussed in relation to vocabulary and collocation learning in this thesis.

The next section will introduce the notion of instructed second language acquisition and its categorizations.

3.2 Instructed second language acquisition

Instructed second language acquisition (ISLA) is a sub-category of SLA which involves all aspects of learning any language other than one's first language (L1) (Loewen, 2010). It is defined as “any systematic attempt to enable or facilitate language learning by manipulating the mechanisms of learning and/ or the conditions under which these occur” (Housen & Pierrard, 2005, p. 3). As Housen and Pierard (2005) note, this broad definition allows for a wide range of instructional approaches, techniques, strategies, methods, activities and practices which mainly occur in language classrooms.

ISLA can be primarily divided into meaning-focussed instruction (MFI) and form-focused instruction (FFI), depending on the emphasis that is placed on either linguistic form or meaning (Loewen, 2010). Generally, FFI involves instruction where the learners' attention is drawn to linguistic forms (Ellis, 2001a). FFI is also referred to by other researchers as “negotiation of form” (Lyster & Ranta, 1997) or “analytic strategy” (Stern, 1990). On the other hand, MFI – also referred to as “experiential strategy” (Stern, 1990) – involves the type of instruction that requires learners to pay attention only to the content of what they want to communicate (Ellis, 2001a). This distinction between MFI and FFI has been criticised by Widdowson (1998), who argues that form-focused instruction has always required learners to address meaning as well as form, whereas meaning-focused instruction still requires learners to process forms in order to encode and decode messages. Though this point has been well received by researchers, Ellis (2001a, 2001b) has argued that it is not new. Ellis (2001a, 2001b) argues that the crucial difference between meaning-focused instruction and form-focused instruction lies in how language is viewed (as a tool or as an object) and the role that the learners play (as a user or as a student).

Although advocates of MFI use the terminology in various ways, some have gone so far as to claim that learners of an L2, especially teenagers and adults, can successfully acquire the target language, whether implicitly (without awareness) or incidentally (without intention), from exposure to comprehensible target language input. MFI proponents generally believe that, like the acquisition of L1 by young children, L2 learners are capable of; (1) analysing L2 input subconsciously and inducing rules and/ or establishing new neural networks underlying these rules, and/ or (2) accessing inherent knowledge of linguistic universals and the way languages can differ (Long & Robinson, 1998).

The later view might seem theoretically sound and coherent to the advocates of MFI. However, it suffers from at least four issues. According to Long and Robinson (1998), an increasing amount of empirical evidence indicates that, unlike young children, older learners do not have the same capacity to attain native-like command of a new language merely from exposure to the language and its use. Moreover, studies show that adult learners with sustained natural exposure may become fluent speakers, but not native-like (Ellis, 2001a). Another criticism of the purely analytic approaches which underlie MFI is the unreliability of some L1-L2 grammatical contrasts from input alone. Additional studies (reviewed in Ellis, 2001a; Long & Robinson, 1998; Norris & Ortega, 2000)

suggest that although learning an L2 through exposure may be possible, it is insufficient. This has led second-language acquisition researchers to realise that L2 learners are unlikely to attain high levels of linguistic competence from purely meaning-centred instruction and that learners should attend to form as well (Ellis et al., 2002). Evidently, form-focussed instruction which relies on inductive learning mechanisms is not only beneficial, but essential for adult learners to acquire an L2 (Ellis et al., 2002).

Long (1991) was among the first to distinguish types of ISLA. He distinguished two kinds of form-focused instruction (FFI), namely focus-on-forms and focus-on-form. According to Long, focus-on-forms (FonFs) is manifested in the traditional approach to grammar teaching, based on synthetic syllabi. The fundamental assumption in FonFs is that language learning is a process of accumulating and addressing discrete linguistic entities or forms either deductively or inductively. In such an approach, learners are required to treat language primarily as an “object” to be studied and to function as learners of the language rather than as users (Ellis, 2001a; Ellis et al., 2002). On the other hand, focus-on-form (FonF) as defined by Long (1991) has two main characteristics: (1) attention to form takes place in lessons whose overriding focus is meaning or communication, and (2) attention to form arises incidentally in response to communicative need (Ellis, 2001a). Long (1991, p. 45) argues in favour of both types of FFI stating that it offers three advantages over either naturalistic SLA or classroom instruction with no focus on form: (1) it accelerates the rate of learning, (2) it influences acquisition processes in ways conceivably beneficial to long-term accuracy, and most crucially, (3) it seems to raise the ultimate level of L2 attainment.

While Long (1991) and Long and Robinson (1998) lay the foundation for FFI, Ellis (2001a) presents a comprehensive framework of what he considers to be FFI. Ellis argues that FFI involves three rather than two broad types; (1) focus-on-forms, (2) planned focus-on-form and (3) incidental focus-on-form. Each of these broad types has different sub-categories and instructional options to achieve FFI according to the primary focus of each type (see the following table for a summary).

Table 3. 1: Summary of R. Ellis' (2001a) FFI types

Type of FFI	Primary Focus	Distribution	Options for achieving FFI
Focus-on-forms	Form	Intensive	<ul style="list-style-type: none"> • Explicit vs. implicit focus-on-forms • Structured input vs. production practice • Functional language practice
Planned focus-on-form	Meaning	Intensive	Options relate to: <ul style="list-style-type: none"> • Input (enriched input, input flood, input enhancement) • Production (focused communicative tasks)
Incidental focus-on-form	Meaning	Extensive	<ul style="list-style-type: none"> • Pre-emptive focus-on-form • Reactive focus-on-form

The notion of FFI in L2 acquisition has primarily been the focus of research in grammar. Thus, the following section will be dedicated to discussing the notion of instruction in relation to vocabulary acquisition.

3.2.1 Instructed vocabulary acquisition

Quite similar to the main principles underpinning the naturalistic noninterventionist and the meaning-focused approaches to SLA, the ‘default hypothesis’ of vocabulary acquisition claims that “we acquire most words in our native language through exposure to language input, particularly written input, rather than by instruction” (Laufer, 2005a, p. 311). One justification for the default hypothesis is the abundance of words that people know that is difficult to account for by direct teaching of vocabulary (Laufer, 2003, 2005a, b). The hypothesis is also supported by research demonstrating that the largest L1 vocabulary development occurs when children reach literacy and they are expected to read approximately a million words of text a year (ibid).

The default hypothesis has found supporters among quite a few L2 researchers and practitioners despite the fact that the hypothesis was primarily developed based on first-language vocabulary acquisition. Krashen (1989) among other researchers has been predominantly active in promoting the prominence of reading for vocabulary acquisition, in naturalistic as well as instructed contexts. However, Laufer, a well-known supporter of instructed second language vocabulary acquisition, FFI in particular, has strongly argued against the default hypothesis (Laufer, 2003, 2005a, b). In her article “Instructed Second

Language Vocabulary Learning: the fault in the ‘default hypothesis’” (2005a), she convincingly argues that the basic assumptions underlying the default hypothesis (i.e. the noticing assumption, the guessing ability assumption, the guessing-retention link assumption, the extrapolation assumption and the repeated-exposures retention-link assumption) cannot be taken for granted in instructed language context. For example, Laufer (2005a, b) asserts that on seeing a new word, the learner does not necessarily ‘notice’ it, i.e. does not recognise it as unfamiliar word. This is either because of the tendency to overestimate one’s understanding of words in text context, or because of confusion with another word. In addition, noticing words as new does not guarantee success in inferring their meaning, and successful guessing does not automatically result in successful retention of meaning. Laufer (2005a) also states that “repeated exposures to the same word are indeed related to its retention, but to ensure repetitions of the same vocabulary, a ‘flood of reading’ is required, which is hard to implement in classroom instruction” (p. 324). Finally, experimental research reported very small vocabulary gains from long and short texts, and these gains cannot be extrapolated to larger quantities of reading (see Laufer, 2005a, for a review of the experimental research). It is worth noting, however, that Laufer (2001, 2003, 2005a, b) does not reject the importance of reading for vocabulary learning or argues against the educational value of reading activities. She mainly argues against the aforementioned assumption as being the primary resource of L2 vocabulary knowledge. Thus, Laufer (2005a) proposed an alternative hypothesis for vocabulary learning in instructed learning context i.e. planned lexical instruction (PLI) which is in line with form-focused instruction in general. The proposed hypothesis states:

“in view of the special conditions which obtain in instructed language learning context, the main source of L2 vocabulary knowledge is likely to be word focused classroom instruction” (Laufer, 2005a, p. 321).

Although most discussions of ‘focus-on-form’ have been done in relation to grammar, Ellis (2001a) rightly emphasises that ‘form’ involves more than grammar. Laufer therefore adopted Ellis’ definition of FFI as “attention to lexical forms and the meanings they realize, where words are treated as objects to be learned” (Ellis 2001a, p. 13). This can be done in a meaning-based task, or in a decontextualised vocabulary activity (Laufer, 2005a, b). Accordingly, attending to lexical items within a communicative task environment is considered as FonF, since these lexical items are necessary for the completion of a communicative or an authentic language task (Laufer, 2005a, b, 2006). On the other hand, teaching and practising discrete lexical items in non-communicative,

non-authentic language tasks is considered as FonFs. FonFs lexical instruction can be incidental, or intentional, whereas FonF is by definition incidental (Laufer, 2010).

Laufer (2001, 2003, 2005a, b, 2006, 2010) strongly argues for the superiority of planned lexical instruction over comprehension-based meaning-focused instruction for vocabulary learning. Laufer (2005a, p. 323) states: “PLI compensates for the relative paucity of input and a limited reoccurrence of words in instructed learning context. It also ensures noticing, provides correct lexical information, and creates opportunities for forming and expanding knowledge through a variety of word focused activities.”

Laufer’s (2005a, b) arguments in favour of the PLI or FFI for vocabulary acquisition are largely true, especially in relation to the EFL instructed language learning context. This is apparently because EFL learners may not know how much vocabulary they do not know while reading so they may not notice the words they do not know. Additionally, the flood of reading required for repeated exposure is probably an unrealistic expectation in the EFL context. Hence, more focused attention to particular vocabulary items would result in better learning. To support the earlier arguments presented regarding the effectiveness of FFI the following section will review the empirical research on instructed vocabulary acquisition.

3.2.2 Empirical research on instructed vocabulary acquisition

Findings of comparative research suggest that the proportions of the acquired words are usually greater in FFI conditions than non-FFI ones. This is evident in many studies (e.g. Ellis & He 1999; Hill & Laufer, 2003; Knight, 1994; Laufer, 2000, 2003; Luppescu & Day, 1993; Paribakht & Wesche, 1997; Sonbul & Schmitt, 2010).¹⁵ Other experimental research investigating whether some types of FFI are more effective than other FFI types (e.g. File & Adams, 2010; Laufer, 2006) has concluded that FonFs conditions yielded superior results as opposed to FonF conditions.

Laufer’s (2003) article comprised three experiments which aimed at checking how much vocabulary was gained from reading with marginal glosses compared to different FFI conditions. In the first experiment, two groups of 60 EFL university students were compared on incidental acquisition of ten unfamiliar, low frequency, target lexical items. One group encountered the words in a text in which the words were glossed in the margin. The learners in this group were asked to answer ten comprehension questions. The second

¹⁵ Due to the lack of space and the abundance of empirical research, only influential studies and the most recent work is reviewed in this section.

group was presented with a list of the ten target words with explanation and translation of meaning. The learners in this group were asked to write an original sentence with each word. An immediate and a delayed post-test were given to both groups in which the learners were asked to provide the words' meanings in L1 or L2. The 'sentence writing' group significantly outperformed the 'reading group' on both tests.

The second experiment's aim was to compare the number of words recalled after a reading activity on the one hand with the number of words recalled after using these words in a composition on the other. The subjects were 82 advanced university EFL learners of English in two parallel classes. The target words were the same ten lexical items used in experiment 1. Each class of learners carried out a different task. The task carried out by one of the classes consisted of reading comprehension with marginal glosses (same as experiment 1). The other class carried out a task that involved writing a composition incorporating the ten target words. The target lexical items were presented on a sheet of paper with explanation in English and translation of meaning for each word. On immediate and delayed post-tests (same as in experiment 1) the 'composition group' retained significantly more word meanings than the 'reading group'.

The purpose of Laufer's (2003) third experiment was to compare three tasks with regard to the number of words recalled after each one. The participants were 90 high-school students in three parallel classes. The target items were ten words with relatively low frequency to ensure that the learners were not familiar with them. One group read a text and looked up the words in a dictionary, the second group wrote original sentences with the target words, and the third group filled in the target words in given sentences. The participants in group 2 and 3 received a list of the target words with explanations of their meaning in order to perform the tasks. Both on the immediate and the delayed post-tests (same as in experiment 1), the 'reading group' attained significantly lower scores than the other two groups.

A more recent experimental study by Sonbul and Schmitt (2010) evaluated the effectiveness of the direct teaching of new vocabulary items in reading passages. The study compared vocabulary learning under a reading only condition to learning plus direct communication of word meanings. Sonbul and Schmitt (2010) assessed the learners on three levels of vocabulary knowledge (form recall, meaning recall, and meaning recognition) using three tests (completion, L1 translation, and multiple choice). Incidental learning which was aided by explicit instruction was found to be more effective than

incidental learning alone for all three levels of knowledge. The results also showed that direct instruction (i.e. FFI) is especially effective in facilitating the deepest level of knowledge, i.e. form recall.

Believing that some types of FFI are not only more effective than input and MFI, but also more effective than other types of FFI, Laufer (2006) compared the effectiveness of FonF vs. FonFs tasks for learning new L2 words under two conditions, namely incidental and intentional.¹⁶ Six intact classes of high school learners ($N= 158$) were assigned to the experiment, three classes of a total of 79 participants for each of the two conditions. Each class contained native speakers of Hebrew and Arabic. The researcher administered a pilot test according to which she chose twelve target words which were unlikely to be familiar to the subjects. In the incidental learning phase (FonF treatment), participants were exposed to the target words during a reading task. After reading the text, the learners answered comprehension questions for which they needed to understand the target vocabulary. Learners were advised to use bilingual dictionaries whenever they needed to. The incidental FonFs group did not read the text, but received a list of the twelve target words with their explanations in English and translations. Then the students worked on two word-focused exercises. Finally, an immediate post-test was conducted which tested their passive/ receptive knowledge of the words.¹⁷ In the post-test, the learners were to provide the meaning for the target words in English or in their L1. In this phase, the analysis of the test results showed that the FonFs group outperformed FonF group (47% retaining of word meanings as opposed to 72%).

In the second phase, under the intentional condition, all participants (both FonF and FonFs) received a list of the twelve target words with definitions of meaning, examples, and translations. Participants were asked to spend 15 minutes on memorising the words and their meanings for an upcoming test. After they had completed memorisation, two tests were carried out; the same post-test of passive knowledge used in phase one and an

¹⁶ It is of paramount importance to note that the notion of incidental vocabulary learning has different indications in the literature. According to Hulstijn and Laufer (2001) and to Hulstijn (2001), incidental learning does not mean that a learner does not attend to the words during the task. He/ she may attend to the words under explicit teaching, but he/ she does not deliberately try to commit them to memory. Incidental learning according to Schmitt (2010), however, is learning which accrues under implicit instruction as a by-product of language usage, without the intention to learning new lexical items. Intentional vocabulary learning, on the other hand, refers to an activity aimed at committing lexical items to memory under explicit teaching (Hulstijn, 2001; Hulstijn and Laufer, 2001).

¹⁷ Students were asked to provide explanation in English or translation in their L1.

active word knowledge test.¹⁸ Results of this phase of the study showed a drastic disappearance of differences between the two conditions. There were no significant statistical differences between the two groups in the immediate post-test nor in the delayed post-test.

It is worth mentioning that the results of the second phase of the study were expected for two main reasons: (1) by definition, intentional learning is a FonFs activity since the target words were decontextualised and became the object of study rather than tools for communication, (2) the subsequent conscious memorising effort of the words that learners invested for an upcoming test can increase the number of learnt words. Thus, it could be concluded that of the two FFI types, the FonFs is more effective than FonF.

In relatively similar study to Laufer's, File and Adams (2010) compared isolated and integrated¹⁹ form-focused instruction for vocabulary development in an English as a second language (ESL) reading lesson. The participants were two classes of adult students of intermediate proficiency from a university preparation programme. The researchers followed a pre-test, post-test and delayed post-test design to examine the influence of FFI on learning and retention of new vocabulary. Eighteen target words were systematically selected from the 5,000-word level to ensure that they are most likely to be unknown to the participants. Of the 18 words only twelve words were selected for the instruction whereas the remaining six were integrated in a text to examine incidental learning through exposure. Two reading treatments (isolated and integrated vocabulary instruction) were conducted in each class. In the isolated treatment, the researcher gave an oral definition of all twelve vocabulary items, and two synonyms and an example of each word was shown on an overhead transparency before the participants read the text. The twelve target words were bolded in the text, however, no further attention was given to them in the reading process. Conversely, in the integrated instruction, the researcher began the oral reading of the text immediately. After reading a sentence that contained one of the twelve target words, the researcher would then return to the target word, draw participants' attention to the form, providing the correct stress, an oral definition and two synonyms. An immediate post-test was conducted after the treatments. Two weeks later the delayed post-test was administered. Paribakht and Wesche's (1997) vocabulary-knowledge scale was employed to measure learning and retention gains for words for both types of form-

¹⁸ L1 translations of the target words were given and the learners were asked to provide the target L2 words.

¹⁹ By definition, integrated FFI corresponds to the notion of FonF instruction, whereas isolated FFI corresponds to the notion of FonFs.

focused instruction as well as for words acquired incidentally. Statistical analysis of the data showed that both types of instruction led to more learning and retention of vocabulary knowledge in both tests than incidental exposure alone. The researchers stress that despite the similar retention rates for isolated and integrated instruction, there was a trend for isolated instruction to lead to higher rates of learning than the integrated treatment. It should be noted, however, that the limited sample size ($N=20$) and the small number of treatments (only two) were probable factors affecting the learning and retention trends, and that a larger sample size and more treatments might have led to stronger and more significant trends.

It is of great importance to point out that vocabulary practice and learning in a computer-assisted setting can be considered a particular case of FonFs (Laufer, 2006). Most of the research conducted to investigate the effectiveness of FFI for vocabulary learning was teacher-centred and did not employ learner-centred or technology-assisted methodologies, with the exception of Hill and Laufer's (2003) study which involved electronic dictionary-checking activities using a computer programme. A learner-centred study by Horst, Cobb, and Nicolae (2005) investigated vocabulary learning through the use of online dictionaries, word banks, cloze exercises, concordances, hypertexts, and self-quizzes. They found that high-school learners, as well as both weak and strong university students learned many of the practised words both receptively and productively. The results suggest that most learners could benefit from FonFs. Most interestingly, the researchers argued for the effectiveness of vocabulary acquisition tools that are based on a corpus. They suggest that such tools expand and vary opportunities for lexical rehearsal, and engage the learners at a deep level of processing. Horst et al. (2005) certainly point out that "not every instance of processing or rehearsal must pass through a teacher" (p. 106).

The studies presented in this section argue in favour of FFI as opposed to MFI and in favour of FonFs conditions as opposed to FonF (see sections 3.2. and 3.2.1 above for rationale). This signals FonFs as a significant and effective instruction type to be employed in the current research. However, the above reported studies used individual words as instructed target vocabulary. A large body of research suggests that the mental lexicon mostly consists of formulaic language and is built from multi-word units such as idioms, phrasal verbs or collocations (cf. chapter 2 above). Therefore, the following section will address the empirical work on the effects of different types of instruction on

second language learners' knowledge of one type of formulaic sequences i.e. collocations which form the focus of this research.

3.2.3 Empirical research on instructed acquisition of collocation

As shown in a previous section, empirical research on vocabulary acquisition has suggested that the majority of words are learned through direct form-focused instruction with comparatively few gains being made through meaning-focused and incidental instruction in an EFL context. Moreover, the research shows that incidentally acquiring meaning for even fairly salient single-word items (through exposure) is a relatively slow process in which acquisition is dependent on the amount of input (Horst, Cobb, & Meara, 1998; Waring & Takaki, 2003). Consequently, Webb and Kagimoto (2009) argued that in this case the learning of collocation incidentally could be a rare occurrence due to the limited number of opportunities to encounter the same collocation twice. This necessitates the introduction of collocation explicitly into the L2 classroom. This is suggested in various experimental research targeting incidental acquisition of collocations.

An early investigation in the context of incidental acquisition of collocations by Marton (1977) was a small-scale study in the context of Polish EFL learners. Findings showed an insignificant increase in the learners' collocational knowledge as a result of two weeks of reading-based exposure to the target collocations. However, the findings of the study can be questioned due to a faulty design. Only the participants in the experiment group, but not those in the control group, took the post-test. Another problem with the design is the fact that the L1 text to be translated into L2 was different in the post-test from the text in the pre-test. A more recent and better controlled experiment by Webb, Newton and Chang (2013) concluded that incidental learning of collocations in the EFL classroom is possible. The experiment showed a strong correlation between the number of exposures to the target collocations (at least ten within a short period of time) and collocational acquisition and development. It should be acknowledged, however, that no delayed test was included in this study. Thus, it is not clear whether these immediate effects were durable or not. According to Schmitt (2010), a delayed post-test is a crucial indication for a stable and durable learning. Additionally, this study did not include a direct FFI teaching condition to allow a comparison with the incidental approach.

Despite widespread recognition of the difficulties learners have in producing collocations and their critical role as part of formulaic language in L2 development (see chapter 2,

section 2.5 and 2.6 for a detailed overview), few empirical studies have addressed the issue of how collocations can be most effectively learned and developed in an EFL context under different FFI conditions. In fact, most research on collocational knowledge in the EFL context has focussed on usage and processing rather than acquisition (as reviewed in chapter 2, section 2.6).

Sonbul (2012) was one of the first to examine the effect of different conditions (instructed and incidental) on improving both explicit and implicit knowledge of collocation. The target items were 18 highly frequent adjective/ noun collocations. The subjects were 30 female Arab speakers of English in an EFL classroom at undergraduate level. The study followed the standard design of classroom acquisition research (pre-test, treatment, post-test). The conditions included in the design were incidental (collocations embedded in a passage), instructed (collocations presented in a list and followed by a short exercise), and control (no exposure). A counter-balanced design was used in which each group of participants received the three teaching conditions but for a different set of collocations. Pre-testing and post-testing phases with the implicit (priming) and explicit (form recall and form recognition) measures had taken place two weeks before and again three weeks after the treatment. Data analysis showed that learners developed explicit collocational knowledge only under the instructed/ direct teaching condition but did not develop implicit knowledge under either condition. The researcher concluded that direct instruction might be the most efficient teaching method for EFL learners to develop explicit knowledge of collocations.

Two classroom studies (Webb & Kagimoto, 2009, 2010) were conducted in an EFL setting to evaluate the effectiveness of various FFI and MFI instruction methods on intentional learning of verb/ noun collocations. Webb and Kagimoto's (2009) study investigated the effects of receptive and productive vocabulary tasks on learning 24 highly frequent collocations. 145 Japanese EFL students of intermediate proficiency were asked to attend to target words in three glossed sentences (the receptive condition) and in a cloze task (the productive condition). Before the treatments, the learners in both groups as well as a control group had taken a pre-test of receptive knowledge only. Three weeks later, the treatment phase took place in a 90-minute session for both groups. In order to determine the effects of the treatments, four tests were then employed to measure receptive and productive knowledge of collocation and meaning: productive knowledge of collocation (cloze), receptive knowledge of collocation (MC), productive knowledge

of meaning (L1-L2 translation) and receptive knowledge of meaning (L2-L1 translation). The results showed that both receptive and productive FFI tasks led to substantial gains in meaning and collocational knowledge, and there was no statistical difference between the two tasks on any of the tests. However, when participants were rearranged into low-level and high-level groups, the receptive task was shown to be more effective for lower-level learners, and the productive task was more influential for higher-level learners. That said, the study has two important limitations: (1) a pre-test measuring productive knowledge was not administered; (2) a delayed post-test was not given in this study.

In another recent study by Webb and Kagimoto (2010), the researchers investigated the effects of the number of collocates per node word, the position of the node word, and synonymy on learning five sets of twelve ($N=60$) adjective/noun collocations. The target items were collocations with a low degree of overlap in translation equivalency/congruency, though the researchers did not specify how they distinguished between high and low degrees of congruency. The participants of the study were 41 Japanese students in two colleges. Like the previous study, this study was conducted in a 90-minute session. The participants were pre-tested for their productive knowledge of collocations (decontextualised L1/ L2 translation). In the treatment, the participants encountered the target collocations in glossed sentences. Three minutes time was allocated for the learning of each set of collocations. An immediate post-test similar to the pre-test was conducted after the treatment. In response to the research questions, the study showed that as the number of collocates per node word increased, more collocations were learned. In addition, the position of the node word had no effect on collocation learning, and synonymy had a negative effect on learning. It is worth noting that, similar to the previous study by Webb and Kagimoto (2009), this study also lacks a delayed-post-test phase. Besides, the researchers did not control for the congruency of the collocations, and they admit that “it is possible that some items may have been easier to learn than others... how the degree of congruency between the collocations in the different sets affected learning is not clear” (Webb & Kagimoto, 2010, p. 273).

While the previous two studies focused on intentional learning in the form of FonFs instruction, Laufer and Girsai's (2008a) study investigated the effect of three instructional conditions on the 'incidental' acquisition of single words and non-congruent verb/ noun collocations: MFI, FFI and the contrastive analysis and translation condition (CAT). Participants were assigned to three groups, each of which represented an instructional condition. In the MFI condition, the participants were assigned to content-based activities

while not attending to the target items. The FFI group carried out text-oriented vocabulary activities focusing on the target items. The CAT group performed text-based translation tasks from L1 into L2 and vice versa. The participants in the CAT group received a teacher-centred contrastive analysis of the target items and their L1 translations during the correction stage upon finishing the tasks. An immediate post-test of active recall (L1/ L2 translation) and passive recall (L2/ L1 translation) was administered one day after the treatments. One week later, a similar delayed test was given to all groups. Results showed that the CAT group significantly outperformed the MFI and FFI groups on all the tests, with the MFI being the least effective. However, a delayed post-test of three weeks would have been better and more indicative of learning which is stable and durable.²⁰

It is to be mentioned that the previous study by Laufer and Girsai was a follow up of a similar preliminary study by the researchers. In the preliminary study (Laufer & Girsai, 2008b) the same design and procedures were followed except that the experiment only included the MFI and the CAT conditions but no non-contrastive FFI condition. The researchers could only conclude that contrastive FFI was superior to message/ meaning-based instruction, but not that it was any different from FFI in general.

Similarly, Szudarski's (2012) six-week study compared the effect of meaning-focused instruction combined with focus-on-forms instruction on the acquisition of collocations by 43 L1 Polish learners, as opposed to meaning-focused instruction only. The target collocations were 50 verb/ noun collocations with frequent delexical English verbs which are non-congruent with the learners' L1. In the first week, a pre-test of receptive (MC) and productive knowledge collocations (L2/ L1 translation and cloze task) was administered. A week later, the treatment phase started and lasted for three more weeks with 45 minutes per week. The participants were divided into two experimental groups, an MFI plus FonFs group and an MFI only group. The first group read stories that contained target collocations and completed explicit activities focusing on the target collocations, while the other group read the same stories and answered comprehension questions with no explicit reference to the collocations. Two weeks after the last treatment, the participants undertook a post-test which was identical to the pre-test. By comparing the results of the experimental groups to the control group, both treatments appear to have led to improvement in collocational knowledge. Findings of pre-test/ post-

²⁰ Though there is no standard period of delay, Schmitt (2010), drawing on memory and mental lexicon research, suggests a three-week ideal delay and a minimum of one week.

test results revealed that the MFI followed by a FonFs condition had a more significant effect on enhancing learners' collocational knowledge at both the productive and receptive level than the MFI only condition. Although this study has a sound design, the findings were relatively predictable. Moreover, an immediate post-test would have allowed an examination of the immediate impact of the treatments in comparison to the delayed results, though this has no serious effect on the overall design.

Finally, the use of computer-assisted language learning facilities and activities, particularly web-based concordances, was studied by Sun and Wang (2003) and Chan and Liou (2005). Sun and Wang (2003) used a concordance program to examine the virtual effectiveness of inductive and deductive²¹ approaches to learning grammatical collocations, as well as the relationship between the difficulty of the collocation pattern and the learners' performance. The researchers randomly divided a group of 81 Taiwanese senior high-school students into two groups (inductive $N= 41$ / deductive $N= 40$). After a 20 minutes pre-test of error correction, the learners were asked to complete a one-hour instruction session of online exercises for four target collocations that used either an inductive or a deductive teaching approach. The four target collocations were divided into two groups of what the researchers called easy patterns and difficult patterns. Immediate post-test results showed that the inductive group improved significantly more than the deductive group in learning the target collocations. The pattern of collocation difficulty was also found to influence the learners' performance with easy collocations being more suitable for an inductive teaching approach. However, as noted by Chan and Liou (2005) and by Webb and Kagimoto (2009), the design of the study had several weaknesses, including the small sample size of collocations and the random and ambiguous nature with which the collocations were allocated into two levels of difficulty. Moreover, the durability of the learning was not assessed in a delayed post-test. Hence, limitations of the study design cast doubt on the generalizability of the results.

Chan and Liou (2005) also investigated the effects of web-based concordancing on learning verb-noun collocations by 32 college EFL students. Five web-based units were designed in the format of semantic grid analysis, bilingual concordance, textual explanation and interactive exercises with an audible online information reader. Three of the units were taught with the use of a bilingual Chinese-English concordance, and two

²¹ Deductive teaching involves presenting rules first, then examples, whereas inductive teaching involves the presentation of examples, then inducing patterns and rules (Sun & Wang, 2003).

units were taught with no concordance. In agreement with Sun and Wang's (2003) results, they also found that explicit online instruction was effective in developing EFL learner collocational knowledge, with significantly higher results for units in which the concordance had been used. Results also showed significant differences in learning between four verb/ noun collocation types (i.e. synonymous verbs, hypernymous and troponymous verbs, de-lexicalised verbs, and collocations that are non-congruent between Chinese and English) with concordances. Concordances were deemed to be most suitable for use in the teaching of non-congruent verb collocations and de-lexicalised verbs. As noted by Webb and Kagimoto (2009), the study design has two main limitations: (1) although a wide variety of interactive activities such as multiple-choice and gap-filling tasks were employed within the online practice units, the research questions focused on the learning impacts of the concordance alone, and there was no discussion of how the different types of tasks might have affected different types of learning gains; (2) overall gains in collocational knowledge were assessed by the use of a single productive (cloze) test after two and a half months of the immediate post-test, making any further analysis of task type largely impossible.

Tsai (2011) conducted a longer and multidimensional study investigating the effect of corpus consultation on learning collocation, for which the researcher coined the abbreviation DALC i.e. data-driven approach to learning collocations. The researcher explored three key dimensions: the learning product, the learning processes and the learner perceptions of DALC. The participants were 186 undergraduate EFL learners in Taiwan. Collocational knowledge was examined at three levels: receptive knowledge, controlled productive knowledge and free productive knowledge. To understand how collocation learning occurred with DALC, the researcher elicited the thinking processes in which learners engaged as they undertook the DALC. A questionnaire was administered to elicit learners' perceptions of DALC. The findings indicate that DALC had a positive impact on all measured levels of learners' collocational knowledge. The elicited thinking processes indicated that the quantitative and qualitative changes in the learners' collocational knowledge could be attributed to the intense cognitive processing they engaged in during the DALC task. According to the questionnaire, the participants' attitude toward DALC was mainly positive, but they were nonetheless concerned about the efficiency of such an endeavour.

3.3 Summary and research gap

To summarize, the studies reviewed in the previous section show that incidental learning of collocations (as in Schmitt's 2010 view) through exposure is possible. For this learning to occur, multiple encounters (at least ten) are needed to establish receptive/ passive and productive/ active collocational knowledge. However, the largest body of research suggests that direct form-focused instruction of collocations, FonFs in particular, leads to higher gains than indirect instruction or incidental exposure. However, the studies examining acquisition of collocations under different FFI conditions are limited in several ways. The majority of the studies targeted verb/ noun collocations (with the exception of Webb & Kagimoto, 2011 & Sonbul, 2012) and ignored other categories of lexical collocations. Despite the acknowledgment of the inherent difficulty of collocations for L2 learners, especially those that do not have equivalents in the learners' L1 (non-congruent collocations), only a few studies addressed the notion of non-congruency and examined the different conditions under which EFL learners might develop collocational knowledge (e.g. Chan & Liou, 2005; Laufer & Girsai, 2008a, b). In Chan and Liou's (2005) study the use of a data-driven approach with the help of online concordances proved to be significantly beneficial for learning de-lexicalised and non-congruent verb/ noun pairs, despite the limitations of the study (see the previous section). On the other hand, Laufer and Girsai (2008a, 2008b) consistently state that when it comes to the learning and teaching of collocations with no L1 equivalents, adopting a contrastive analysis and translation (CAT) method as part of an explicit form-focused instruction (FFI) is evidently more effective than meaning-focused instruction (MFI) and other explicit FFI with no contrastive analysis and translation. However, as far as the current researcher knows, no attempt has been made to investigate the effect of both form-focused instructional approaches (i.e. data-driven and CAT) combined on the acquisition of non-congruent lexical collocations. Hence, my research aims to fill this methodological gap through examining the effect of a corpus-assisted contrastive analysis and translation approach on the acquisition of non-congruent adjective/ noun collocations.

The next section is intended to provide a rationale for combining the two instructional approaches for the purpose of teaching non-congruent collocations in the current EFL context.

3.4 A corpus-assisted contrastive analysis and translation approach to learning collocations: the rationale

This section provides my rationale for the utilisation of a data-driven approach (DDL) in the form of concordance data from a bilingual corpus (English/ Arabic) and the use of translation tasks for the purpose of teaching non-congruent adjective/ noun collocations in a contrastive manner.

3.4.1 The rationale for data- driven learning (DDL)

“It is my belief that a new understanding of the nature and structure of language will shortly be available as a result of the examination by computer of large collections of texts.”

(Sinclair, 1991, p. 489)

Since the mid-1980s, the corpus-based approach to linguistics and language education has gained eminence. This is because corpus analysis²² can be enlightening in nearly all branches of linguistics or language learning: grammatical, lexical, contrastive, translation and so on (McEnery & Xiao, 2011). Leech (1997) believes that there is a convergence between teaching and language corpora and attributes this to the ‘trickle down’ approach whereby techniques and resources used in research gradually become available for teaching; Johns (1991, 1994) sees it instead as a ‘trickle up’ process because those developing language teaching techniques naturally adopt the resources available for research.

Leech (1997) reviews three points of convergence between teaching and language corpora: their indirect use in teaching (language testing, materials development, and reference publishing), their direct use (teaching about, teaching to exploit, and exploiting to teach), and further teaching-oriented development (L1 and L2 developmental, LSP, and bilingual/ multilingual corpora).

The direct use of corpora in teaching, which is the focal area of the corpus-aided field, involves different ways in which corpora may be utilised. ‘Teaching about’ involves providing courses on corpus linguistics or in which they are a sizeable and significant component of the course. ‘Exploiting to teach’ involves making selective use of corpora in courses which would otherwise be taught by other methods (e.g. stylistic, lexical, semantic studies). Leech (1997) argues that the merit of the corpus, in such an approach,

²² Corpus analysis entails empirical analysis of the actual patterns of language use in natural texts (Biber, 2009).

is that it enables data to be delivered in a convenient way (such as print-outs) for the learner or investigator. McEnery and Xiao (2011) believe that unlike direct and indirect uses of corpora, ‘teaching to exploit’ relates to all language learners, who, they argue, benefit from a data-driven learning (DDL) approach, as it involves enabling learners to access and exploit a body of knowledge that is manipulated by the teacher for their benefit. Leech (1997) suggests that ‘exploiting to teach’ actually implies ‘teaching to exploit’ in the sense that supplying students with data in the form of print-outs is a way of helping them gain access to it and exploiting it to learn. He argues against considering the ‘exploiting to teach’ approach an ‘easy way out’ and inferior to ‘teaching to exploit’ learning sessions. He proposes that this so-called ‘easy way out’ in fact ensures that the maximum number of learners, who might in some cases be technophobic or lack necessary search skills, are able and willing to participate in this kind of learning experience, without being discouraged. This paper-based approach complies with what the notion of DDL entails and is not incompatible with it. It is supported and embraced by many researchers in the field (Boulton, 2009a, b, c, 2010; Breyer, 2006; Mukherjee, 2006). This standpoint is also held by the current researcher, who considers it an effective way of carrying out data-driven language teaching and learning in this EFL context (see the following sub-sections for justification).

Johns (1986, 1991, 1994) was among the first to realise the potential of corpora for language learners and to advocate the direct use of corpora in teaching. Johns (1991, 1994) proposed the DDL approach, which gives language learners access to the facts of linguistic ‘performance’ as an alternative to a rule-based approach, which endeavours to encapsulate the linguistic ‘competence’, thus departing from a deductive to an inductive approach of language learning. This shift in learning routines has extensive impact on the teachers, who become coordinators of research, as well as on the learners, who learn how to learn through exercises and activities that encourage the observation and interpretation of patterns of use (Bernardini, 2004). According to Johns (1994) the learners’ utilisation and interrogation of the corpus data not as a substitute teacher, but as a rather special type of informant is at the heart of this approach. These characterizations of DDL by Johns (1991, 1994) correspond to some extent to McEnery and Wilson’s (1997, p. 6) description of corpus-assisted language learning as:

- “Directed learning - learners are directed by the teacher but led by themselves through the corpus consultation.
- Mediated learning - the corpus is not a source of didactic learning: rather, it is a medium through which learning may be achieved. Students learn

through the process of interacting in some way with the corpus (e.g. via practical grammatical analysis), rather than from its explicit content.”

According to Johns (1991, 1994), the extraordinary increase in processing speed and storage capacity of affordable computers with flexible and powerful concordancing packages, as well as the increased availability of ready-made and easily-accessed corpora has given rise to the DDL approach. According to Flowerdew (1996), the concordance is the most important computing tool for the data-driven approach. Concordancing is “a means of accessing a corpus of text to show how any given word or phrase in the text is used in the immediate contexts in which it appears” (Flowerdew, 1996, p. 97). By grouping the uses of a given word or phrase on the computer screen or in print-outs, the concordance lines show the patterns in which the particular word or phrase is typically used, and this facilitates rapid scanning and comparison (Flowerdew, 1996; Johns, 1997). The most common format produced by concordances is keyword-in-context (KWIC) in which the keywords are positioned one below the other at the centre of the page, with a fixed number of characters of context to the left and to the right (Johns, 1994), as shown below in figure 3.1.

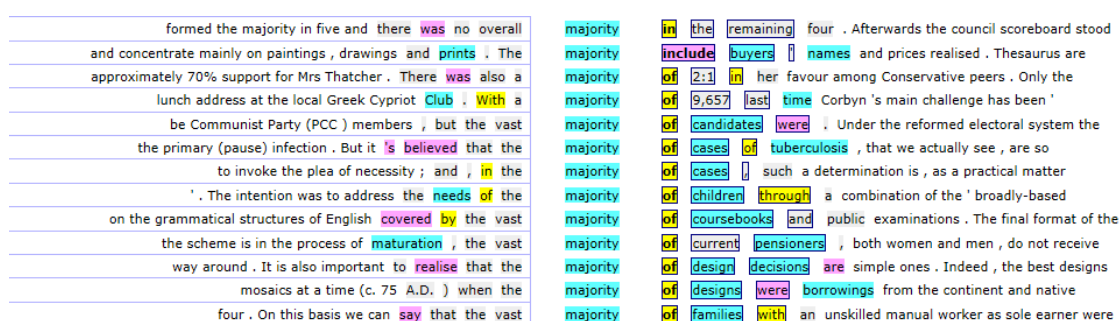


Figure 3. 1: Concordance of 'majority' from the British National Corpus (BNC) in KWIC format (taken from the Brigham Young University, BYU, website).

Having introduced the general merits of corpora and concordances in language teaching and learning, in the following sub-sections I will address their merits in relation to this research, providing the rationale underpinning the adopted data-driven approach for the learning of collocations. This incorporates notions such as authenticity, profusion and autonomous learning.

3.4.1.1 Authenticity

Corpus data have long been established as “the real language data” (Bernardini, 2004, p. 15), and many researchers consider this the main advantage a corpus has to offer. The easily-accessible and huge bodies of naturally occurring texts in corpora has helped

researchers, language teachers and learners to gain a better understanding of how language is actually used, as opposed to how language is perceived to be used i.e. intuition (Tsui, 2004). Corpus- based examples and extracts can also be viewed as an authentic substitute of what is described as “constructed and artificial” (Johns, 1994, p. 28) and “concocted” (Carter & McCarthy, 1995, p. 154) textbook examples. Braun (2005, p. 48) describes the data provided by corpora as being:

- “rich, providing more (and more diversified) information than dictionaries or reference grammars;
- illustrative, providing actual patterns of use instead of abstract explanations;
- up-to-date, revealing trends in language use and evidence for short-term historical change.”

This perception of the authenticity of corpus data is shared not only by linguists and researchers, but also by language learners and teachers. For example, Chambers’ (2005) small-scale study revealed a general agreement among the participants with regard to the perceived authenticity of corpus language. The participants described the corpus data as “authentic”, “real”, “up-to-date” and “relevant”, as opposed to the “unreal and sometimes stupid” invented examples in textbooks (Chamber, 2005, p. 120). Similarly, on investigating learners’ perceptions of corpus data, Tsai (2011) reported a positive attitude towards the language data in corpora, noting that it was perceived as more authentic than textbook language. Along the same lines, Farr (2008) reported that teachers perceived the access to real language data as the main asset of corpus-assisted language learning. Nonetheless, some researchers have cast doubts on this alleged intrinsic authenticity of corpus data.

Widdowson (1978) was among the first to argue against the inherent authenticity of corpus data which he generally referred to as “comprehension pieces” and “extracts” intended as demonstrations of language as use. Although these “extracts” are genuine instances of language use, they are not authentic discourse. He affirms that genuineness is a characteristic and absolute quality of the text itself, but notes that authenticity is a characteristic of the relationship between the text and the reader. Later (2000, 2003), he echoed his previous crucial distinction between text and discourse with regard to corpora and asserted that ‘text’ (represented in concordance lines) is a kind of static decontextualised semantic patchwork which exists as an object for analysis, and is considered as a product. Language learning, on the other hand, is concerned with

discourse, that is, contextualised text, which involves processual aspects.²³ Widdowson points out that “however the language is to be contextually abstracted..., they have to be recontextualized in the classroom so as to make them real for learners, and affective for learning” (2000, p. 8). In other words, real or genuine language texts are only useful if the learners and their teachers are able to authenticate them by creating a relationship to the texts (Braun, 2005) through a personal process of engagement with them (Van Lier, 2000).

To sum up, what is important here is the way in which a corpus and corpus data are used as pedagogically relevant to the learners’ needs. This authentication makes the difference between the use of a corpus from the perspective of what Widdowson (1980, 2000) calls “linguistics applied” (i.e. taking linguistic findings of corpus analysis more or less directly to the classroom), and from the perspective of “applied linguistics” (i.e. a pedagogical treatment of corpus descriptive findings to make them appropriate as prescriptions that activate the process of learning). (See Section 3.5.1 for how the corpus-assisted CAT approach was hypothesised to induce authentication for the purpose of collocation learning.)

3.4.1.2 Profusion

Many researchers affirm that the large repertoire of language use offered by corpora is undoubtedly what makes them valuable both as a resource for the creation of rich and interesting learning materials and for direct exploitation by learners. The focus of direct uses of large computerised corpora with multi-millions of words in language teaching and learning lies very strongly on concordance-based materials and activities (Braun, 2005). This abundance of data (generated through concordancing) highlights common and frequent patterns in language use and makes them more salient through the concentration and manipulation of instances of a language phenomenon (Aston, 1995; Barlow, 1996). According to Braun (2005), these concordance-based materials and activities have certainly brought a “healthy” focus on form back to language learning and teaching.

Collocations are recurring patterns which are pervasive in language, and which can also benefit from such profusion of corpus data. Corpus-data consultation can be an efficient

²³ It is worth noting that the distinction between “text” and “discourse”, and the perception of “text” as a static entity devoid of any contextual features has varied across the literature (see Tanskanen, 2006 for a detailed review).

method in providing learners in EFL classrooms with intensive multiple exposures to the same collocations that might in normal cases take considerable period of time for foreign language learners to encounter incidentally. This frequent occurrence of collocations is essential for noticing, thus converting input to intake, which then results in learning (Schmidt, 1990). Repeated exposure to the same collocation using corpus data could also be seen as the flood of input/ readings recommended by researchers (e.g. Laufer, 2005a, b) for vocabulary retention. As Thurstun and Candlin (1998, p. 270) put it, corpora offer “the opportunity to condense and intensify the process of learning through exposure to multiple examples of the same vocabulary item in context, and to promote awareness of collocational relationships.” Furthermore, the abundance of corpus data provides a wealth of resources for a principled “recycling” of the previously studied words/ collocations. In fact, many researchers (e.g. Nation, 1990; Schmitt, 2008; Schmitt and Schmitt, 1995) view the “recycling” of vocabulary as a process of paramount importance for consolidating vocabulary knowledge. Hence, it seems that the pedagogical application of corpora in the form of concordance lines would be a feasible and practical method in facilitating collocation learning and raising learners’ awareness to this phenomenon.

It is worth noting though that some researchers argue that this wealth of data and the sheer size of many available corpora make them difficult to manage and process by most teachers and learners. Meunier (2002) suggests that corpus results could be messy, ambiguous or misleading. Other researchers’ opposing views are less concerned about this messiness and ambiguity. For example, Cobb (1999) argues that learners would not be distracted by the flow of discourse as the words could be seen in multiple contexts rather than in isolation. Despite Braun’s (2005) concerns regarding what could be an overwhelming and time-consuming task for teachers, he views the messiness as part and parcel for using real-language materials, and he affirms that removing some unwelcome or unclear lines from a concordance before presenting it to learners is indeed a workable solution. Nonetheless, to overcome the so-called problems of size and the problems relating to the diversity and ambiguity of content, some researchers have suggested the use of small genre-specific corpora, the use of sub-corpora derived from large corpora and the use of language for specific purposes (LSP) corpora (Braun, 2005).

For the purpose of this research, the current researcher proposes the use of one type of specialized corpora i.e. parallel or translation corpora. A parallel corpus is a corpus which consists of original texts and their translations and which also lends itself to the kind of

DDL exploitations that involve LSP and learner corpora (Römer, 2009). According to Kjellmer (1992, cited in Römer, 2009), contrastive work (research based on parallel corpora) is valuable for the selection of “elements the learner is likely to mistreat because they are different [...] from those in his [or her] native language” (p. 375). Römer (2009) suggests that exploiting a parallel concordance and observing the translation equivalents of a lexical item in the L1 could be extremely helpful in coming to terms with the meaning(s) of this item. Thus, it is argued by the current researcher that a parallel corpus could be a useful DDL exploitation tool, with its wealth of real language data in the L1 and L2. Parallel corpora comprise two or more corpora in different languages, each of which contains translated texts from one language into the other, or texts that have been produced simultaneously in two or more languages (Hunston, 2002). A parallel corpus could also be a feasible solution for the alleged problem of content ambiguity and a way of highlighting collocational and phraseological differences between L1 and L2. Hunston (2002) suggests that translators as well as language learners can use parallel corpora to find potential equivalent expressions in each language and to investigate the differences between them.

3.4.1.3 Learner autonomy

The notion of learner autonomy is closely associated with DDL as an inductive approach to language teaching and learning. The novelty of this approach is presented in a well-known quote by Johns (1994, p. 297) as he states:

“What distinguishes the DDL approach is the attempt to cut out the middleman as much as possible and give direct access to the data so that the learner can take part in building his or her own profiles of meanings and uses. The assumption that underlies this approach is that effective language learning is itself a form of linguistic research, and that the concordance printout offers a unique resource for the stimulation of inductive learning strategies -- in particular, the strategies of perceiving similarities and differences and of hypothesis formation and testing.”

One of the most prominent aspects of the DDL approach is the shift in the role of the teacher and the learner. In a traditional EFL classroom, the teachers are considered to be the source of linguistic knowledge, while the learners are mostly perceived as the passive recipients of that knowledge. A DDL approach promotes more autonomy however, and the teacher becomes the director and coordinator of the corpus data exploitation and exploration conducted by learners (Bernardini, 2002, 2004; Johns, 1991, 1994). The notion of autonomy within what the current researcher proposes to call ‘mainstream DDL’ involves corpus consultation that allows learners to take greater control over their

learning. This means that learners can make choices about what they would like to learn by setting their own tasks, answering their own questions, and formulating hypotheses for themselves via self-accessed exploitation of corpora (O'Sullivan, 2007). One could argue that autonomy in that sense may be faced with mainly (but not exclusively) three sources of concerns: learners, teachers and DDL resources.

As previously mentioned, learners' attitude towards the use of corpus data as an authentic source for language learning is generally positive. However, researchers increasingly cast doubt on what they call "full hands-on DDL" in the "mainstream DD" approach. Kennedy and Miceli (2001), for example, argue that in interacting with online corpora, learners have difficulty thinking of and formulating appropriate questions, choosing suitable corpora, understanding the results, and refining their questions with subsequent searches. Other researchers (e.g. Mukherjee, 2006; Boulton, 2009a, b) argue that hands-on DDL represents an overwhelming leap for many learners and teachers, and they doubt whether this extremely autonomous corpus-based approach can be fruitful in the reality of ELT classrooms.

Another source of concern regarding autonomous learning in the DDL approach comes from the shift in the teacher roles. Despite the fact that there is no wholesale abandonment of teacher pedagogical actions in the mainstream DDL approach, Boulton (2009a, b) suggests that teachers may perceive DDL as a threat to their role, especially a loss of control, power and respect as the ultimate knower. The resources themselves (i.e. the corpora and software) constitute a matter of concern as well. They are perceived to offer "too many degrees of freedom [...] for the ordinary learner" (Schmied, 2006, p. 104). Additionally, computer rooms might be badly equipped, unavailable when needed, subject to breakdown, too small, have no technical backup, or simply be non-existent (ibid). Thus, Boulton (2008, 2009a, b, c, 2010) among others has suggested what is perceived as more pragmatic DDL i.e. printed corpus-data with associated activities. Boulton (2009b, 2010) justifiably argued that the use of printed corpus data associated with linguistic activities constitutes a form of DDL on the basis of the following points:

- Johns (1991), who is broadly considered as the father of DDL, made extensive use of printed concordances.
- The key element of the teacher acting as a research coordinator holds true for printed concordances. Boulton (2009b) points out that perhaps the prefabricated

materials do not entirely “cut out the middleman” as suggested by Johns (1994), but that the teacher takes on a new role as a director.

- Although providing handouts clearly lessens the scope for learners to be able to take greater responsibility for their own learning as a crucial feature of DDL, the main process still involves exploring the data, noticing patterns, formulating hypotheses and generalising to other situations. This suggests that the essence of autonomous learning is still present.

In alignment with Boulton’s views, Mukherjee (2006) points out that “DDL activities can be plotted on a cline of learner autonomy, ranging from teacher-led and relatively closed concordance-based activities to entirely learner-centred corpus-browsing projects” (p. 12). Researchers assert that autonomisation is itself a gradual process (Mukherjee, 2006; Boulton, 2009b) which can still be attained in a less than ideal DDL environment where concordances are provided as printed materials by teachers (Allan 2006).

In relevance to this research context, it should be noted that one of the main purposes for adopting a DDL approach to learning collocation is to foster autonomous learning and to gradually equip the learners with the necessary techniques before going on to explore corpora and utilise appropriate online software. This might also help the learners reach the stage at which they will be able to continue their language learning outside the classroom and perhaps even after they finish their education.

In brief, it is the current researcher’s belief that fostering a DDL approach in which the learners are provided with printed concordances from parallel English/ Arabic corpora could help them make a gradual autonomous exploitation of a wealth of genuine language instances to explore and observe the use of the target non-congruent collocations in both languages. It might also raise their awareness and increase their observation of differences between the forms and meanings of non-congruent collocations in both languages, thus resulting in better retention and production.

3.4.2 The rationale for contrastive analysis and translation

This section is intended to argue for the effectiveness of contrastive analysis (CA) as a pedagogical approach for teaching and learning non-congruent collocations. It also aims

at justifying the use of translation tasks as complementary instructional tools which help the learners to observe the difference between the target collocations in their L1 and L2.

3.4.2.1 Lexical Contrastive analysis: a cognitive perspective

In the 1960s, the notion of ‘contrastive analysis’ (CA) became mainstream. According to Lado (1957), contrastive analysis of two languages is a procedure which enables one to predict problems encountered by L2 learners or to explain errors made by them. Then, more effective language-learning materials, based precisely on these learning problems, can be developed (Hadlich, 1965). According to Lado (1957) the significance of CA for teaching for example, entails the teacher making comparisons between the learners’ native and foreign languages and predicting and diagnosing the difficulties the learners may encounter in learning linguistic patterns, in order to provide them with adequate materials. This notion of CA is based on the Contrastive Analysis Hypothesis (CAH) proposed by Lado in his influential work *Linguistics Across Cultures* (1957). According to this hypothesis, transfer and distribution of forms and meanings from the learners’ first language and culture to their foreign language and culture, both receptively and productively, has a major impact on L2 acquisition. Lado argued that “those elements that are similar to the [learner’s] native language will be simple for him, and those areas that are different will be difficult” (p. 2).

Despite the fact that Lado’s CAH may seem sound and credible, his interpretation of CA has been criticised on theoretical, empirical and pedagogical bases (James, 1980). The criticism of CA focused on its predictive and explanatory claims and on its behaviouristic-structuralistic rationale (Kupferberg, 1999). James (1980) considers CA to be an “interlinguistic” enterprise which perceives language not merely as a system to be described but as a system to be acquired. James reintroduces CA in cognitive terms as a process which takes place “when two languages come into contact in the bilingual brain” (James, 1996, p. 143). This process often leads to metalinguistic generalisations (transfer) about the target language, some of which may be incorrect. Consequently, James (1996) noted the need for learners to observe and notice the relationship between their native language and the foreign language so that (1) they can attain what he calls “cross-linguistic awareness”, which in turn may hinder erroneous generalization, and (2) they can convert input into intake necessary for learning. According to Schmitt (1990), there are several determinants of noticeability of a given aspect in the foreign language, namely

functionality, frequency, skill level, task demands and perceptual salience.²⁴ To James (1996) and in relation to CA, perceptual salience is the most important determinant factor of noticeability. He provides two potential sources for salience in any target language form: (1) the target language form itself could be inherently salient, thus, universally noticeable; (2) the salience may be contrast-dependent or cross-linguistic. Several empirical studies show that explicit instruction which induces input salience in the form of contrastive meta-linguistic input, and engages the learners' attention in various recognition and production tasks is conducive to the acquisition of difficult L2 forms (e.g. Ammar & Lightbown, 2005; Kupferberg, 1999; Kupferberg & Olshtain, 1996; Sheen, 1996).

By definition, CA is not restricted to one area of linguistic knowledge. However, in empirical research CA has been mainly applied in the area of grammar. Contrastive analysis in the area of vocabulary teaching and learning, i.e. lexical contrastive analysis, was initially rejected by Hadlich (1965). While he did not question the validity of contrastive analysis at the levels of syntax and pronunciation, he believes that the application of contrastive analysis to vocabulary learning is not only "incorrect", but could even be "harmful". Based on results obtained during the experimental development of elementary audio-lingual materials for Spanish, Hadlich (*ibid*) concluded that when pairs of words which are known traditionally and proved analytically to be problematic are juxtaposed, explained, contrasted and drilled, learners tend to continue confusing them. When they are presented as if no problem existed, students have little or no difficulty with them. Hadlich (1965, p. 427) further states:

"Words, after all, must be learned within the grammatical and situational restrictions of the second language. A word cannot be said to have been learned until the student can respond with it directly to the needs of communication, without external mediation... Therefore, no matter how it is presented, contrastive information... must be unlearned or at least ignored before a word can be really learned."

Hadlich's claims, however, could be refuted on different empirical and theoretical levels. Empirically, Laufer (2008a, b) argues that similarly to grammar, L2 cross-linguistic form-focused instruction which entails comparison with L1 and translation is advantageous to the area of vocabulary teaching and learning (see section 3.2.3 for details on empirical research supporting this assumption).²⁵ From a theoretical point of view about L2

²⁴ Schmidt's (1990) "Noticing Hypothesis" is discussed in more detail in section (3.5.1).

²⁵ It is worth noting that Laufer's (2008a, b) notion of contrastive analysis did not entail contrastive input, the cross-linguistic contrast was provided to the learners by the researcher.

acquisition, Selinker (1992) argues that L2 learners often conduct a cognitive inter-lingual comparison, or some kind of CA between the linguistic form they have noticed in the input, and knowledge of their native language. This suggests that some sort of L1 mediation takes place in the process of internalizing a given linguistic aspect. Therefore and in support of Ellis's (2008, p. 375) recognition that "acquisition and representation are inseparable", the current researcher argues that research on representations in the bilingual mental lexicon and psycholinguistic research on vocabulary acquisition could be used to refute Hadlich's (1965) claims. The next two sections will be allocated to presenting this argument.

3.4.2.2 Lexical transfer and the representations in the bilingual mental lexicon

'Lexical transfer' or 'cross-linguistic influence'²⁶ is defined as "the influence that a person's knowledge of one language has on that person's recognition, interpretation, processing, storage and production of words in another language" (Jarvis, 2009, p. 99). To a great extent, lexical transfer has an effect on the different dimensions of word knowledge including word use i.e. collocations (see chapter 2, section 2.5).

Research on lexical transfer is concerned with how different dimensions of word knowledge (form, meaning and use) relate to one another in the mind, and how lexical transfer operates in the minds of bilinguals and multilinguals. Jarvis (2009) distinguishes between two broad types of lexical transfer: the lemmatic transfer and lexemic transfer. The scope of lexemic transfer contains both the graphemic and phonological structure of a certain form of a word (Jarvis, 2009). On the other hand, the lemmatic transfer scope relates to the semantic (e.g. polysemy, synonymy, antonymy, etc.) and syntactic (e.g. a word's syntactic category and grammatical gender, etc.) and word properties (ibid). Collocational knowledge encompasses both syntactic and semantic specifications simultaneously, hence, it is part of the lemmatic transfer.

The consequences of lexical transfer, whether lemmatic or lexemic, can be seen in learners' and bilinguals' faulty and erroneous language use. According to Jarvis (2009), this negative transfer generally occurs through one of the two mental processes in the bilingual mental lexicon: (1) the construction of learned cross-linguistic associations and (2) processing interference. Learned cross-linguistic associations involve formed mental

²⁶ The terms are used interchangeably in the literature (cf. Jarvis, 2009; Jarvis & Pavlenko, 2008).

links between stored representations of lemmas (node words in this context) from two or more different languages. In contrast, processing interference could take place through the activation of words (lemmas) in one language when the speaker is trying to use another language (Jarvis, 2009). However, Jarvis (2009) credibly argues that none of the types of lemmatic transfer (including collocational transfer) seem to be induced to any significant degree by processing interference or activation levels. Instead all types of lemmatic transfer seem to result mainly from the ways that L2 users construct lexical representations in one language in accordance with their knowledge of corresponding words in another language. This argument by Jarvis (2009) seems compelling, with the construction of learned cross-linguistic associations being more relevant to the Revised Hierarchical Model (RHM), and the processing interference being more relevant to the Bilingual Interactive Activation Model (BIA) of the bilingual mental lexicon.

The BIA model is a model of bilingual word recognition based on the interactive activation model. It proposes that “proficient bilinguals activate information about words in both languages in parallel, regardless of their intention to function within one language alone” (Sunderman & Kroll, 2006, p. 391). This implies that the less proficient a bilingual is, the less the parallel activation occurs. The RHM, on the other hand, is “a developmental model that captures the interlanguage connections between lexical and conceptual representations as learners become more proficient in the L2” (Sunderman & Kroll, 2006, p. 392). The focus of this model is on how semantic representations²⁷ are developed and accessed during language processing.

The RHM suggests that lexical representations for words in each language are independent while their conceptual system is integrated. During the early stages of SLA, words in the L2 are assumed to be linked to their translation equivalents. The activation of the translation equivalent in L1 facilitates access to meaning for the new L2 words, because words in the L1 are hypothesised to correspond directly to their equivalents in the L2 (Sunderman & Kroll, 2006). Additionally, the model proposes that for all but the most proficient and balanced bilinguals, word-to-concept connections are stronger for the L1 than for the L2 (ibid). Thus, the model presumes that translation from the L1 to the L2 is more likely to be conceptually mediated (i.e. a trail of activation from the L1 word, to its associated concept, to the corresponding L2 word) (Sunderman & Kroll, 2006;

²⁷ Semantic representations involve mental links that map lemmas to concepts, and lemmas to other lemmas (e.g. collocations, synonyms) (Jarvis, 2009).

Jarvis, 2009; Kroll et al., 2010). The RHM further presumes that the strength of word-to-concept connections for the L2 increases and the presence of lexically mediated processing decreases, as proficiency in the L2 increases. To exemplify, Sunderman and Kroll (2006) pointed out that during the early stages of language learning, Spanish learners presumably associated the word *gato* to the translation equivalent *cat* in English. The English word *cat* will have advantaged access to the meaning; therefore, the word-to-concept connection is greater in the L1 than in the L2. According to the model's hypothesis, the connection between *gato* and the concept will strengthen and the dependency on the L1 translation equivalent will diminish with increasing proficiency in the L2 (ibid). In fact, RHM works hand in hand with Jiang's (2000) psycholinguistic model of an L2 vocabulary acquisition in instructional setting which will be presented in the next section.

3.4.2.3 Psycholinguistic model of an L2 vocabulary acquisition in instructional setting

In the second-language learning classroom, L1 and L2 lexical development processes differ significantly due to two practical constraints that L1 acquisition is not subject to. The first constraint is the lack of sufficient input in terms of quality and quantity. This poverty of input makes the extraction and creation of lemmatic and lexemic specifications about a word, and the integration of such information with the word's other specifications extremely hard, if not impossible, for L2 learners (Jiang, 2000). The second constraint in L2 learning is the existence of an established semantic/ conceptual system with an L1 lexical system closely associated with it. Because of the presence of the established L1 lexical system, L2 learners may rely on that system to learn new words in L2 (Jiang, 2000).

Given these constraints and based on Levelt's (1989) model of lexical representation, Jiang (2000) proposed a psycholinguistic model of L2 vocabulary acquisition. In this model, most L2 words go through three processing stages in lexical development. At the first stage, L2 words are initially mapped to L1 translations, not to meaning directly. For each time an L2 word is encountered, its L1 translation is activated to provide syntactic and meaning information (Jiang, 2000, 2002, 2004). The lexemic information (i.e. pronunciation, morphology and orthography) is gradually deactivated because it does not contribute to L2 word use. Strong links are established between L2 words and the lemmatic components of their L1 translations as experience in L2 increases (ibid). In other words, L2 words are no longer mapped to L1 translations but to L1 meaning

directly. The second stage is what Jiang (2000, 2002, 2004) considers as a unique process of form–meaning mapping in L2 vocabulary acquisition. He calls it “L1 lemma copying”. He also calls the resulting lexical use “L1 lemma mediation” (ibid). At this stage, the lemma spaces of L2 words are occupied by the lemma information from their L1 translations and the L2 processing is mediated by L1 lemma information. Jiang (2000) argues that once the semantic information is copied from the L1 translation, it stays in the L2 lexical entry and continues to mediate L2 word use even with continued exposure to the L2. Thus, L2 words will continue to be used on the basis of the semantic specifications of their L1 translations even by highly proficient L2 users. The third stage of this model is the “L2 integration stage” in which the syntactic, semantic and morphological specifications of a word are extracted from exposure and use and integrated into its lexical entry. However, Jiang cast a pessimistic view on reaching the alleged lexical competence in the third stage due to the aforementioned constraints in the L2 learning classroom. He suggests that most L2 words are fossilized in the second stage i.e. L1 lemma mediation stage. Nevertheless, the current researcher would argue that the notion of L1 mediation does not seem to constitute a crucial problem in the case of congruent collocations, because the transfer of knowledge from L1 would be mostly successful, resulting in correct combinations. In the case of non-congruent collocations on the other hand, the transfer of knowledge from L1 would mostly be unsuccessful negative transfer, resulting in erroneous combinations (see chapter 2 section 2.5 for empirical evidence).

3.4.2.4 Translation

For a considerable period of time and across different educational contexts and countries, translation was one of the key tools for teaching and assessing language competence, including vocabulary. Over time and with the emergence of different language teaching approaches, the use of translation as a teaching and assessment tool has gradually declined (Tsigari & Floros, 2013). The reasons for this decline were mainly related to: (1) false perceptions of the notion of translatability in connection with language pedagogy; (2) the equally false interpretations of the translation task as a common attempt to utilise a grammar-translation method to teaching a language; and (3) the insufficient attempt from translation studies to consider ways of informing other areas of language-related activity (Tsigari & Floros, 2013). However, translation is re-emerging as an important tool that serves the various purposes of language teaching and assessment (ibid).

In relation to the lexical domain and vocabulary teaching and learning, translation has long been classified as part of the ‘social strategies’ to learning vocabulary (as classified by Schmitt, 1997) in which the teacher is the source of information including translation in L1. It is considered to be the simplest way of providing definition and communicating word meanings (Nation, 2001). Moreover, the examination of words in a range of contexts and uses through translations, concordances or dictionaries is considered part of a rich instruction which involves learning the meanings, comparing and contrasting words, etc. (Nation, 2001). Translation tasks were also widely used in numerous studies as an assessment tool providing evidence of the learners’ receptive and productive vocabulary knowledge. They have been proven to activate different aspects and levels of language processing, such as awareness of similarities and differences between L1 and L2, distinguishing patterns in each language, increasing positive transfer ability, and enhancing mental flexibility and memorisation (cf. Belpoliti & Plascencia-Vela, 2013; Goundareva, 2011; Laufer & Girsai, 2008a, b; Machida, 2008).

It is of paramount importance to note that translation tasks are used in this research as a tool that is well established in the literature of vocabulary teaching and assessment. Most relevant for this research, the translation tasks are intended to be utilised as a complementary instrument to emphasise and raise learners’ awareness of a presumed automatic mental process that takes place in the bilingual mental lexicon (i.e. lexical transfer), and to conduct self-initiated lexical contrastive analysis with the help of the bilingual corpus-data. In fact, James (1996) points out that translation is a predominantly effective way to raise cross-linguistic awareness since “uniquely, in the act of translation two manifestations of MT and FL are juxtaposed and language juxtaposition is the very essence of Contrastive Analysis” (p. 147).

3.4.3 Summary

The above discussion on the rationale for a corpus-assisted contrastive analysis and translation approach for the acquisition of non-congruent lexical collocations leads to several concluding points: (1) CA is operationalised in terms of the cognitive processes, ignited by perceptual salience which is provided by cross-linguistic instruction/ input; (2) collocational transfer seems to be induced primarily by the ways that L2 users construct lexical representations in one language in accordance with their knowledge of corresponding words in the other language; (3) the construction of lexical representations presumably takes place through concept mediation and dependency on the L1 translation

equivalent as shown by the RHM of the bilingual mental lexicon and by Jiang's (2000) psycholinguistic model of vocabulary acquisition; (4) cross-linguistic awareness is crucial for the purpose of establishing the right lexical links, which may not be established through inherently salient target language forms, but through contrast-dependent salience. With these conclusions in mind, the current researcher hypothesises that providing learners with real language corpus- data which comprise texts in their L1 and L2 in juxtaposition, and engaging them in the act of translation is a form of CA that would result in the acquisition of the target non-congruent lexical collocations.

Whereas the previous sections focussed on the rationale for data-driven learning and contrastive analysis, the next section will apply this rationale specifically to the learning of non-congruent collocations from a theoretical perspective.

3.5 How collocation learning occurs with the corpus- assisted contrastive analysis and translation approach: theoretical underpinnings

Contrastive analysis is a fairly neglected pedagogical approach in any vocabulary teaching and learning practice, let alone a corpus-assisted one. This approach may not fit neatly into a single theoretical framework of SLA. However, in this research the current researcher would argue that the hypothesised positive results (i.e. learning of non-congruent collocations) of the proposed approach can be accounted for by a synthesis of hypotheses in SLA. The researcher would argue that hypotheses such as the ‘noticing hypothesis’, the ‘involvement load hypothesis’ and the ‘pushed output hypothesis’ could account for the cognitive processes whereby the learners engage with the linguistic environments as the prelude for learning. These hypotheses will be discussed in detail below, sections 3.5.1 - 3.5.3.

3.5.1 Collocation learning as a result of ‘noticing’

Awareness is a complex psychological construct comprising a number of different levels. Schmidt (1990) distinguished three levels of awareness: perception, focal awareness and understanding. Focal awareness, also referred to as ‘noticing’ (Schmidt, 1990) and ‘attention’ (Schmidt, 2001), is necessary in order to understand virtually every aspect of SLA as proposed by Schmidt (1990, 1993a, 1993b, 1994, 2001). In 1990, Schmidt proposed his influential ‘Noticing Hypothesis’, arguing that for the conversion of input into intake for learning, noticing is both necessary and sufficient. This hypothesis was later modified into a weaker version: more noticing leads to more learning (Schmidt, 1993b, 1994). However, the strong version of the hypothesis, favoured by Schmidt, had a significant extension claiming that noticing should be specifically focussed on the linguistic aspect to be learned rather than being global. He states “[n]othing is free... In order to acquire vocabulary one must attend to both word form (pronunciation, spelling) and to whatever clues are available in input that can lead to identification of meaning” (Schmidt, 2001, p. 30). The ‘noticing’ hypothesis thus provided the theoretical underpinning of FFI.

Schmidt and Frota (1986) claim that those learners who notice most learn most. However, the crucial question is what determines noticing of a linguistic aspect? As mentioned briefly in section (3.7.2.1) in relation to CA, there are several determinants of what learners will notice in the foreign language; expectation, frequency, perceptual salience, skill level and task demands (Schmidt, 1990, 2001). In relation to this research, three

types of noticeability determinants were available for learners in order to facilitate the process of converting the input into intake i.e. learning the target non-congruent collocations.

Frequency and establishing collocational connections

Frequency as the first relevant determinant for noticeability is believed to enhance the likelihood of an item to be noticed in input, and thus to be learned. The learning process as a result of input frequency in itself could be accounted for elsewhere within the connectionist approach to SLA as opposed to the generative approach. The generative model for SLA perceives language as a “separate faculty of mind” and knowledge about language as a “complex set of rules” (Ellis, 1999, p. 23). Thus, proponents of this approach perceive language acquisition as rule-governed. In the past two decades or so, many cognitive scientists have doubted these core assumptions underlying the generative model for SLA arguing that it lacks a plausible process explanation (Ellis, 1998). Thus, many have turned to connectionist models. The connectionist models perceive the mental representation of language as exemplar-based in which learning takes place on the basis of associative processes between elements and creating links between them. These links become stronger as these associations keep recurring (Mitchell & Myles, 2004). In language learning, connectionism argues that learners are sensitive to the frequency of occurrence of particular language forms, and that they extract probabilistic patterns on the basis of these re-occurrences (Lightbown & Spada, 2013; Mitchell & Myles, 2004). Advocates of connectionism subsequently analyse the way frequency and repetition influence, and eventually bring about, form in language, and the way this knowledge affects language comprehension and production (cf. N. Ellis, 2002; Ellis et al., 2008).

The current researcher would argue that learning collocations within this model²⁸ and as a result of frequency of input may entail forming, restructuring or strengthening the connections between the node words and its collocates through an intensive exposure to the collocations in the corpus data. In fact, input frequency is at the heart of DDL as Jaen (2010, p. 18) nicely puts it:

“We believe that Data-Driven Learning is an approach particularly suitable not only to help students notice and explore linguistic patterns which are made salient by the

²⁸ It is of paramount importance to note that the current research accords with the connectionist model to SLA only inasmuch as it accounts for the mental representation and associative learning and collocation acquisition mechanism. It did not employ the research methodology conventional to connectionists, namely simulations of language acquisition which are run using computer models comprising many artificial neurons connected in parallel (Ellis, 2003).

concordance because of their frequency and stability, but also to make them aware of the combinations which are not naturally used by native speakers.”

However, this researcher would quote Schmidt (1990, p. 152), “noticing depends on more than input frequency”, and would argue that learning non-congruent collocations with a corpus-assisted contrastive analysis and translations approach has more to offer than just frequency of input as argued in the next sections.

Perceptual salience and cross-linguistic input

Salience of input (perceptual salience in particular) is another determinant of what is noticed by learners as argued by Schmidt (1990, 2001, 2003a). According to Ellis (2006) the linguistic forms that L2 learners fail to take on and to use routinely in their second language processing are those which, though available as a result of frequency, context or recentness, fall short of intake because of several associative learning factors. One of the factors involved in this is low salience of co-occurring forms, which results in failure in the selection process of these particular linguistic forms according to the connectionist learning research summarized by Ellis (2006). The main question that arises here is, are collocations inherently salient, and do they thus by definition attract learners’ attention in the absence of intervention by any external party?

When it comes to the salience of collocations in language input, this researcher would argue that collocations are not inherently salient enough to be noticed by learners (see section 3.2.3 for empirical evidence). Lewis (2000) warns: “Do not assume students are noticing collocations and recording them for themselves. They won’t unless you train them to” (p. 163). However, the most significant argument in support of low collocational salience is manifested in Wray’s (2000, 2002) model of adult second-language learners’ acquisition of collocations. In this model, Wray argues that in acquiring collocations, L2 learners adopt an analytic word-focused mode of processing, as opposed to a holistic mode of processing by which natives and young L2 learners acquire collocations. To illustrate her model, Wray (2002) suggested that on encountering a collocation like *major catastrophe*, the adult language learner would break it down into a word meaning ‘big’ and a word meaning ‘disaster’ and store the words separately, without any awareness of the pair’s association. Later, when they need to express the idea again, they would have no memory of *major catastrophe* as pair, and any combination of words with the right meaning would seem equally possible. Some of these pairs would be native-like, others would not. Accordingly, Wray (2002) suggested that for adult L2 learners collocations

are “separate items which become paired.” Therefore, they do not usually establish the appropriate “strength of association” between words. Another conclusion to be drawn from Wray’s (2002) model is that the mental links between constituents of collocations are weak due to the lack of their salience as chunks in the L2 mental lexicon. This means that it is necessary to induce salience externally so that collocations could be noticed as linguistic chunks, as advocated by many researchers (e.g. Ellis, 1996, 1997; Ellis et al., 2008; Lewis, 2000; Siyanova & Schmitt, 2008). According to Ellis (2006), FFI which involves selective attention and awareness raising is a pedagogical reaction to low salience of L2 forms, such as collocations (see section 3.2. for an overview of FFI).

In relevance to this research, one source of perceptual salience within FFI is particularly important in the teaching and learning of non-congruent collocations i.e. perceptual salience that is contrast-dependent proposed by James (1980, 1996, 2005). James (ibid) argues that this type of cross-linguistic input salience would help the learners spot the snag and suppress the mother-tongue transfer. Since EFL learners’ collocational knowledge is largely plagued with negative transfer from L1, especially in producing non-congruent collocations,²⁹ the current researcher, as informed by the connectionist model of SLA, would argue that the corpus-assisted contrastive analysis and translation approach would have a positive impact on suppressing this transfer. It would hypothetically reconstruct and strengthen the mental links between the node words and their collocates, thus helping the learners to establish them as lexical chunks in their mental lexicon. In addition, with the perceptual salience of collocations provided by the proposed contrastive approach in this research, the L2 learners may find it easier to become aware of more or less fossilised characteristics (i.e. collocations) of their interlanguage, thus potentially initiating a process of knowledge restructuring. The parallel English/ Arabic corpus data is one way of providing this contrast-dependent perceptual salience as noted by Bernardini (2004, p. 40):

“The ease of access to instances of language performance makes it possible for learners to rely less on one or two individuals with their idiosyncrasies and their limited intuitions. If they can also work with corpora in their native language, this may convince them of unreliability of their own intuitions about their mother tongue, resulting in a highlighted attention to un(typical) ways of saying in any languages they know.”

²⁹ See section 2.5 for a detailed overview.

Since the translation tasks are another way of providing contrast-dependent perceptual salience, this researcher hypothesises that this corpus-assisted contrastive analysis and translation approach would constitute an important addition or alternative to standard DDL. Most significantly, it would develop the idea that authenticity may be a condition of the learners' engagement with the corpus data or the perception that the corpus data is relevant to their concerns.

The last of Schmidt's (1990) determinants of noticing as relevant to this research is task demands. Schmidt argues that certain tasks may make certain language forms (collocations) salient. Task demands according to Schmidt offer one of the fundamental arguments that what is learned is what is noticed. The demands of translation and corpus consultation tasks, which according to my hypothesis result in learning, will be more rigorously discussed in the following sections.

3.5.2 Collocation learning and the 'Involvement Load' hypothesis

The 'task-induced involvement load' hypothesis (Laufer & Hulstijn, 2001) proposes that retention of previously unfamiliar words is conditional upon the amount of learners' involvement while processing these words. Involvement is operationalised by tasks designed to differ in three motivational and cognitive dimensions: need, search and evaluation. The 'need' dimension is the motivational component of involvement. Laufer and Hulstijn (*ibid*) convincingly argue that there is at least one theory that explicitly incorporates the dimension of 'need' under motivation i.e. need creates tension. The researchers point out that a mild degree of tension could have a positive impact on information processing, and thus could indirectly affect learning. Accordingly, the dimension of 'need' exists in a task when the lexical item is perceived to be necessary for task completion.

On the other hand, 'search' and 'evaluation' are the two cognitive dimensions (information processing) of involvement. They are presumed to be dependent upon noticing and intentionally allocating attention to the form-meaning relationship. The researchers identified 'search' as the attempt to find the L2 lexis form expressing a given concept, or, conversely, to find the meaning of an unknown L2 lexis. Laufer and Girsai (2008a, b) suggest that examples of 'search' involve: trying to find the L2 translation of an L1 lexis by consulting a dictionary, or trying to guess and infer the meaning of an L2 lexis from context. Additionally, 'evaluation' denotes some sort of selective decision

about the lexis' form or meaning, based on a criterion of semantic and formal appropriateness of the lexis and its context. It also involves a comparison of a given lexical item with other lexical items, a specific meaning with its other meanings, or the lexical item with other lexical items, in order to assess whether a word does or does not fit its context. Laufer and Hulstijn (2001) have suggested two degrees of prominence for the 'evaluation' dimension i.e. moderate and strong. A 'moderate evaluation' involves recognising differences between lexical items (as in a fill-in task), or differences between several senses of a lexical item in a particular context. On the contrary, a 'strong evaluation' entails a decision as to how additional lexical items will combine with the new ones in an original sentence or text.

The proposal of the involvement hypothesis by Laufer and Hulstijn (2001) was an attempt to operationalise several concepts that have been used in connection with good retention, one of which is depth of processing. Originating from cognitive psychology, the depth of processing hypothesis proposed by Craik and Lockhart (1972) suggests that the likelihood of a piece of information to be committed to long-term memory depends on the depth with which it is initially processed in short-term memory. Based on this, the researchers have two further assumptions:

- Words processed with a higher involvement load will be retained better than those processed with a lower involvement load;
- Tasks designed with a higher involvement load will better facilitate vocabulary retention than those with a lower involvement load.

Empirical evidence in full or partial support of the involvement load hypothesis has been suggested (e.g. Hill & Laufer 2003; Laufer & Girsai, 2008a, b; Laufer & Hulstijn 2001; Tsai, 2011; Webb 2005). In her experimental studies, Laufer (2008a, 2008b) argued that translation tasks embody the three elements of need, search and evaluation. They entail 'need', because the words that have to be understood (when translating into L1), or produced (when translating into L2) are predetermined by the task. The element of search is present when learners, being unfamiliar with the L2 words, have to perform a search for their meaning when translating into L1, or a search for their forms when translating into L2. Most significantly, Laufer and Girsai (2008a, b) argued that an element of evaluation is necessary to perform a translation activity since there is typically more than one translation alternative for a particular sentence. Therefore, in carrying out a translation task, learners have to choose the alternative that fits the text they create. In the

process of translation into L2, the learners' decision will be based on the way other words in the text combine with the new word. According to the model of involvement, Laufer and Girsai (ibid) maintain that the evaluation element is strong. Finally, they postulate that translation tasks can be effective in vocabulary learning, since they are tasks with high involvement load.

This is very much true in relation to this research, however, due to the autonomous nature of the corpus-assisted contrastive analysis and translation approach, the involvement load is presumably higher than the load induced by the act of translation only. Similar to Laufer and Girsai's (2008a, b) argument, the 'need' component is present as the learners will have to understand the target collocations in order to translate them into L1, or produce them when translating into L2 as required by the task. In the 'search' component, the learners in Laufer and Girsai's studies (ibid) inferred meanings of the target lexical items from context or asked the teacher for them. Later they received a contrastive cross-linguistic explanation of the target lexical items provided by the teacher. As a result, this researcher considers the 'search' component of the load to be relatively moderate. In the autonomous approach proposed in this research, the learners are expected to 'search' and 'evaluate' the meanings and forms of the non-congruent target collocations through the consultations of bilingual corpus data. Accordingly, this researcher would hypothesise that the task design entails high involvement load engaging the learners in deeper processing which would render better retaining of the target non-congruent collocations.

3.5.3 Collocation learning and the 'Pushed Output' hypothesis

There is a broad agreement among most language learning researchers that output is necessary to increase linguistic competence i.e. L2 learners must practise producing the language if they are to learn to use their interlanguage system routinely and confidently (Mitchell & Myles, 2004). Based on her observational data of the French immersion program in Canada, Swain (1985) argued that part of the learners' inability to exhibit a full mastery of French, especially in speaking and writing skills, was that they had had little opportunity to engage in producing 'comprehensible output' through negotiation of meaning. According to Swain (1985), negotiating meaning involves the notion of being 'pushed' toward not only a mere conveyance of a message, but rather a precise, coherent, and appropriate delivery of that message. She tentatively used the term 'output' to include speaking, writing, collaborative dialogue and/ or verbalizing tasks, which serve the language learning process through different functions (Swain 2000). Other than the

typical 'practice' function, the 'pushed output' hypothesis proposed by Swain advocates three functions for learner output. The first one is the noticing/ triggering function. The claim behind this function is that learners might notice that they do not know how to write or say precisely what they wish to convey. This awareness about their linguistic gap would trigger cognitive processes which push learners to generate new linguistic knowledge or to consolidate their current existing knowledge. The second function is hypothesis testing, which serves the claim that "output may sometimes be, from the learners' perspective, a "trial run" reflecting their hypothesis of how to say (or write) their intent" (Swain, 2005, pp. 476). The third role of the output is the metalinguistic/ reflective function. Using language produced by the self or by others to reflect on language mediates second language learning. Swain and Lapkin (1995) confirm that the cognitive processes (identified in their experiment) represent processes similar to those hypothesised by other theorists and researchers which involve extending L1 knowledge to L2 contexts, extending L2 knowledge to new target-language contexts, and formulating and testing hypotheses about linguistic forms and functions.

In relation to vocabulary acquisition and lexical competence, a number of studies have given the 'pushed output' qualified support (e.g. de la Fuente, 2002; He and Ellis, 1999). Additionally, there is evidence from empirical research that output tasks were more effective when compared to input tasks or activity-based tasks for the purpose of learning new words (Browne, 2002; Laufer & Girsai, 2008a, 2008b). According to Laufer and Girsai (*ibid*), translation into L2 is a manifestation of pushed output. Using the pushed output hypothesis in this research means that (1) the subjects were required to actively produce language in order to translate (2) a translation was only considered good if they used, rather than avoided problematic words or structures (i.e. collocations), like in the case of free production (3) Upon encountering a gap in their lexical knowledge, subjects were required to engage in a thinking process in which they extend their L1 knowledge to L2 context, extend their L2 knowledge to new target language contexts (4) they are also expected to engage in restructuring, testing their new knowledge and reflecting on their previous knowledge about the target non-congruent collocations with the help of the bilingual corpus-data. Hence, this researcher would hypothesise that the corpus-assisted CAT might be an effective pushed output task for learning the non-congruent collocations.

3.5.4 Summary

As suggested earlier, collocation learning with a corpus-assisted contrastive-analysis and translation approach does not fit neatly into a single theoretical framework of SLA. This researcher argued that it could be accounted for by a synthesis of hypotheses in SLA. Informed by the connectionist model to SLA, the ‘noticing’ hypothesis provides a sound basis for the pedagogical use of corpora and concordance-based tasks as a prelude to collocation learning. The connectionist model perceives word knowledge as one of the neural networks with complex clusters of connections. The connections within the mind are strengthened by repeated exposure to exemplars. This could be achieved by being exposed to a wealth of real language in corpora. The contrast-dependent perceptual salience provided by the bilingual corpus-data could also account for the strengthened mental connections and representations of the target collocations. The ‘involvement load’ and the ‘pushed output’ hypotheses are viewed by the current researcher as being interdisciplinary with the ‘noticing’ hypothesis. However, more grounded in information processing approaches to SLA, the two hypotheses provide theoretical underpinnings for this cross-linguistic form-focused approach, and shed light on the autonomous processes in which the learner cognitively manipulates translation tasks and corpus- data in working memory which may result in learning.

3.6 Research hypotheses

Based on the literature and the theoretical underpinnings reviewed in the preceding sections the following hypotheses in favour of the proposed corpus-assisted contrastive analysis and translation approach were formulated.

H1. The corpus-assisted CAT condition will lead to the learning of a significantly larger number (if any) of adj. /noun collocations than the non-corpus-assisted CAT condition.

- d) The corpus-assisted CAT condition will lead to the passive recall of a significantly larger number (if any) of adj. /noun collocations than non-corpus assisted CAT condition.
- e) The corpus-assisted CAT condition will lead to the active recall of a significantly larger number (if any) of adj. /noun collocations than the non-corpus assisted CAT condition.
- f) The differences between the conditions in active and passive recall (if any) will be retained in a delayed post-test.

H2. The corpus-assisted CAT condition will lead to the learning of a significantly larger number (if any) of adj. /noun collocations than the corpus-assisted non-CAT condition.

- d) The contrastive analysis and translation conditions (both) will lead to the passive recall of a significantly larger number (if any) of adj. /noun collocations than the non-contrastive and translation tasks.
- e) The corpus-assisted CAT condition will lead to the active recall of a significantly larger number (if any) of adj. /noun collocations than the corpus-assisted non-CAT condition.
- f) The differences between the conditions in active and passive recall (if any) will be retained in a delayed post-test.

Chapter 4: Methodology

In the previous chapters, I have provided a rationale for employing a corpus-based contrastive analysis and translation approach to the teaching/ learning of non-congruent adjective/ noun collocations. It was hypothesised that the proposed approach will result in the learning of non-congruent collocations both receptively and productively. This chapter details the research methodology employed to test this hypothesis. It starts with the paradigm and philosophical stance which informed the design in section 4.1. The characteristics and sampling of the participants are described in section 4.2. A detailed account of collocation extraction and the attempt to establishing non-congruency in Arabic, as well as a description of the experimental instruments will follow in section 4.3. The procedures followed in all experimental groups are presented in section 4.4. Sections 4.5 and 4.6 are allocated for the data collection methods, measures, and data analysis methods used to address and test the research hypotheses. It will be followed by validity and reliability issues in section 4.7 and ethical considerations in 4.8. The chapter concludes with a summary of the research methodology 4.9.

4.1 The philosophical stance

Many researchers argue that there is no need to worry about the philosophical stance of research as this is best dealt with by philosophers who have time to devise theories of being and knowledge (Grix, 2004). However, many others take different views (e.g. Clough & Nutbrown, 2002; Grix, 2004). They believe that researchers need to know the core assumptions that underline their work. These assumptions should inform their choice of research questions, methodology, methods and even sources, if they are to present clear, precise and logical work, and engage with and debate other work. In fact, Mackenzie and Knipe (2006, p. 2) state that “without nominating a paradigm as the first step, there is no basis for subsequent choices regarding methodology, methods, literature or research design.” Interestingly, the previous notion of a ‘research journey’ is quite controversial since many researchers think of it as cyclical or multidimensional rather than linear. The elements forming the basis of a research process are another debatable matter. According to Grix (2004), ontological and epistemological assumptions form the foundations of the whole research’s edifice, whereas methodology, methods and resources are closely connected to and built upon ontological and epistemological assumptions.

Ontology is defined as the study of “claims and assumptions that are made about the nature of social reality, claims about what exists, what it looks like, what units make it up and how these units interact with each other. In short, ontological assumptions are concerned with what we believe constitutes social reality” (Blaikie, 2000, p. 8). There are two main ontological positions: those based on foundationalism and those based on anti-foundationalism. Foundationalists believe that reality exists independently of our knowledge of it (Grix, 2004), and there are central values that exist and that can be rationally and universally grounded (Flyvbjerg, 2001). Conversely, central to the anti-foundationalist view are the beliefs that: (1) the existence of the world depends on our knowledge of it; (2) reality is socially and discursively constructed by human actors; (3) there are no central values that can be rationally and universally grounded (Grix, 2004).

Epistemology, on the other hand, is “concerned with the theory of knowledge, especially in regard to its methods, validation and the possible ways of gaining knowledge of social reality, whatever it is understood to be” (Grix, 2004, p. 63). In short, epistemology is concerned with “claims about how what is assumed to exist can be known” (Blaikie, 2000, p. 8). Epistemological positions are divided into two overarching positions, namely positivism and constructivism. Positivism “is an epistemological position that advocates the application of the methods of the natural sciences to the study of social reality and beyond” (Bryman, 2001, pp. 11-12). Constructivism “is predicated upon the view that a strategy is required that respects the differences between people and the objects of the natural sciences and therefore requires the social scientists to grasp the subjective meaning of social action” (ibid). Grix (2004), among other authors (e.g. Mack, 2010; Mackenzie & Nipe, 2001), advocates a logical directional relationship between the key building blocks of research i.e. ontology, epistemology, methodology, methods and sources. He states:

“It is of paramount importance that students understand how a particular view of the world affects the whole research process. By setting out clearly the interrelationship between what a researcher thinks can be researched (her ontological position), linking it to what we can know about it (her epistemological position) and how to get about acquiring it (her methodological position), you can begin to comprehend the impact your ontological position can have on what and how you decide to study.” (p. 66)

Grix (2004) also asserts that choosing one of these epistemological positions will lead a researcher to employ a different methodology than he/ she would if they choose the other. Another point made by Grix (2004) worth mentioning here is his argument that ontology is often wrongly collapsed together with epistemology, with the former simply seen as a part of the latter. Despite the fact that they are closely related they need to be kept separate (ibid).

A different group of researchers (e.g. Creswell, 2014; Crotty, 2003) have different views. Crotty (2003) believes that methods, methodology, theoretical perspective and epistemology constitute the basic elements of any research process. He defines the theoretical perspectives as “the philosophical stance informing the methodology and thus providing a context for the process and grounding its logic and criteria” (p. 3). Theoretical perspectives, as seen by Crotty, *roughly* correspond to Blaikie’s use of the term ‘ontology’, and refer to how one views the world. In addition, epistemology is defined as “the theory of knowledge embedded in the theoretical perspective and thereby in the methodology” (Crotty, 2003, p. 3). It is about “the nature of the relationship between the knower or would be knower and what can be known” (Guba and Lincoln, 1998, p. 201). Moreover, epistemology is concerned with providing a philosophical grounding for deciding what kind of knowledge is possible and how researchers can ensure that they are both legitimate and adequate (Maynard, 1994).

It needs to be noted that Crotty (2003) believes that ontology should be placed alongside epistemology in informing the theoretical perspective, and that ontological issues and epistemological issues tend to emerge together. This is because each theoretical perspective represents a certain way of understanding *what is* (ontology) along with certain way of understanding *what it means to know* (epistemology) (ibid). Crotty confirmed that writers in the research literature have trouble keeping ontology and epistemology apart conceptually. For example, realism (an ontological view proclaiming that realities exist outside the mind) is often taken to suggest objectivism (an epistemological view proclaiming that meaning exists in objects independently of any consciousness). Crotty (2003) suggests that scholars such as Heidegger and Merleau-Ponty, who frequently invoke a ‘world always already there’, are far from being objectivists. We may assume that the world is there whether human beings are aware of it or not. However, a world with no conscious beings to engage with it is still an intelligible world and not a world of meaning (Crotty, 2003). It becomes a

world of meaning only when meaning-making beings make sense of it (ibid). Hence, an ontological stance of realism and an epistemological position of objectivism turn out to be compatible.

Given that state of affairs, Crotty plausibly suggests that researchers can deal with the ontological issues without expanding their schema to include ontology, a viewpoint adopted by the current researcher. Thus, I will apply the term ‘paradigm’ to refer to “a shared set of ontological and epistemological assumptions and their attendant methodological principles concerning how to conduct a research” (Lynch, 1996, p. 13). Assumptions, concepts or propositions are presumed to be logically related and orient thinking and research (Crotty, 2003).

The following sections will be dedicated to briefly depicting the entwined relationship between the key building blocks of the current research.

4.1.1 The paradigm of this research

The present study was guided by a symmetrical philosophical stance: ontologically and epistemologically committed to a foundationalist/ positivist position. As opposed to constructivism, positivism is based on several significant premises summarized by Grix (2004) as follows:

- It is based on a realist, foundationalist ontology which perceives the world as existing independently of our knowledge of it.
- A positivist view believes in the possibility of making causal accounts. Hence, many positivists seek to employ scientific methods in their analysis of the social world as they seek objectivity in research.
- Positivists place great emphasis on explanation in research as opposed to understanding.
- They believe that we can establish regular relationships between social phenomena by using theory to generate hypotheses, which can be then tested by direct observation. Thus, they emphasize verificational and observational dimensions of empirical practice.

The current researcher generally views language as an artefact co-constructed by its speakers, and language learning as a constructed reality. This is not in a contradictory position with the positivist ontological and epistemological standpoint of this research.

As indicated in the previous section, reality (i.e. language) is out there, however, not without conscious meaning-making beings (i.e. speakers/ learners) making sense of it. In that sense too, collocation learning is a reality existing and occurring out there in the world. It is this researcher's aim to try to investigate and explain how reality occurs under certain proposed conditions rather than trying to understand how it comes into being in the first place. Grix (2004) suggests that from positivist to constructivist viewpoints, researchers range from those seeking to explain social reality to those attempting to interpret or understand it. Hence, in this researcher's view it is of paramount importance to distinguish between constructivism as a 'theory of knowledge acquisition by learners' as adopted by applied linguists on the one hand, and constructivism as an epistemological view that enables the researcher (as a knower or would-be knower) to know what can be known in a legitimate and adequate means of investigation, on the other.

To that end, a positivist methodology has been employed in this research. Empirical substantiation of knowledge through objective manipulation and control over variables is a typical characteristic of positivist work. These investigations are concerned with causal inferences as stated earlier in this section. In the next section, I will discuss how these causal inferences will be investigated in relation to this research.

4.2 Overall research design

Many research studies in applied linguistics are intended to establish unambiguous causal links through the application of experimental research designs (Dörnyei, 2007). Experimental design in its true sense involves a random assignment of participants in two types of groups; 'experimental groups' which are exposed to a particular treatment or condition and 'control groups' which are similar to the experimental group in every aspect except for the exposure to that special condition. Any differences in the comparison between the results of the two groups should be attributed only to that particular condition (Johnson and Christensen, 2004). Unfortunately, in educational contexts such true experimental designs and tightly controlled research environments are very rarely feasible and therefore the common method uses intact class groups i.e. quasi-experimental designs (Dörnyei, 2007).

To serve the purpose of this research, a pre, post and delayed post-test quasi-experimental design was devised, or to be more precise, a non-equivalent (pre-test/

post-test) control-group design. This design involves the selection of an experimental group A and a control group B without random assignment (Creswell, 2014). Both groups take a pre-test and post-test, but only the experimental group(s) receive the treatment. In this research, three experimental groups received different treatments of collocation instruction while the control group was given the pre, post and delayed post-tests but received no treatment at all. The control group was included in order to provide a baseline for comparison.

To address and test the research hypotheses, measurements of the participants' collocational knowledge were taken two weeks before, immediately after, and finally three weeks after the treatments. Collocational knowledge was measured at two levels: active recall (Arabic-English translation) and passive recall (English-Arabic translation) (see Figure 4.1).

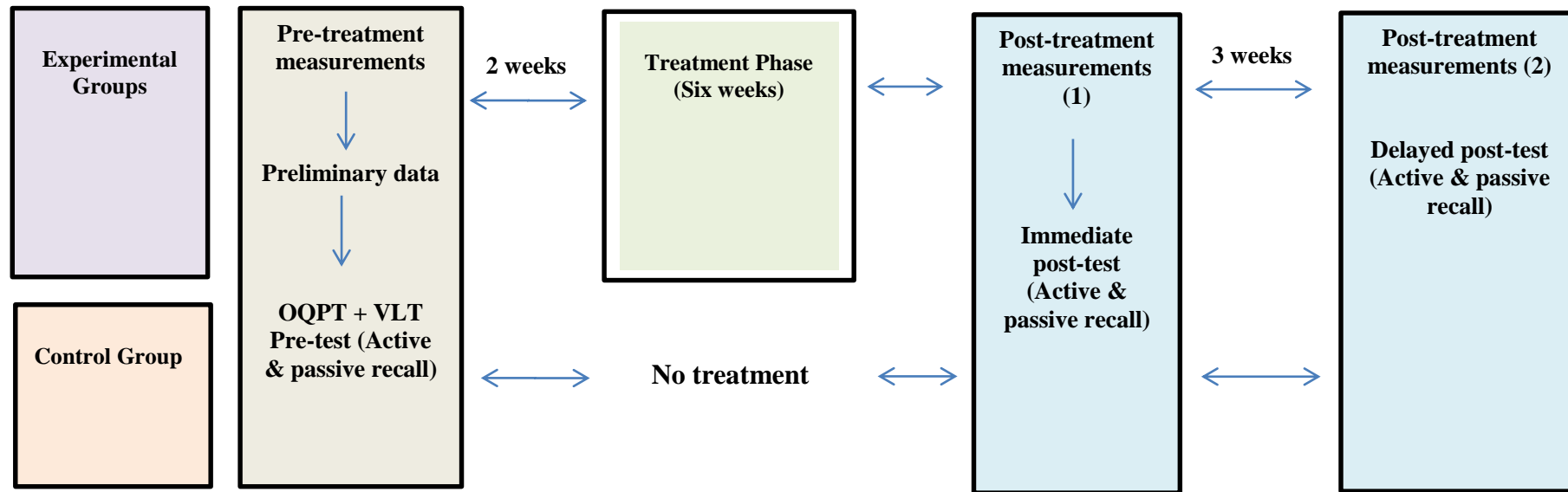


Figure 4. 1: The overall research design and data collection procedure.

4.2.1 Participants

At the outset of this study, 177 female undergraduate EFL students at a university in Saudi Arabia expressed their willingness to participate in the experiment. Classroom context researches are prone to participant attrition; unfortunately, this context was no exception. The number of students greatly decreased at an early stage of the research due to their low scores in the vocabulary level test VLT (the most frequent 2000 (K2) and 3000 (K3) words) i.e. they scored less than 13/30 in either or both of the levels ($N= 16$). Another group of students were eliminated due to their repeated absences throughout the treatment phase or due to their absence in the testing phase ($N= 32$).

The remaining 129 participants were first and second-year EFL students majoring in English. They ranged between 18-20 years of age, and had never lived in an English-speaking country. They are homogeneous in the fact that they all speak Arabic as their mother tongue. Moreover, English in Saudi Arabia is taught in public schools starting from the first year of middle school. Thus, the participants' English backgrounds were similar since they had studied English for six years prior to entering university and have been exposed to the language from an average age of 11-12.

4.2.2 Sampling

As mentioned earlier, this research follows a non-equivalent control group design, thus randomization as in true experimental design was not attainable. Moreover, the use of intact classes in quasi-experimental design is favourable in many educational research settings because it causes less disruption to the existing school system (Porte, 2002). Therefore, cluster random sampling of participants was employed in this research. Cluster random sampling involves selecting groups (e.g. intact second/foreign language classes) to serve as participants rather than individuals (Mackey & Gass, 2005). The current researcher had access to four intact classes that had been assigned by the University administration. Initially, all these students were allocated to the study. Later, students in the four classrooms were randomly assigned to experiment group 1 (-DDL +CAT, $N= 33$), experiment group 2 (+DDL +CAT, $N= 32$), experiment group 3 (+DDL -CAT, $N= 32$) and a control group ($N= 32$). Prior to this assignment, it was crucial to make sure that any variation in research results

between groups could not be attributed to variations in the participants' English proficiency levels or vocabulary knowledge levels. To address these two issues, two commonly used and freely available tests were administered to the participants: a Quick Oxford Placement Test (QOPT) and the Vocabulary Level Test (VLT) by Schmitt, Schmitt and Clapham (2001, version 2). Another reason for the VLT and for making sure that all participants achieved a similar level in lexical coverage and vocabulary knowledge is the fact that the students were required to carry out translation tasks. It was thus necessary to have an insight into whether or not they were likely to have the lexical resources necessary to cope with the translation tasks, both receptively and productively.

1. *QOPT*

Given their educational background as mentioned earlier, students at this academic level (year one and two) were expected to be mainly of intermediate level of English language proficiency. In order to validate this claim, all participants in each of the four groups were given the QOPT. The placement test scores showed that the majority of participants in each group were of lower-intermediate level of English proficiency i.e. they scored between 30 to 39 out of 60. They also showed that each of the groups had a number of participants of an upper-intermediate level (scored between 40 to 47 out of 60). However, two statistical tests (i.e. Kruskal-Wallis and Chi-Square)³⁰ showed that there was no significant statistical difference in the QOPT scores between the groups ($p > .05$), and no significant difference in the distribution and number of students of lower or upper intermediate levels of English proficiency in each group ($p > .05$) (see Table 4.2 and Figure 4.3).

Table 4. 1: Descriptive statistics and normality test (QOPT)

Parameter	Groups	N	Mean	Median	Std. Deviation	Min	Max	P-value
QOPT	Group1	33	37.15	37	5.15	30	47	.014
	Group 2	32	36.94	36	4.04	30	47	.119
	Group 3	32	36.25	33	6.02	30	47	.000
	Group 4	32	36.50	35	4.61	30	47	.022

³⁰ The normality of distribution of the data was checked using Shapiro-Wilk test before carrying out data analysis in order to choose the most appropriate statistical test (see chapter 5, section 5.1 and sub-section 5.1.1 for a detailed overview of normality of distribution assumption and Shapiro-Wilk test). The Chi-Square test for independence is used here to determine if two categorical variables (upper and lower) are related as it compares the frequency of cases found in the four groups.

Table 4. 2: QOPT (Kruskal-Wallis test for between groups comparison)

Parameters	Groups	N	Mean Rank	P-value
QOPT	Group 1	33	67.00	0.556
	Group 2	32	70.45	
	Group 3	32	57.52	
	Group 4	32	64.97	

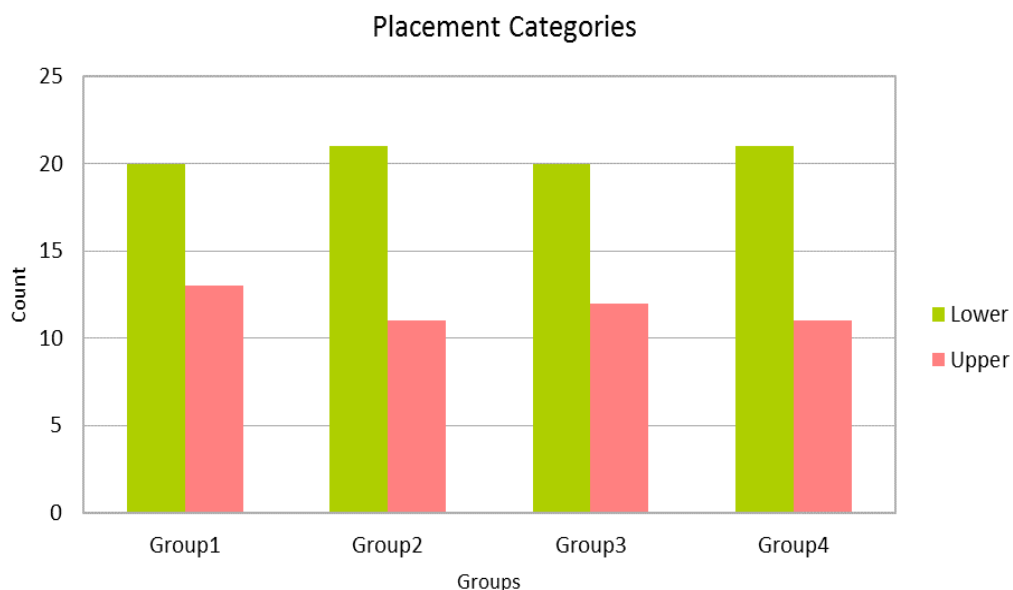


Figure 4. 2: Placement test categorisation (Chi-Square test)

2. VLT

Schmitt, Schmitt and Clapham's (2001) vocabulary level test (K2 and K3, version 2) was given to the participants. On the K2 test, the groups achieved the mean scores of 25.42, 24.81, 24.50 and 24.72. On the other hand, they achieved the mean scores of 19.85, 18.94, 20.09 and 19.91 on the K3 test (see table 4.3). A Kruskal-Wallis' statistical test was run to check if there were any statistical differences between the groups on each VLT. The results revealed that the differences between participants' scores on both tests between in all four groups were not statistically significant (K2 and K3 $p > .05$) (see table 4.4 below).

It is of crucial importance to state that there was no attempt to control for knowledge or lack of knowledge of the words comprising the target collocations due to time constraints in the classroom context. Therefore, the VLT scores were used as a

periphery and baseline for their lexical coverage based on Read's (1988) and Schmitt's et al.'s (2001) arguments that knowing lower-frequency words tends to imply knowing higher-frequency ones.³¹

Table 4. 2: VLT Descriptive Statistics and Normality test

Parameter	Groups	N	Mean	Median	Std. Deviation	Min	Max	(P-value)
K2	Group1	33	25.42	25	3.10	15	30	.011
	Group 2	32	24.81	25	3.65	17	30	.050
	Group 3	32	24.50	25	3.41	18	30	.097
	Group 4	32	24.72	25	3.71	17	30	.047
K3	Group 1	33	19.85	19	3.80	14	28	.049
	Group 2	32	18.94	18.50	3.96	13	28	.166
	Group 3	32	20.09	19	4.07	15	29	.024
	Group 4	32	19.91	19	4.34	14	28	.029

Table 4. 3: VLT (between-groups comparison)

Parameters	groups	N	Mean Rank	P-value
K2	Group 1	33	70.45	0.765
	Group 2	32	64.72	
	Group 3	32	60.58	
	Group 4	32	64.08	
K3	Group 1	33	67.23	0.701
	Group 2	32	58.36	
	Group 3	32	68.39	
	Group 4	32	65.95	

4.3 Materials

In this section, the materials used in the research will be presented. These include the extraction and selection of the target non-congruent collocations, the worksheet for all the experimental groups, and the design of the corpus-data sheets.

4.3.1 Extraction and selection of the target collocations

As reviewed in chapter 2 (section 2.7.3) a complementary approach used by Sonbul (2012) was adapted to define collocations. A collocation is thus defined both from statistical and phraseological viewpoints. Statistically, collocations are defined as

³¹ All constituent words of the target collocations belong to the (K1) and (K2) levels as they appear in either the BNC or COCA.

two-word pairs which co-occur above chance (i.e., with a minimum frequency of five occurrences and a minimum MI score of 1). Phraseologically, collocations are basically non-idiomatic two-word pairs for which native speakers show a degree of sensitivity to usage restrictions, and which Arabic native speakers would perceive as non-congruent. The following section will present the stages of the extraction and selection of the target collocations in the present study, according to the statistical and phraseological approaches.

4.3.1.1 Statistical extraction of the collocations

The statistical extraction of the target collocations was carried out systematically as follows:

- As the current researcher was targeting adjective/ noun combinations, the node nouns were extracted from the most frequent 3,000 lemmas in the BNC (Leech, et al., 2001) which resulted in 1284 nouns.
- Collocates of each noun of the 1284 were then checked and extracted from the British National Corpus (BNC) according to two criteria.
 - Firstly, collocates should be adjectives that belong to the most frequent 3,000 lemmas in the BNC (Leech, et al., 2001) or to the General Service List (West, 1953).
 - Secondly, the node noun and the collocate adjective should have at least 50 occurrences (frequency threshold) in the BNC (within a window of ± 3) and an MI score of 3 or above. This step resulted in a very long list of adjective/ noun combinations.

Since the current researcher is also employing a phraseological approach to defining collocations, investigating the intuition of native speakers of English in producing these pairs was necessary (see chapter 2 section 2.7.2 for details and justifications). However, because the current research is only looking at non-congruent collocations, the long list had to be filtered before checking native speakers' intuition. A criterion to establish non-congruency from the point of view of native speakers of Arabic thus had to be established first. Moreover, checking the native speakers' sensitivity to every single item in the list is impractical and rather impossible due to the length of the list. A selection of a random sample of collocation would shorten the list, therefore, minimise the possibility of finding a good number of non-congruent

collocations. The following section briefly details an attempt to attaining non-congruency with the Arabic language.

4.3.1.2 Non-congruent English collocations with Arabic

Non-congruent collocations are broadly defined as collocations that do not have translational equivalents in L1 and thus are difficult to produce by L2 learners (Nesselhauf, 2003; Yamashita & Jiang, 2010, see also section 2.5). However, the presence or absence of an exact L1 translation equivalent is not sufficient, due to polysemy and prototypicality of meaning (Peters, 2015). Thus, congruency might not be as easy to operationalize as previously hypothesised. Swan (1997, p. 158) already referred to the role of prototypicality in translation equivalence as he states: “Languages may have exact translation equivalents when used in their central sense but not when they are used in more marginal or metaphorical ways.” Peters thus tentatively argued that the degree to which a collocation is presumed as congruent could differ from one learner to another.

This researcher would extend Peters tentative argument by suggesting that the notion of congruent vs. non-congruent collocations differs from one language to another and sometimes within the same language as in the case of Arabic. Selecting non-congruent (adjective + noun) collocations for the purpose of teaching in this research was a very demanding and challenging task for the following reasons:

- Firstly and most importantly, the richness in polysemy phenomena and the different varieties and forms of the Arabic language (i.e. the Classical Arabic of the Quran, Modern Standard Arabic³² or Colloquial Arabic³³) (Hasanuzzaman, 2013), and the necessity of making a decision on which form of Arabic.
- Secondly, the lack of a systematic framework, at the time of carrying out this task, to rely on when determining non-congruency of collocation, especially adjective/ noun pairs, in Arabic or in any language for that matter.
- Finally the non-existence of any lists of non-congruent adjective/ noun collocations from previous research.

³² Modern Standard Arabic is the language used in writing, reading and high register speech. It is derived from the Classical language of the Quran (Bishop, 1998).

³³ Colloquial Arabic is the language which is spoken regularly in all daily interactions and which Arabic speakers learn as their L1 (Bishop, 1998).

To overcome some of these issues and to generate a list of non-congruent adjective/ noun collocations, the current researcher decided to follow in Li and Schmitt's (2009) footsteps. In Li and Schmitt's study, collocations were judged by a panel of judges who identified English lexical phrases in the written assignments of an MA student and who tracked the participants' progress in the use of lexical phrases. However, in this study a panel of native speakers of Arabic was employed to judge the congruency of the English collocations with the Arabic ones, which had been extracted statistically in the first step. According to Moon (1997), "while not infallible, it is assumed that native judges can make a reasonable identification of the formulaic language [non-congruent collocations in this context] because those features have the property of "sounding right" and are "regularly considered by a language community as being a unit". Moreover, Bahns et al. (1986) pointed out that formulaic language with semantic-pragmatic functions can only be identified by native speakers' intuition.

The judges were required to be proficient in both English and Arabic in order to identify non-congruency and to make sure that the translations of the English collocations constitute units in Arabic as well. The judges not only had to be native speakers of Arabic, but they also had to have majored in Arabic/ English translation or Arabic language or to have experience in translation into/ from English. As for their English language proficiency, Newcastle University's entry level for non-native English speakers (ILETS 6.5) was considered acceptable.

A panel of two judges from similar backgrounds was initially set up. The first one was the current researcher as she has a BA degree in English/ Arabic translation from King Saud University (KSA) and experience in carrying out translation and interpretation work. She holds an MA degree in TESOL, and is currently a PhD candidate in Applied Linguistics. The second judge has a BA and MA in English and was also a PhD candidate in Applied Linguistics with at least 2 years' experience in English/ Arabic translation.

Similar to Li and Schmitt's research (2009), the second judge was given a brief description of the study and its aims by the first judge (the researcher). Additionally, he was presented with a short explanation of the common understanding of the notion of non-congruency in collocation (i.e. no word for word translation). Given the fact that the Arabic language constitutes more than one variety, and to overcome this issue as was mentioned earlier, this researcher, with the help from the second judge, decided to focus

on Modern Standard Arabic as well as Colloquial Arabic as used in the Gulf area. Unlike Li and Schmitt's five-point scale of lexical appropriateness, the current researcher presented a three-point scale as she was not interested in degrees of congruency (if such a notion exists at all). Therefore, the other judge was given the list of statistical collocations and was instructed to identify each collocation as either congruent, non-congruent or unsure, using his intuition. He was also instructed to provide an appropriate Arabic translation to what he believed were non-congruent collocations.

Judges in Foster (2001) and in Li and Schmitt (2003) reported that tiredness, lack of concentration and difficulty in marking lexical phrase boundaries led to missing obvious examples of lexical phrases. However, more confidence was gained by judges in Li and Schmitt's study after a certain amount of revision comprising reviewing the identification process, taking breaks during lengthier identification sessions etc. In this context a one-to-one revision of the statistical collocation list was administered to make up for any missing examples of non-congruent collocations. The judges compared notes on the selected collocations for their non-congruency. Only collocations identified by the two judges as non-congruent were added to a collocation list.

The list was quite interesting in the sense that some collocations such as *heavy losses* might have different translation versions according to Modern Standard Arabic *khasa'er fadiha* or to the colloquial Arabic *khasa'er kabera*, however, in both cases the English equivalent would be *big losses* thus showing non-congruency. In other cases, collocations like *deep trouble* and *vast numbers* are translated in Arabic as *big trouble/ problem* and *big/ huge numbers* which are fairly acceptable collocations. Thus, the non-congruency lies in using the exact combinations together (i.e. word-for-word). Collocations such as *naked eye* and *good faith*, more restricted combinations with marginal or idiomatic sense, were easier to identify as non-congruent since they are translated as *abstract eye* and *good sincerity*. It was interesting to find that there appeared to be a correlation between non-congruency of the adjective + noun collocations with their Arabic counterparts and the degree of restriction in collocation usage i.e. the more restricted the combination is, the easier it seems to identify it as non-congruent (see chapter 2 section 2.9.2 for details). However, this does not mean that free combinations might not constitute non-congruent collocations. It is also worth noting that although free non-congruent collocations may actually have congruent acceptable substitutes, they are still less likely or unlikely to be

produced by the EFL learners. Accordingly, those free non-congruent collocations might be underused despite them being strong collocates (with high MI scores).

The generated shorter list of only non-congruent collocations ($N= 75$) was then passed on to two more judges to agree or disagree with the opinions of the first two judges. The second pair of judges was similar to the first pair in terms of the following characteristics; (1) they both are PhD candidates in Applied Linguistics at Newcastle University, (2) they speak Arabic as their L1. One of the two judges has a BA majoring in Arabic and the other one has a BA degree in English. Both have teaching experience in their majors at a university level. They received the same background information regarding the study and the same instruction regarding the notion of non-congruent collocations in its primary sense. Just like the first panel of judges, they worked individually at first and then they compared notes. The second pair of judges agreed with the first pair on the non-congruency of the collocations except for two items (*sharp contrast* and *strong feelings*) which were accordingly eliminated. This step resulted in a 73-items list which can be fairly claimed to contain non-congruent collocations according to the statistical approach.

4.3.1.3 Phraseological status of the collocations

In order to check the phraseological status of the chosen collocations and to check English native speakers' sensitivity and intuition towards the pairs, a 73-item pilot test (clued recall) was developed. The test was administered to a group of eleven native speakers to test their knowledge of the 73 collocations.³⁴ Each item included the second word of the collocation (noun), the first letter of the first word (adjective), and a meaningful context (adapted from the BNC). Here is an example:

1. (**R----- years**) have witnessed changes in the overall structure of art education course.

Test takers were instructed to fill in the blank with the word that completes the phrase and that begins with the letter provided. In the latter example, for example, they are expected to come up with the word 'recent' to complete the collocation 'recent years'. Test takers were also requested not to make random guesses and to leave the item blank if they did not know the answer. In the end, 45 items were chosen where at least eight out of the eleven native speakers were able to recall the first word of

³⁴ The English native speakers were approached by the researcher by e-mail, and only 11 volunteered.

the collocation. Finally, 30 collocations were chosen³⁵ for teaching and they were randomly divided into three sets of ten collocations each (see the table below).

Table 4. 4: Sets of target non-congruent collocations

Set 1	Set 2	Set 3
Open air	Vast numbers	Early summer
Key areas	Broad agreement	Hard facts
Vast majority	Heavy emphasis	Heavy losses
Immediate future	Ill health	Low risk
Recent years	Naked eye	Instant coffee
Hard copy	Fine arts	Poor condition
Round trip	Steady progress	Heavy traffic
Domestic violence	Fresh start	Long tradition
Careful attention	Huge success	Safe return
Common sense	Careful planning	Good faith

4.3.2 Treatments worksheets

Since the three experimental groups were exposed to different treatments, this section is allocated to describing the worksheets used by learners in the three groups for learning collocations.

4.3.2.1 Worksheets for experimental group 1 (-DDL +CAT) and experimental group 2 (+DDL +CAT)

As indicated in chapter two (section 2.3), the distinction between passive and active knowledge of vocabulary may not be as simple as it seems, as there exists a great discrepancy in the use and interpretation of active and passive knowledge in the various studies (Read, 2000). In this research, the researcher follows the distinctions of Nation (2001) and Laufer et al. (2004) and refers to the ability to provide a word meaning as passive knowledge and to the ability to provide the word form as active knowledge. In that sense, the ability to supply the translation form of the target collocations in response to the learners' L1 translation equivalents is considered by the researcher as an active recall, and their ability to supply the meaning of the target words as passive recall (Laufer & Girsai, 2008a, b; Takala, 1984).

³⁵ According to their availability in the parallel United Nations (Arabic/ English) corpus.

The three sets of the target collocations (in Table 4.5) were included in six translation worksheets used by participants in experimental groups 1 and 2 as they were the groups with (+CAT) treatment. Three worksheets comprised English into Arabic translation tasks, and three included Arabic into English translations tasks. Each of the English /Arabic translation sheets included ten sentences that were adapted from the English/ Arabic parallel corpus i.e. some of the sentences were shortened or simplified. The participants were expected to translate the full sentence as they were believed to have an adequate lexical knowledge of K2 and K3.³⁶ The sentences were checked to be of matching word level. For each sentence, the Lextutor research tool was used to check the words' K-levels. If any of the words in a sentence was not at K1, K2 or maximum K3 level, it was substituted with a synonym that belongs to one of these levels. Here is an example:

- In **recent years** tourism has made an increasing impact on farming.

The Arabic sentences in the Arabic/ English translation sheets were translations of English sentences adopted from the same parallel corpus and comprised Arabic translations of the target English collocations. The translations of the target collocations were also bolded. It is worth mentioning that the translation worksheets in both experimental groups were identical. The following is an example:

- لا توجد أي سجلات عن بدايات حياة جيمي مكراي المهنية ولكن لابد من أنه قد أحرز تقدماً مطرداً.

4.3.2.2 Worksheets for experimental groups 3 (+DDL -CAT)

Since this experimental group was not intended to carry out contrastive analysis and translation tasks, different worksheets were designed for the participants in this group. Despite the fact that the participants in this group would not practice passive (E/A translation) and active recall (A/E translation) of the form and meaning of the target collocations, they were still subject to tasks aiming at practising passive and active knowledge of the target collocations. According to Waring (1997), another way of demonstrating and practising passive knowledge of L2 vocabulary is by asking the learners to choose the correct answer from several form options for a given meaning or to choose the correct answer from several meaning options for a given word. Whereas the E/A translation task is considered passive recall, the MC task is considered a passive knowledge task of form recognition.

³⁶ According to the VLT.

Active knowledge of vocabulary is associated with speaking and writing on the understanding that learners can retrieve the appropriate written or spoken word form for the meaning they want to express (Nation, 2001). On that basis, fill-in-blanks tasks were treated as active recall tasks. An additional rationale for using gap-filling questions to practice the learners' controlled productive/ active knowledge of collocations is that gap-filling questions, to a certain degree, resemble real-life communication situations where the learner needs to retrieve words or collocations in response to the given contextual clues (Laufer, 1998).

Each of the three sets of the target collocations was put into a MC worksheet and into a fill-in-blanks worksheet, resulting in six worksheets in total. The sentences in all the worksheets which included the target collocations were adapted from the E/A parallel corpus. They were also checked against Lextutor for the words' K-levels. Words that did not belong to the K1, K2 or K3 levels were substituted by simpler synonyms. The following is an example of an active recall (fill-in-blank) task:

- Milk was not greatly used by villagers, partly because of the..... **condition** of the animals.

Each item in the recognition MC task included four choices: the correct adjective and three plausible distracters (three adjectives either synonymous, contextually relevant or close in meaning). The collocability of the distracter adjectives with the node noun was set to be of very low MI scores ($MI < 1$) indicating very weak or non-collocates (for MI scores see appendix B). Here is an example:

- Homes in the **majority** of Detroit suburbs cost \$10,000–100,000.
a. greater b. big c. vast d. enormous

4.3.3 Designing the corpus data sheets

The use of computers and computer programs by learners might be essential to DDL although this is not always the case. DDL can also be used through printed materials instead of computer programs for the presentation of data to the learners. This can be more effective for those students who might be technophobic (Bernardini, 2002). Where the luxury of computer-equipped laboratories does not exist, printed materials would seem to be more economic and more accessible for the researchers (see chapter 3 section 3.4.1 for details).

In many DDL research contexts, “designing” may not seem to be the right word to describe the process of printing out data from corpora since almost all the well-known and established monolingual and bilingual corpora (e.g. BNC, COCA, ICA, CCA, UMIST, etc.) have their concordancers.³⁷ However, in this research context the task was not as easy as printing out concordance lines. One important reason for this is the limited existence of bilingual parallel English/ Arabic corpora and the relatively small sizes of the existing ones³⁸ (e.g. E-A Parallel Corpus, 2003, University of Kuwait, 3M words; Arabic English Parallel News, 2004, 2.5M words; Arabic Blog Parallel Text, 2008, 102K words, etc.). The content and the size of a corpus are closely interdependent aspects (Gavioli, 2000). This means that a small corpus may not guarantee adequate representation of general English (Gavioli, 2002), and inclusion of the target collocations. Almost all of the parallel corpora were behind a paywall except for the E-A Parallel Corpus which is accessible only by staff and students of Kuwait University through the university’s server. Hence, the current researcher opted for the best available and freely accessible parallel corpus, namely the English-Arabic Parallel Corpus of United Nations Texts (EAPCOUNT).

EAPCOUNT is one of the largest available parallel corpora containing the Arabic language. It was intended as a general research tool, and started in 2006 as a PhD research project at the University of Carthage, by Dr. Hammouda Salhi. It was completed and revised in 2010 as a result of collaborative work between Dr. Salhi and some of his students. It was motivated by the increasing demands for cross-lingual research and information retrieval (Salhi, 2010). The EAPCOUNT comprises 341 texts aligned on a paragraph basis, so texts in English are shown along with their translational counterparts in Arabic. It consists of two sub-corpora; one contains the English originals and the other their Arabic translations. The English sub-corpus contains 3,794,677 word tokens. The Arabic sub-corpus has slightly fewer word tokens (3,755,741). This means that the whole corpus contains 7,550,418 tokens.

³⁷“A concordancer is a programme that searches a corpus for a selected word or phrase and presents every instance of that word or phrase in the centre of the computer screen, with the words that come before and after it to the left and right” (Hunston, 2002. P. 39).

³⁸ Compared to some of the English monolingual corpora.

The existence of the EAPCOUNT in a 'raw'³⁹ form constituted another problem for the current researcher. In order to be able to carry out the task of searching and sorting the target collocations in this research, a concordancing programme was needed. According to Talai and Fotovatnia (2012), language teachers can utilise a concordancing technique for presenting DDL exercises to the learners. Several concordancing programs are commercially available, such as WordSmith, MonoConc and ParaConc. Some others are free such as Wconcord and ConcApp. However, although these tools work perfectly well on English and other languages with Roman script, they are not very effective tools for processing Arabic (Alsulaiti, 2004). Thus, the researcher utilised two different tools; one to process the English texts (WordSmith) and the other to process the Arabic texts (Examine32 Text Search tool). It is worth noting, though, that a few concordancing tools are now available for searching and analysing Arabic corpora such as AntConc 3.3.5, KACST Arabic Processing Tool and ConCorde. However, these tools were still in development at the time of this study, so it was not possible to use them.

As the target collocations were English, the researcher started by processing the English corpus texts to generate concordance lines in a KWIC format. A KWIC format denotes that several sentence examples with the target word are generated. The lines may comprise incomplete sentences and are organised one below the other for the purpose of centralizing the intended word or grammatical point in the middle of each line. Through using this technique, the attention of the learners is attracted to the intended word or lexical item and its immediate context in different sentences.

To carry out this task, all the English text files in the corpus were uploaded into WordSmith 6.0. The researcher then began her search using the node word of each of the target collocations (the noun) to search for collocates in the corpus. The researcher then copied the first fifteen concordance lines along with their file numbers.⁴⁰ Two crucial matters should be noted here: (1) to enhance the chances of the learners noticing the target collocation, five occurrences of the intended collocation in the concordance lines were targeted,⁴¹ (2) if the five occurrences did not appear in the first fifteen concordance lines

³⁹ A raw or unannotated corpus consists mainly of the text itself without additional information (McEnery & Wilson, 2001).

⁴⁰ The file numbers were needed to help the researcher find the Arabic counterparts.

⁴¹ In an ideal situation in which learners have access to the corpus, they can encounter the target collocations more than once. Repeated exposure of the lexical units results in the strengthening of connections (gradual reinforcement of the association) (Cleeremans, et al., 1998; Ellis, 2003; Schmidt, 1993b, 1994; Williams, 2009).

(which is unlikely), the researcher inserted a concordance line which comprised the target collocation. Eventually, fifteen concordance lines for each of the 30 node nouns were extracted with ten non-target collocations and five target collocations of the node word. Using the file numbers of the concordance lines, a file was created for each of the Arabic counterpart texts. The following screen shot shows an example of the concordance lines generated using WordSmith 6.0 in search for collocates of the word ‘violence’.

N	Concordance	Set	Tag	Word #	Sen	Sen	Parx	Parx	Hea	Hea	Sec	Sec	File
1	integrated part of assistance to all victims of violence and trauma and persons with			7,870	200	76%	0	43%			0	43%	APLC_MSP2_2000_
2	health, community development, conflict and violence prevention. In addition, the SCE-VA			7,940	202	95%	0	43%			0	43%	APLC_MSP2_2000_
3	integrated part of assistance to all victims of violence and trauma and persons with			449	13	76%	0	18%			0	18%	APLC_MSP2_2000_
4	health, community development, conflict and violence prevention. In addition, the SCE-VA			519	15	95%	0	21%			0	21%	APLC_MSP2_2000_
5	safe while others are wracked by organized violence. The village has suffered a growing			4,350	203	100	0	17%			0	17%	A_54_2000_en.snt
6	can live free from hunger and safe from violence, drinking clean water, breathing clean			4,493	208	67%	0	18%			0	18%	A_54_2000_en.snt
7	of communities and individuals from internal violence. The need for a more human-centred			13,228	637	100	0	52%			0	52%	A_54_2000_en.snt
8	Women have become especially vulnerable to violence and sexual exploitation, while children			14,058	678	31%	0	56%			0	56%	A_54_2000_en.snt
9	are sometimes the principal perpetrators of violence against the very citizens that			14,150	682	48%	0	56%			0	56%	A_54_2000_en.snt
10	hunger and squalor and from the fear of violence or oppression. These rights are best			23,766	1,11	93%	0	94%			0	94%	A_54_2000_en.snt
11	us to do " and especially from the violence of civil conflict and the fear of			24,616	1,11	59%	0	97%			0	97%	A_54_2000_en.snt
12	, ignorance and disease; the fight against violence and terror; and the fight against the			25,013	1,14	75%	0	99%			0	99%	A_54_2000_en.snt
13	, exclusion, environmental degradation and violence. Regarding human rights and gender			829	22	100	0	11%			0	11%	A_54_718_en.snt
14	noted that armed conflict and other forms of violence impacted more severely on children,			2,808	89	36%	0	82%			0	82%	A_54_754_en.snt
15	Administration's capacity to combat crime and violence in East Timor. The refurbishment and			1,525	54	86%	0	7%			0	7%	A_54_769_en.snt
16	of the encounter and with threat of lethal violence, including the use of firearms,			459	12	48%	0	61%			0	61%	A_54_773_en.snt
17	decision-making process, multiple issues of violence against women, growing post-conflict			5,714	176	55%	0	31%			0	31%	A_54_807_en.snt
18	women, growing post-conflict domestic violence, trafficking of girls and young women			5,720	176	66%	0	32%			0	32%	A_54_807_en.snt
19	the growing number of attacks and acts of violence against United Nations and associated			8,039	239	31%	0	62%			0	62%	A_54_839_en.snt
20	and property were subjected to acts of violence, and staff security continues to be			5,395	165	69%	0	65%			0	65%	A_54_841_en.snt
21	the increasing number of inhuman acts of violence and atrocities affecting primarily the			205	2	85%	0	57%			0	57%	A_54_888_en.snt
22	urges all armed groups to end all acts of violence, to respect human rights, to comply			287	4	41%	0	80%			0	80%	A_54_888_en.snt
23	insurgents to call an immediate halt to their violence and on the Nepalese Government to			438	11	49%	0	72%			0	72%	A_54_895_en.snt
24	encounter daily the human impact of rising violence and armed conflicts, widespread			485	11	49%	0	5%			0	5%	A_54_959_en.snt
25	in increasing inequity, social injustice and violence worldwide. The eradication of poverty			986	33	95%	0	11%			0	11%	A_54_959_en.snt
26	and preserving human life. Organized armed violence is depriving millions of people all over			2,228	80	12%	0	24%			0	24%	A_54_959_en.snt

Figure 4. 3: The concordance of ‘violence’ in WordSmith concordance tool

After finishing with the English concordance lines, the researcher started her research for the Arabic counterparts in the E/A parallel corpus. As mentioned earlier, existing concordancers do not support the processing of Arabic texts, so the researcher utilised software called Examine32 Text Search 6.00. This software enables its user to conduct two different types of searching, one of which is the text search, which implies that the user enters the desired word or phrase to search for. The user needs to select the specific folder that contains the files he/ she is looking for. Based on the specified word or phrase, the software can scan all the files contained in the indicated folder, even sub-folders, then return the results. To search for the Arabic counterparts of the English concordance lines, the researcher searched each of the saved files mentioned above and manually extracted the lines. This was done by cutting the Arabic sentences from the beginning to the end of the concordance line. The researcher paid careful attention to the process of producing correctly matching English and Arabic texts. The following screen shot shows the search for the word ‘violence’ in the Arabic texts.

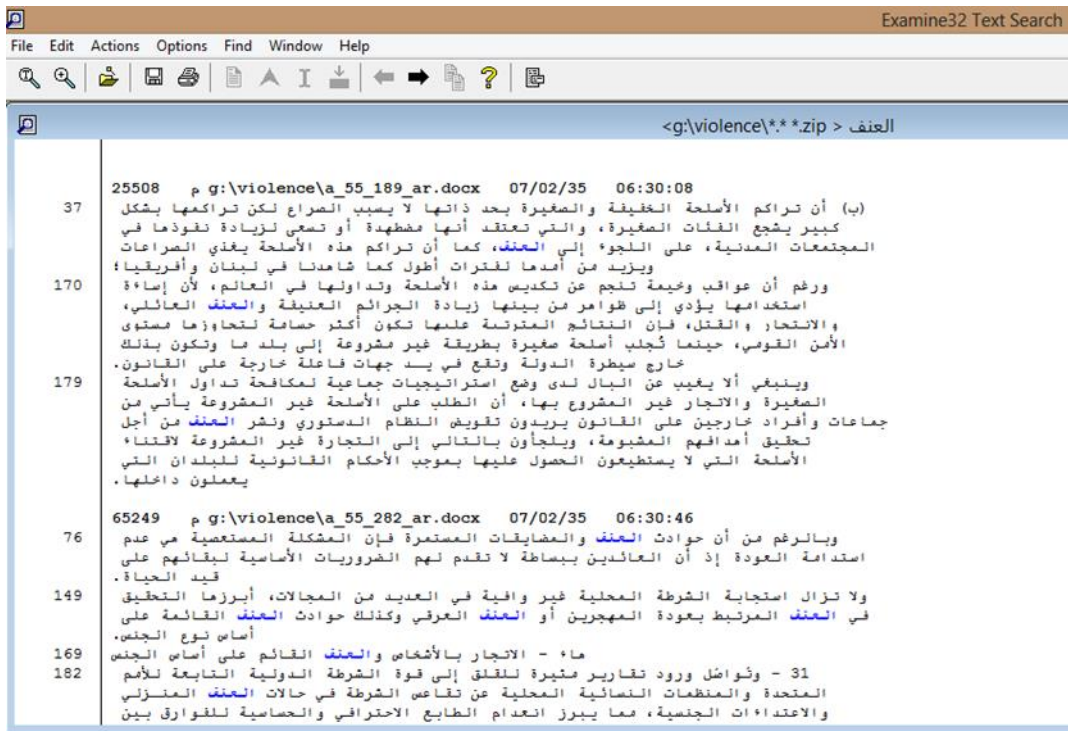


Figure 4. 4: The concordance of ‘violence’ in Examine32 Text Search tool

Upon compiling the English concordance lines and their Arabic counterparts, the researcher needed to present them in an adequate parallel manner. Thus, she adopted the layout of Kuwait University English/Arabic Parallel Corpus (Al-Ajmi, 2003) which places the concordance lines vertically paralleled as shown below

It is worth mentioning that the monolingual sheets comprised the same concordance lines but with only the English part included. It was presented and arranged in a (KWIC) format (see appendices G and H for samples of the bilingual and monolingual corpus-data).

نصوي نتائج البحث على 4			
الموضوع :	كل الموضوعات	الكتاب :	الذكاء العاطفي
النص الحر :	كصير		
النص العربي	English Text		
تضرب مثلا بالتجربة التي شاهد فيها المتطوعون لقطات من فيلم كصير المزيد ...	For example in one experiment volunteers watched short film clips More...		
والتنس كصير المزيد ...	shortness of breath More...		
على الأكل على المدى كصير المزيد ...	at least in the short term More...		
إن إن الشعور الجميل على المدى كصير يقابله انهيار حياتهم المزيد ...	a short-term good feeling in exchange for the steady melt-down of one's life More...		

Figure 4. 5: An example of Kuwait University E/A Parallel Corpus layout (adopted from Alsulaiti, 2004)

4.4 Procedure (experimental groups)

The intervention for all experimental groups lasted for six weeks with 55-60 minutes of a three-hour class per week for each group. Prior to the intervention and in order to familiarise the students with DDL as a new concept and approach, a 45-minute session with an introduction to the notion of corpora, specifically bilingual corpora, their format, and their usage for language learning was given to experimental group 2 (+DDL +CAT). The session was also intended to familiarise the learners with the idea of using bilingual corpus-data to compare and contrast their mother tongue with English and to come to an understanding of the similarities and differences between the two languages in terms of individual words and the overall lexical system. This pre-treatment stage aimed at distinguishing this contrastive FFI from bilingual glosses which simply state the meaning of L2 words.

A similar session was conducted with experimental group 3 (+DDL -CAT) though it involved information on basically monolingual corpora, their formats and their usage for language learning. Handouts which included a summary of the sessions were distributed to the students in both groups. As no corpora had been involved in their teaching, students in experimental group 1 (-DDL +CAT) were not subject to any introductory sessions on corpora. Similarly, students in the control group did not receive any information since they were not subject to any treatment or condition.

The teaching sessions for all experimental groups were divided into two parts. In the first part, all experimental groups were given a reading passage along with worksheets which included three MC questions to assess general comprehension of the reading texts. The comprehension questions were written in such a way that none were related to the target collocations i.e. no knowledge of or reference to the target collocation was required in order to answer the questions. The students were given approximately 20 minutes to carry out the reading and MC tasks. By allocating a specific time for the completion of the tasks in the three groups under the different conditions, this researcher was hoping to exert some control over the time-on-task factor which may affect the learning outcome.

The texts were chosen from the New Headway plus Intermediate, Special Edition (Liz and John Soars, 2012). This book is used in the foundation year for students who are not majoring in English, but was chosen to ensure that the level of the reading passages matched the students' expected proficiency level (intermediate). The passages in Chapter 1: "Wonders of the modern world", Chapter 2: "The life of a hard working king", Chapter 3: "Agatha Christie", Chapter 8: "Giving your money away", Chapter 10: "The beautiful game" and Chapter 11: "How well you know your world" were the most suitable passages as they allowed the inclusion of the target collocations. Each passage was shortened slightly (with a maximum length of 561 words), and was adapted to include one occurrence of each target collocation twice (see figure 4.6 below for illustration and appendix E for samples of the reading passages).

Ch1: Wonders of the modern world → Collocation set 1 ← Ch8: Giving you money away
Ch2: The life of a hard working king → Collocation set 2 ← Ch10: The beautiful game
Ch3: Agatha Christie → Collocations set 3 ← Ch11: How well you know...

Figure 4. 6: Collocation sets occurrences in reading passages

4.4.1 Experimental group 1 (-DDL +CAT)

The treatment procedure in this group was similar, but not identical to that in Laufer and Girsai's studies (2008a, b). In the teaching sessions for this group, the participants were initially given reading passages as stated previously, and were

instructed to read the passages silently for (10-15 minutes). After they had finished the reading, they were asked to answer MC comprehension questions (5 minutes). Upon completion of the task, the researcher went over the answers with the learners. That being done, the translation tasks followed.

In the first three sessions, the students were requested to translate ten English sentences into Arabic (passive recall) and pay attention to the translation of the bolded word combinations (i.e. the target collocations). In these sessions, the reading passages were not collected and the learners could use them for more clues about the meaning of the collocations if they wanted to. The target collocations were bolded in each sentence. The researcher monitored and provided help when needed. After the students reported finishing the translations, the researcher gave corrective feedback as well as explicit contrastive instructions. For example, the researcher pointed out that while in most of the cases the nouns have equivalents in Arabic, the adjectives that collocated with them were totally different (e.g. *heavy emphasis* in English can be *extreme emphasis* in Arabic). She suggested that students should be careful not to provide automatic transliterations which might lead to the production of weak or unacceptable word combinations.

Earlier studies on vocabulary acquisition have shown that productive learning of word pairs can be more effective than receptive learning of word pairs at increasing productive knowledge of meaning, and the receptive task is more effective than the productive task at contributing to receptive knowledge of meaning. However, results from other studies (e.g. Webb & Kagimoto, 2009) indicate that both receptive and productive tasks were effective in learning collocation and meaning, and that there was little difference between the effects of the two types of tasks. In addition, this researcher could not assume that all of the target collocations were already part of the participants' passive/ receptive knowledge. Therefore, the first three teaching sessions were intended to focus on the learners' receptive knowledge of the target collocations and to raise their receptive awareness of the cross-linguistic differences.

The same procedures were followed in the next three teaching sessions. However, the learners were requested to translate the sentences from Arabic into English (active recall). Moreover, in these sessions the passages were taken away so that the students could not copy the collocations from them. Instead, whenever the new words were

deemed necessary for a task by learners, they could ask the teacher (i.e. researcher) for help. The answers to students' questions and explanations were given in English. The learners received the same kind of corrective feedback and contrastive analysis upon finishing the translation tasks. These sessions aimed to establish their active/productive knowledge of the target collocation. Moreover, it is widely acknowledged in the empirical studies that learners cannot be expected to learn a word fully on first exposure (Schmitt & Schmitt, 1995; Schmitt, 2008). In fact, negligence in the recycling process will result in many partially-known words being forgotten, wasting all the effort already put into learning them (Nation 1990). The three final sessions, therefore, addressed the recycling issue.

4.4.2 Experimental group 2 (+DDL +CAT)

The first part of the treatment procedure of experimental group 2 was identical to experimental group 1 and 3. In the second part however, the participants were given collocation learning worksheets along with sheets that included the concordance lines from a bilingual English/ Arabic corpus. In each session for the first three teaching sessions, the students were instructed to translate ten English sentences into Arabic with the help of the corpus data. They were requested to pay careful attention to the translation of the bolded word combinations (collocations) in each sentence.

In order to translate the collocations in particular, the students were asked to consult the corpus data sheets and search for the combinations, observe the Arabic meanings of the individual words comprising the collocation as well as the holistic meaning of the combination. For example in the sentence *In **recent years** tourism has made an increasing impact on farming*, the students were expected to notice the following when translating the collocations:

- 1) Fi al-sanawat al-akhirah
In det-years det-last
In recent years

As can be seen in the previous example, the word *recent* does not have an exact equivalent in Arabic, since the Arabic translation of it does not imply or convey the meaning of being recent. It could be rather associated with *last* and translated as the *last few years* instead of the two-word combination *recent years*.

In the next three teaching sessions, the learners were exposed to the same three sets of collocations. However, in these sessions they were requested to translate different sentences from Arabic into English. Upon the completion of the translation task in each teaching session, the teacher (the researcher) went over the translations with the class. The corrective feedback was on the general translation of the sentence i.e. no attempt was made to further explain the meaning of the collocations or give any contrastive analysis instruction.

4.4.3 Experimental group 3 (+DDL -CAT)

Similar to experimental groups 1 and 2, the first part of the treatment procedure was the reading and MC comprehension tasks. The second part comprised collocation learning worksheets and sheets that included concordance lines from the parallel corpora. However, unlike those given to the other group, these concordance lines were monolingual i.e. only the English part of the same concordance lines was included. In each of the first teaching sessions, the participants were asked to carry out a MC task in which they were supposed to choose the most suitable adjective that goes with each noun. They were instructed to consult the corpus data to help them understand, decide their answers or check their decisions. The researcher monitored while the students carried out the task.

In the next three teaching sessions, the students were asked to use corpus data to fill in the blanks with the missing adjective that most appropriately goes with the noun. They were given the same instructions regarding corpus consultation as in the previous sessions. At the end of each teaching session and upon completion of the multiple-choice and fill-in-the-blanks task, the researcher went over the items with the participants.

4.5 Collecting data on collocational knowledge: measures

Word knowledge entails many components of knowledge: the word's spelling, pronunciation, meaning, syntax, morphology, lexical relations, etc. (Nation, 2001). Moreover, knowledge of vocabulary falls on a receptive/ passive- productive/ active continuum, rather than existing as an all-or-nothing dichotomy (see chapter 2, section 2.3). Collocational knowledge, being one aspect of lexical knowledge, also operates along a continuum. However, this research draws on Laufer et al.'s. (2004) emphasis

that the most important component of word knowledge is the knowledge of the form/meaning relation, that is, the ability to retrieve the meaning of a given word form, and the ability to retrieve the word form of a given concept (as indicated earlier in this chapter). Emphasis on form/meaning relation was addressed in the collocation teaching sessions. Therefore, the learning product was measured with the recall of meaning as a passive/receptive knowledge test (E/A translation) and the recall of form (A/E translation) as an active/productive knowledge test.

To examine the changes in the learners' collocational knowledge brought about by the three teaching conditions, measurements were taken at three points in time: two weeks prior to the intervention, immediately subsequent to the intervention and three weeks after the treatment period. A rationale for the length of delay between the post-tests was provided by Schmitt (2010) who affirms that a delayed post-test of three weeks indicates stable and durable learning. The total duration of each of the tests was 90 minutes approximately. Note that the measurements were taken from the four groups, but only the experimental groups had received the different collocation treatments. The items in the pre, post and delayed post-tests were exactly the same, however, the sequencing of items was different to avoid a washback effect.⁴² Moreover, the set of items used in the collocation learning worksheets distributed in the teaching sessions were different from the items in the tests.

The next section gives a procedural account of how the collocation tests were developed.

4.5.1 Pre, post and delayed post-tests of passive collocational knowledge

This test included 30 English sentences which comprised the target collocations. These sentences were extracted from EAPCOUNT⁴³ through the WordSmith tool. Firstly, the Concord tool in WordSmith was used to generate concordance lines of each of the target collocations. Since the concordance lines were of incomplete sentences, the researcher accessed the full context of each of the collocations to extract meaningful sentences. The sentences were then shortened and simplified when necessary by substituting words that do not belong to the K1, K2 or K3 world levels with simpler synonyms. It is worth noting that unlike the English/Arabic

⁴² Washback effect refers to the effect that tests have on teaching and learning (Shohamy, 1993).

⁴³ The English-Arabic Parallel Corpus of United Nations Texts.

translation tasks in the treatment sessions, the participants were not asked to translate full sentences. They were instructed to translate the underlined word combinations i.e. collocations only. This is mainly because of the constraints of the class time and due to the fact that the participants might not have been able to finish the translation of thirty sentences during the test time allocated for this part (30 min.).

4.5.2 Pre, post and delayed post-tests of active collocational knowledge

The active recall test included 30 Arabic sentences with one target collocation in each sentence. The sentences were the Arabic counterpart translations of the original English sentences. Following the same process as in section (4.4.3), this researcher extracted 30 English sentences in which the target collocations occurred. They were different from the ones in the passive recall test. After that, she used Examine32 Text Search 6.00 to find the counterpart Arabic translations. Similar to the passive recall test, the participants were asked to translate only the bolded Arabic word combinations into English. The time allocated to finishing this part of the test was 30 minutes. To minimise the possibility of the collocations in the active recall test being remembered in the passive recall test, the participants were given a 15-20 minute distracting task (10 addition math problems), followed by a brief 5-10 minute discussion about a general topic. Additionally, the order of the target collocation in the passive recall test was different from the active recall test.

4.6 Marking the tests and analysing the data

The previous section has outlined the instruments for eliciting learners' collocational knowledge prior to and after the experimental treatment. This section goes on to detail the methods of marking and analysing collocation tests.

4.6.1 Marking (English ↔ Arabic translation tests)

The translation tests were manually marked by the current researcher. Marking the English/ Arabic translations of the collocations was rather straightforward. The Arabic translation was considered correct by the assessor if the participant was able to understand the meaning of the English collocation and produced an acceptable⁴⁴ translation in the modern Arabic language. If the Arabic translation did not show any understanding or an incorrect understanding of the English collocations, then the

⁴⁴ According to the intuition of the assessor as a native speaker of Arabic and her strong knowledge of Modern Standard Arabic.

answer was considered wrong. Accordingly, one point was given to each of the acceptable answers, while no point was given for incorrect answers. For example, in translating the collocation ‘heavy losses’ into Arabic, the adjectives *fadeha*, *kabira* and ‘*haa’ela* were considered acceptable, while the adjective *thaqeela* was considered to be a transliteration of the English collocation which does not indicate understanding of meaning.

In comparison, marking the Arabic/ English tests was less straightforward. This is due to the fact that some of the target collocations (four of the target collocations) were not highly restricted combinations in Arabic which allowed for a relatively wider range of possible answers. For example, the Arabic collocation *A’adad haila* can be translated into different acceptable English collocations other than the desired response *vast numbers*. Some of the produced collocations which could be accepted collocates⁴⁵ are *huge numbers* and *large numbers*, but not *big numbers*, which would be the exact meaning of the collocation in Arabic. It is of paramount importance to note that although these acceptable collocations could be treated as indicators of collocational knowledge to a certain degree, students were not given any points for them, because they did not serve the purpose of this research. The idea behind the collocational treatments in this study was to expand the learners’ vocabulary size by presenting new word combinations and to help them to establish or strengthen the links between the non-congruent combinations. Hence, only the intended target collocations were considered as correct answers in the post-tests, and one point was given for each correct answer. The same criteria were used in marking the pre-tests since the learners were not able to produce the target non-congruent collocations even though they were given the chance to produce more than one translation if they desired.

To ensure reliability of the marking process, this researcher re-checked the marking of all passive and active recall tests from each group after an interval of one month. Recruiting a second assessor for that purpose was an idea taken into consideration, however, it was not feasible due to the abundance of test-papers. The same marking criteria were employed in the second marking stage.

⁴⁵ According to the BBI Dictionary of English Word Combinations, Oxford Collocations Dictionary for Students of English or the BNC.

4.6.2. Analysing the data

The pre, post and delayed post-tests were administered to the experimental groups and the control group, so twelve sets of test scores (i.e. three tests x four groups) were gathered and subjected to statistical analyses, using SPSS 21 (Statistical Package for the Social Sciences) for within-group and between-group comparisons. Descriptive statistics, including the mean or median, and standard deviation (SD), were calculated to examine the participants' performance on the pre, post and delayed post-tests. Descriptive statistics were used to "characterize or describe a set of numbers in terms of central tendency and to show how the numbers disperse, or vary, around the centre" (Brown & Rodgers, 2002 p. 122). However, these cannot be used to make inferences about or assess the strength of the relationship between the independent (causal) variables, and dependent (effect) variables. Hence, inferential statistics were necessary to make comparisons within and between groups. Different parametric and non-parametric statistical tests were used to assess the differences in the learners' performances within each group and between groups. Detailed justifications for each statistical test used for the comparisons are provided along with the results in the next chapter (Chapter 5).

4.7 Validity and reliability of the research

Validity is considered to be the methodological goal of the researcher. Reliability, on the other hand, is an essential element in the attainment of validity. However, the relationship between validity and reliability is believed to be unidirectional. As Fred (2011) put it, reliability does not require validity, but validity depends on reliability. This means that an instrument can be reliable when measuring something, while not measuring the right thing. On the other hand, an instrument cannot be judged valid if it is not reliable, i.e. accuracy entails consistency, not vice versa (Fred, 2011). The following sections introduce the notions of reliability and validity, and outlines the reliability and validity issues for this research.

4.7.1 Reliability

Reliability is commonly used in relation to the question of whether or not the measures devised for a given research design are consistent (Dörnyei, 2007). Consistency of results is either defined by measuring rater reliability, instrument reliability, or both whenever applicable (Mackey & Gass, 2005).

The most obvious ways of determining instrument reliability is the test-retest and equivalence of forms of a test as in pre and post-tests (Mackey & Gass, 2005; Fred, 2011). The test-retest method entails administering a test or measure on one occasion and then re-administering it to the same sample on another occasion. Results obtained from the two tests should show little variation over time (Bryman, 2012). Equivalence of test forms, i.e. testing using the same materials and instruments, is believed to be of great importance. This is due to the obvious fact that it would be inappropriate to have one version of a test be easier than the other as the resulting gains of the treatment would be falsely high or falsely low. For this research, the test-retest method of measuring instrument reliability was not feasible due to practical issues such as availability of participants, which obstructed the procedure. In fact, Mackey and Gass (2005) affirm that it is not always possible to administer tests twice to the same group of participants. Nonetheless, the equivalence of forms reliability was ensured since this research did not utilise different sets of test items at different times of testing. The same exact set of items were used in the pre, post and delayed post-tests, thus fulfilling the equivalence of test forms requirement. The only difference between the tests was the order of the items, which was changed in each test in order to reduce the washback effect.

As mentioned earlier, another type of reliability is referred to as rater reliability. The defining feature of rater reliability is that scores by two or more raters (i.e. inter-rater) or one rater at Time X and that same rater at Time Y (i.e. intra-rater) are consistent (Mackey & Gass, 2005; Cohen, Manion & Morrison, 2007; Fred, 2011). The latter type of consistency has been taken into consideration in the process of marking the translation tests (see section 4.7.1). To recapitulate, the researcher used her knowledge of Modern Standard Arabic as well as her mother-tongue intuition as a criterion for marking the English/ Arabic translation tests. As for marking the Arabic/ English translation tests, only the produced target collocations were marked as correct, and each correct answer was given one point. Approximately a month later, the researcher re-checked the marking of all the tests of passive and active collocational knowledge for all groups. According to Gwet (2014), intra-rater reliability can be measured using the Intraclass Correlation Coefficient (ICC), which is the preferred measure for continuous or scale data. ICC was run and yielded an alpha coefficient of .938 (see tables below), which suggests that the level of agreement between the scores in the two marking periods was very high (Larson-Hall, 2010).

Thus, consistency of scoring and reliability of results can be claimed. Note that in the cases where the scores were different between the two marking stages, an average score was used for the analysis.

Table 4. 5: intra-rater reliability

Cronbach's Alpha	N of Items
.938	12

Table 4. 6: Intraclass Correlation Coefficient

	Intraclass Correlation ^b	95% Confidence Interval		F Test with True Value 0			
		Lower Bound	Upper Bound	Value	df1	df2	Sig
Single Measures	.559 ^a	.494	.628	16.202	128	1408	.000
Average Measures	.938 ^c	.921	.953	16.202	128	1408	.000

4.7.2 Validity

Validity is primarily concerned with the integrity of the conclusions that are generated from a piece of research (Bryman, 2012). However, the notion and quality of validity is more complex than it appears. According to Messick (1995, p. 741) validity is not a test's property, rather it is "an overall judgment of the extent to which empirical evidence and theory support the adequacy and appropriateness of the interpretations based on the assessment". Logical distinctions also exist between empirical evidence for measurement validation i.e. its evidential basis as well as its consequential basis or functional impacts on social systems and values that result from the assessment (Messick, 1989). To that end, many types of validity are distinguished in research methodology textbooks, including construct validity, content validity,⁴⁶ predictive validity, face validity, internal validity, external validity, etc. Being the most common areas of concern in quantitative research (Macky & Gass, 2005), internal and external types of validity are discussed in relation to this research. Rather than having a set of mini-validities, this section,

⁴⁶ Messick (1994) considers content validity as one aspect under the broader notion of construct validity, along with other aspects such as substantive, structural, generalizability, external and consequential aspects.

following the guidelines of Fred (2011), discusses different facets of a more global construct validity.

4.7.2.1 Multiple facets of validity

A more global notion of validity involves two main facets: trait accuracy and trait utility (Fred, 2011). Trait accuracy corresponds with the well-established meaning of construct/ measurement validity (ibid) concerning the question of whether or not the measurement accurately measures and reflects the concept it was designed to measure (Bryman, 2012). The degree to which a procedure is valid for trait accuracy is determined by the degree to which the procedure corresponds to the definition of the trait (Fred, 2011). Trait utility, on the other hand, is concerned with whether measurements are utilised to measure the intended trait (ibid).

In this research, acquisition of lexical collocations is defined as the ability of the participants to actively and passively recall the target collocations. Accordingly, translation tests that measure the active and passive recall of the target collocations are believed to be valid measures. The study has thus attained trait accuracy and utility i.e. construct validity. No matter how simple and straightforward this may appear, both facets (i.e. trait accuracy and utility) are defined by other components i.e. content coverage (parallel to content validity) and face appearance (parallel to face validity).

In experimental research, the main problem is teasing out a cause and effect relationship to establish the effects of treatment. Typically, the treatment's objective is to enhance learning or change the attitude or behaviour of the participants. This exact objective needs to be considered when planning the measurement tool because its main goal is to assess the achievement of the treatment objective (Fred, 2011). In addition, Mertens (1998, p. 294) states that "If all students are taking the same test but all the students were not exposed to the same information, the test is not equally content valid for all the groups." Consequently, the validity of the measurement procedure is not evaluated by computing a correlation coefficient, but by aligning different components of the measurement procedure with the treatment objectives (Fred, 2011). Regarding content coverage of the measurement procedure in this study, the following points are worth mentioning:

- The participants in both experimental group 1 (+DDL +CAT) and experimental group 2 (-DDL +CAT) have practised active and passive knowledge of the target items in the form of translation tasks (E/A and A/E). Hence, the current research can safely claim the validity of the content coverage of the measurement tool for these groups.
- The participants in experiment group 3 (+DDL –CAT) were purposefully not given practice translation tasks in their treatment. This is because the current researcher intended to not only assess the effect of DDL, but also the effect of presence or absence of contrastive analysis and translation tasks. Nevertheless, the validity of the results according to content coverage should be accepted for several reasons: (1) prior to the treatment phase, the participants in this group were exposed to the target collocations and the translation tasks in the pre-test; (2) in the treatment phase, the participants were subject to monolingual tasks and focussed on practicing active and passive knowledge of the target collocation; (3) although the tasks were monolingual, the current researcher relied on the argument of the Revised Hierarchal Model (RHI) of the bilingual lexicon which states that “During early stages of SLA, words in the L2 are hypothesized to be associated to their translation equivalents. Because words in the L1 are assumed to have direct access to their respective meanings, the activation of the translation equivalent in L1 facilitates access to meaning for the new L2 words” (Sunderman & Kroll, 2006, and see chapter 3 section 3.4.2.2 for details). This argument was also supported by Laufer and Girsai (2008a, 2008b).

In relation to accuracy, face appearance is concerned with whether a measurement procedure appears to the public eye to measure what it is supposed to measure (Fred, 2011). Face validity is closely related to content validity in that it aims to convince others that the designed measurements have content validity (Mackey & Gass, 2005). In regards to utility, face appearance is important for many people such as examinees and people outside a study. To illustrate, people outside a study may not see the relevance of a certain measurement tool and, consequently, not consider the results from such measurement suitable for answering the researcher’s question (Fred, 2011). According to Bryman (2012), face validity can be established by asking people with expertise in a particular field to check whether or not the measure appears to be representative of the trait it is designed to measure.

With regard to this research, the translation tests were shown to academic staff members (supervisors), as well as to a number of PhD students in the school of ECLS at Newcastle University who had experience in the field of second/ foreign language teaching and learning. After reading the research hypotheses and checking the content and instruction language of the tests, they agreed that the instruments appeared to be valid in relation to the research's main and sub-hypotheses.

3.7.2.2 *Internal validity*

One main type of validity is internal validity, which is concerned with the question of whether a conclusion that involves a causal relationship between two or more variables holds water (Bryman, 2012). In other words, it refers to the extent to which the differences that have been found for the dependent variable are directly related to the independent variable (Mackey & Gass, 2005). Internal validity is of critical importance in any research involving a cause and effect relationship (Fred, 2011). A researcher must control for all the potential factors that could possibly account for the results and eliminate or at least minimise threats to internal validity (Mackey & Gass, 2005).

For this research, several attempts were made to control for extraneous variables and essential variables that may affect the results. For example, the participants' English proficiency and vocabulary levels were controlled in all groups (experimental and control), so that no variation in the research results could be attributed to the variations in the proficiency or vocabulary levels between them. Moreover, the prior collocational knowledge of the participants was controlled to verify the causal inference within and between the groups.

Participants' mortality (i.e. attrition), as one way of compromising internal validity, was also taken into considerations in this research. According to Mackey and Gass (2005), some studies in second language research seek to measure language development over time, so they typically carry out immediate post-tests as well as one or more delayed post-tests to identify the longer or shorter effects of treatments. They assert that in order to appropriately address research questions and hypotheses, it is best to make sure that all participants are present for all sessions. Hence, only

the results of the participants who attended all the treatment sessions were considered in testing the research hypotheses.

One serious design issue constituting a threat to the internal validity of research relates to the comparability of tests (Mackey & Gass, 2005). One way to ensure comparability is to establish a fixed group of sentences in all tests. In this research, comparable vocabulary difficulty levels between the sentences in the treatment sessions and in the tests was maintained. This was attained by consulting a word frequency index (i.e. Luxtutor) for each sentence to make sure that the component words belonged to the most frequent 1000, 2000 or 3000 word level.

Note that there was no attempt to control for all the input the participants might have had from the curriculum or outside the treatment sessions. Considering the fact that these students are majoring in English, controlling for extra input was simply impossible. However, within the experiment the exposure to the target collocations was strictly monitored and controlled. The time-on-task factor did not greatly differ between treatments in the three experimental groups.

4.7.2.3. External validity

External validity “relates to the degree to which findings can be generalised/transferred to populations or situations” (Fred, 2011, p. 96). Deficiencies in a study’s internal validity limit the findings’ generalisability to a greater population. Many researchers argue that although a study that looks at causation might be designed so that a change in the dependent variable is only due to the independent variable, the results of this study can still not be generalised to the target population or situation, because the sample is simply not representative of that population or comparable to any other situation (Mackey & Gass, 2005). Additionally, it is incumbent upon researchers to make sure that the sample be of sufficient size to allow for generalisation of results. Larger samples mean a higher likelihood of only incidental variations between the sample and the population (ibid). Accordingly, cluster random sampling was employed in this research to ensure representativeness, and four intact classes of over 30 students in each were allocated for the research.

4.8 Ethical considerations

Ethical issues have received substantial attention in research literature. Dörnyei (2007) affirms that ethical issues are inevitable in social research (including research in education), because the research concerns people's lives in the social world. In second/foreign-language research, minimising potential ethical issues entails obtaining approval from institutions and informed consent from individuals to collect data from human subjects (Creswell, 2014; Dörnyei, 2007; Mackey & Gass, 2005). Additionally, learners might need to be notified that they might be allocated to a group that, theoretically, might benefit less than a treatment group. For example, in research on the effect of second/foreign-language instruction the control group may not receive equal instruction time to that of the experimental groups (Dörnyei, 2007; Mackey & Gass, 2005). Moreover, Duff and Early (1996, p. 21), in their discussion of participants' anonymity, state that "Although it is common practice to change the names of research subjects, this in itself does not guarantee subject anonymity. In reports of school-based research, prominent individuals or focal subjects tend to be more vulnerable." In some cases, the identification of students might have consequences for how other teachers perceive them, and accordingly might have an impact on their grades or letters of recommendations (Mackey & Gass, 2005). Dörnyei (2007) also reports that misusing test scores might entail real potential risks.

To alleviate these concerns and to comply with the research ethics regarding this research, Creswell's (2014) guidelines were followed during different stages of the research.

- Prior to conducting the study, the current researcher sought approval from Newcastle University through an institutional review board. She also sought approval from the university in Saudi Arabia where she conducted the study.
- At the beginning of the study, the current researcher approached teachers and participants and informed them of the general purpose of the research. Additionally, she informed the participants that they were not obliged to participate or to sign consent forms. They were informed that their participation or the lack of it would not affect their grades or assessment in any way and that their data and results would be anonymised. Additionally, the participants were informed that one group (the control group) would not be expected to benefit from the study as much as the other groups, although the precise nature of this benefit was not stated. The researcher then obtained consent from the participants (see appendices K and L).

- While collecting the data, the current researcher aimed to build trust and minimise any disruption by carrying out the treatment and the testing phases within the usual class times.

4.9 Summary

This chapter has provided a detailed account of the data collection and analysis methods to elicit data regarding the anticipated learning product as a result of three different treatments: (-DDL +CAT), (+DDL +CAT) and (+DDL -CAT).

Table 4. 7: Summary of research methods

Group	Treatment (part 2 of the teaching session)	Data collection method	Data analysis method
(+DDL+CAT)	Parallel corpus + E/A Translation worksheets	Pre, post, delayed post-tests of E/A translation	Statistical analysis (SPSS)
(-DDL +CAT)	Reading text+ E/A Translation worksheets		
(+DDL -CAT)	Monolingual corpus + MC/ fill in blank worksheets		

The learning product of the three treatments was measured by changes that occurred in the size of the learners' receptive and controlled productive knowledge of the target non-congruent collocations. The elicitation instruments were collocation (E/A – A/E) translation tests. It was hoped that such a multi-treatment investigation would create a more complete and comparative picture of the proposed pedagogical approach. The analysis and findings follow in the next chapter.

Chapter 5: Data Analysis and Results

In the previous chapter the instruments and processes of data collection were described. In this chapter, the analysis of the quantitative data gathered from the participants ($N= 129$) is presented and justifications for the utilised statistical procedures are provided. These data were obtained from the three stages of translation tests (pre, post and delayed post-treatment testing) given to three experimental groups and one control group. The chapter is divided into two main sections. The first section is devoted to a presentation of the data gathered from each of the groups and a description of the treatment effect among participants of that particular group. The second section is divided into sub-sections which correspond to the major themes of the research hypotheses. SPSS software was used to examine the quantitative data obtained from the pre, post and delayed post-tests.

5.1 Parametric versus non-parametric statistical tests

Data analysis using SPSS can be quite straightforward, however the selection of the appropriate test depends entirely on the decision of the researcher (Norusis, 2006). Therefore, in order to analyse the data obtained for this study, it was decided to calculate the statistics in the form of means (M), median (MD) and standard deviations (SD) and to use parametric and non-parametric tests. The decision to use parametric or nonparametric statistical tests is not random. Some scholars distinguish between parametric and non-parametric tests based on the level of measurement represented by the data being analysed. Thus, inferential statistical tests which evaluate interval data are categorised as parametric tests, whereas tests that evaluate nominal data and ordinal data are categorised as non-parametric tests (Sheskin, 2004).⁴⁷ According to other researchers, especially in the field of second-language and applied linguistics research, the distinction is not only made on the basis of the type of data, but also on the assumption of normality of the distribution of the data (Lowie & Seton, 2013).⁴⁸

⁴⁷ The interval scale of measurement is a numeric scale in which not only the order of the values is known, but also the exact differences/ intervals between the values (test scores are a typical example) (Dörnyei, 2007; Larson-Hall, 2010).

⁴⁸ The normality of distribution of data means that if the data were plotted, the result should be a symmetrical, bell-shaped curve, where the greatest frequency of score accumulates in the middle and the smaller frequencies fall towards the extremes (Dörnyei, 2007).

To make an objective decision on the normality of the data, it is recommended that a test of normality should be run (Dörnyei, 2007; Larson-Hall, 2010; Kinnear & Gray, 2012; Lowie & Seton, 2013). Fortunately, data does not have to be perfectly normal, because most procedures work well with data that is only approximately normally distributed (Dörnyei, 2007) and other procedures can work very well with non-normal data i.e. non-parametric tests.

It is commonly, although not accurately, accepted that parametric statistical tests, which are run when the data is normally distributed, provide a more powerful test of an alternative hypothesis than their non-parametric counterparts (Dörnyei, 2007; Sheskin, 2004). This assumption is rejected or at least not fully accepted by many researchers. For example, Larson-Hall (2010, p. 58) states:

“if we understand that the term “power” means the probability of finding a statistical difference when one exists, using either a parametric test or a non-parametric test when the data do not follow the assumptions can result in the loss of power to find statistical differences when they do in fact exist.”

He thus suggests that researchers should use either parametric or non-parametric statistical tests depending on which has more power to find statistical significance. In addition, it is argued by Sheskin (2004) that either choice is of little consequence in most instances. This is because most of the time, a parametric test and its non-parametric equivalent are utilised to evaluate the same set of data, and they lead to similar or identical conclusions (ibid). Nevertheless, the current researcher decided to check for the normality of distribution assumption in order to choose the best statistical test with the most powerful significant results. This is discussed in the following section.

5.1.1. Checking assumptions

The data obtained from the pre, post and delayed post-test can be categorised as interval data. However, being of an interval type alone does not make the data eligible for parametric tests. Therefore, the Shapiro-Wilk test was run before carrying out data analysis to check for the assumption of normality of distribution. The Shapiro-Wilk Test is a very powerful numerical method of assessing normality, which is more appropriate for small sample sizes (< 50 samples), but can also handle sample sizes as large as 2000 (Razali & Wah, 2011). The following tables show the normality of

distribution of the data obtained from pre, post and delayed post-tests, divided into passive/ active test results for each of the four groups.

The data from a particular test is considered by the researcher as normally distributed if the data set from all groups had p value $> .05$. In case of any violation of the normality assumption of the data distribution in any group or data set, the whole set of data was considered non-normal and non- parametric statistical tests were utilised accordingly.

Table 5. 1: Shapiro-Wilk test of normality of pre-tests

Treatment		Shapiro-Wilk		
		Statistic	df	<i>P.</i> value
pre-passive	group 1	.956	33	.118
	group 2	.959	32	.266
	group 3	.935	32	.053
	control	.934	32	.052
pre-active	group 1	.876	33	.001
	group 2	.850	32	.000
	group 3	.856	32	.001
	control	.906	32	.009

As mentioned earlier, the Shapiro-Wilk test was run on the data obtained from the pre-test for both the active knowledge results and the passive knowledge results for all groups. The results of the pre-test for the passive knowledge indicated that the data was normally distributed ($p > .05$). However, the results of the pre-test for the active knowledge were not normally distributed ($p < .05$).

Table 5. 2: Shapiro-Wilk test of normality of post-tests

Treatment		Shapiro-Wilk		
		Statistic	df	<i>P.</i> value
Post-passive	group 1	.811	33	.000
	group 2	.811	32	.000
	group 3	.833	32	.000
	control	.967	32	.421
Post-active	group 1	.946	33	.024
	group 2	.941	32	.082
	group 3	.962	32	.306
	control	.910	32	.011

As for the results of the post-test for the passive knowledge, the Shapiro-Wilk normality test results showed that the data was not normally distributed ($p < .05$). The post-test results for active knowledge were also found to be not normally distributed ($p < .05$).

Table 5. 3: Shapiro-Wilk test of normality of delayed post-tests

Treatment	Shapiro-Wilk			
	Statistic	df	<i>P.</i> value	
delayed-passive	group 1	.869	33	.001
	group 2	.869	32	.001
	group 3	.913	32	.014
	control	.977	32	.702
delayed-active	group 1	.955	33	.117
	group 2	.971	32	.517
	group 3	.966	32	.170
	control	.937	32	.060

Similarly, the results of the normality test run for the delayed post-test results for passive knowledge of collocation shows none-normally distributed data, whereas the results for active knowledge were normally distributed.

Another crucial assumption was also checked for the purpose of choosing the most appropriate statistical procedure to compare between groups: homogeneity of variance. The assumption here is that the variance within each of the populations is the same. To check whether different groups show similar variance is to compare the standard deviation (*SD*) (Lowie & Seton, 2013). If one *SD* is more than twice as big as that for another group, this means that the variance is not homogeneous (ibid).

5.2 Effect of treatments: within group comparisons

The normality tests' results were used in this section to select the appropriate statistical test. A paired-sample t-test (a parametric statistical test aimed at research designs where researchers want to compare two sets of scores obtained from the same group, or when the same participants are measured more than once (Dörnyei, 2007) was utilised. Additionally, a non-parametric statistical test equivalent to the paired-sample t-test called the 'Wilcoxon signed-rank test' was also used. Both procedures examine two different results from the same group (i.e. within-group comparison). In order to compare results from three related samples, as in the pre, post and delayed post-tests for the same group, the non-parametric Friedman test (the counterpart of the parametric one-way analysis of variance (ANOVA)) was the most appropriate statistical test (Larson-Hall, 2010; Corder & Foreman, 2011).

This section also reports on the effect size of each treatment within the groups. Effect size is one of the main variables involved in statistical inference which constitutes ‘power analysis’.⁴⁹ Effect size measures the degree to which a null hypothesis⁵⁰ is wrong (Grissom & Kim, 2005). It needs to be computed to provide information about the magnitude of an observed phenomenon since an existing statistical significance alone may have no practical or theoretical importance (Dörnyei, 2007). Coe (2002, p. 1) states:

“It allows us to move beyond the simplistic, 'Does it work or not?' to the far more sophisticated, 'How well does it work in a range of contexts?' Moreover, by placing the emphasis on the most important aspect of an intervention – the size of the effect - rather than its statistical significance..... it promotes a more scientific approach to the accumulation of knowledge. For these reasons, effect size is an important tool in reporting and interpreting effectiveness.”

Nonetheless, reporting effect size is continuously ignored by researchers (Cohen, 1992; Grissom & Kim, 2005; Dörnyei, 2007). Calculating and interpreting effect size is quite problematic as there are no universally accepted and straightforward indices. However, this process is easier when parametric statistical tests, such as a t-test, a paired-sample t-test or a one-way ANOVA, are utilised. It becomes more complicated when non-parametric tests are used. Authors such as Leech and Onwuegbuzie (2002) note that researchers who exploit non-parametric tests generally either do not report effect size estimates or report parametric effect size estimates. It is however acknowledged that these effect size estimates are adversely affected by a violation of normality and heterogeneity of variances. Thus, such estimates may not be well advised for use with the type of data which generally motivates a researcher to employ non-parametric tests. Accordingly, the current researcher utilised different formulas in accordance with the parametric and non-parametric statistical procedure being used in each section.

To calculate the effect size for the t-test, the formula $(r = \frac{t^2}{t^2 + (N_1 + N_2 - 2)})$ ⁵¹ recommended by Pallant (2007) and Dörnyei (2007) was used. The effect size for the paired-sample t-test was calculated using the formula recommended by Pallant (2007) which is $(r = \frac{t^2}{t^2 + (N_1 - 1)})$. Additionally, the formula $(r = \frac{SSM}{SST})$ ⁵² recommended by Pallant (2007),

⁴⁹ Power analysis utilises the relationship between the main variables involved in statistical inference: sample size, significance criterion and population effect size (Cohen, 1992).

⁵⁰ The null hypothesis suggests that there is no correlation between variables in the population or that there is no difference between the mean of populations (Grissom and Kim, 2005).

⁵¹ r = effect size, t = t value in the t-test, N = number of population.

⁵² SSM = sum of squares between groups, SST = total sum of squares.

Dörnyei (2007) and Lowie and Seton (2013) was utilised to calculate the effect size for the one-way ANOVA. However, there is no easy way of finding the effect size for the Friedman test, so the current researcher has performed Wilcoxon signed-rank tests and t-tests to derive the effect size between the pre-test and post-test results, between the pre-test and the delayed post-test results, and between the post-test and delayed post-tests results. The utilised effect size formula for the Wilcoxon signed-rank test is $(r = \frac{z}{\sqrt{N}})^{53}$ as recommended by Field (2013) and Pallant (2007). The same formula was also utilised to calculate the effect size whenever a Mann Whitney U test was performed.

5.2.1 Effect of (-DDL +CAT/ group 1) treatment on collocational knowledge

This section will first look at the effect of the treatment on the participants' passive and active knowledge of the target collocations in comparison to their entry level knowledge i.e. the participants' performance in the pre-test. Then, more test results will be presented to compare the computed results of the delayed post-tests with the previous two. The table below shows overall descriptive statistics on the learners' performances on the pre-test, post-test and delayed post-tests for passive and active collocational knowledge.

Table 5. 4: Descriptive statistics (-DDL +CAT/ group 1)

group 1 (N= 33)	Passive recall Max. = 30	Active recall Max. = 30
Pre-test	16.03 (17.00) SD 3.459	2.55 (2.00) SD 2.463
Post-test	27.06 (28.00) SD 3.211	19.85 (21.00) SD 5.263
Delayed post-test	26.48 (27.00) SD 3.374	17.03 (18.00) SD 5.676

The table shows clear discrepancies between the participants' passive and active collocational knowledge in the different testing phases. Notably, the participants' pre-passive collocational knowledge is greater than their pre-active knowledge as indicated by the mean (16.03 > 2.55) and the median (17.00 > 2.00) cores of the two tests. The fact

⁵³ z= z value in the Wilcoxon signed-rank test, N= number of population.

that progress made in the post-testing phase and retained in the delayed post-testing phase is not identical between passive and active knowledge was thus to be expected.

The following section considers the treatment’s effect on each level of the participants’ collocational knowledge in more detail.

5.2.1.1 Effect of (-DDL +CAT/ group 1) treatment on passive knowledge

As indicated by the results of the Shapiro test of normality, the scores of passive/receptive collocational knowledge obtained from the (-DDL +CAT/ group 1) treatment were normally distributed for the pre-test, but not for the post and delayed post-tests. Accordingly, the non-parametric Friedman test was run in order to compare the three test scores obtained from the experimental group 1, and to check for statistical differences between them. The statistical test rendered results as follows.

Table 5. 5: All passive recall tests (-DDL +CAT/ group 1)

		Chi-square	df	<i>p</i> . value
	Pre-passive			
Treatment/ group 1 (<i>N</i> = 33)	Post-passive	55.983	3	.000
	Delayed-passive			

Overall, the Friedman test shows that there was a significant statistical difference in the participants’ scores of passive knowledge across the three testing time points (pre-test, immediate post-test, three weeks delayed post-test), $\chi^2 (3, N= 33) = 55.98, p < .05$. Inspection of the median values showed an increase in the scores of the non-congruent collocation passive knowledge from pre-treatment (*MD*= 17.00) to post- treatment (*MD*= 28.00). It also shows a slight drop in the median value of the scores in the delayed post-test (*MD*= 27.00) in comparison to those of the post-test.

Having established that there is a statistically significant difference between the three testing phases, the next step was to run post-hoc tests (individual Wilcoxon Signed Rank tests) to compare the pre-test with both post-test and delayed post-test scores, and then compare post-test with delayed post-test scores. These measures would not only allow the current researcher to spot existing significant statistical differences and progressive changes in the participants’ passive knowledge of the target non-congruent collocations (if any), but it would also enable for the calculation of the actual size of these differences.

The first Wilcoxon test was run between the pre-test and post-test scores, and was intended to show the statistical differences between the performance of this experimental group at the entry level and immediately after the treatment.

Table 5. 6: Pre/ post-tests of passive recall (-DDL +CAT/ group 1)

		<i>z</i>	<i>p. value</i>	<i>r</i>
Treatment/ group 1 (<i>N</i> = 33)	Pre-passive	-5.027	.000	.9
	Post-passive			

The test showed a statistically significant increase in passive knowledge scores for the non-congruent collocations following (-DDL +CAT) treatment, $z = -5.027$, $p < .025$.⁵⁴ Experimental group 1 attained a median score of ($MD = 17.00$) in the pre-test ($SD = 3.459$), and progressed to ($MD = 28.00$) in the post-test ($SD = 3.211$).

Notwithstanding the relatively high entry-level performance on the collocational passive knowledge, experimental group 1 still seemed to benefit from the (-DDL +CAT) treatment, as evidenced by the progress in the post-test scores. Additionally, the magnitude of the difference in the means was large, as shown by the effect size ($r = .9$).⁵⁵ This indicates considerable positive variance in the participants' passive collocational knowledge after the treatment.

Table 5. 7: Pre /delayed post-tests of passive recall (-DDL +CAT/ group 1)

		<i>z</i>	<i>p. value</i>	<i>r</i>
Treatment/ group 1 (<i>N</i> = 33)	Pre-passive	-5.022	.000	.9
	Delayed-passive			

Similarly, the statistical test results obtained when the pre-test and the delayed post-test results were compared showed significant statistical differences in the median scores for the pre-test ($MD = 16.03$, $SD = 3.459$) and delayed post-test ($MD = 26.48$, $SD = 3.374$), $z = -5.022$, $p < .025$ with the magnitude of the difference between the means score still high ($r = .9$). This means that the participants' passive collocational knowledge was still higher than their entry-level knowledge even three weeks after the post-treatment test.

⁵⁴ A Bonferroni correction to the alpha value ($.05/2 = .025$) was applied to control for Type 1 errors as recommended by Pallant (2007).

⁵⁵ Effect size for the Wilcoxon signed-rank test was interpreted according to Cohen (1988) criteria of .1= small effect, .3= medium effect, .5= large effect

Table 5. 8: Post/ delayed post-tests of passive recall (-DDL +CAT/ group 1)

		Z	P. value	r
Treatment/ group 1 (N= 33)	Post-passive	-2.027	.043	.4
	Delayed-passive			

As shown in the descriptive statistics, there was a slight decrease in the median scores in the delayed post-test in comparison to the immediate post-test. However, a third Wilcoxon signed rank test revealed that there were no significant statistical differences between the median scores of the two groups $z = -2.027$, $p > .025$, which indicates that this drop was not statistically significant. The size of the difference, however, was found to be medium. Nonetheless, it can still be concluded that the (-DDL +CAT) treatment had a positive effect on the participants' passive knowledge of the target collocation.

5.2.1.2 Effect of (-DDL +CAT/ group 1) treatment on active knowledge

The table below presents Friedman test's results showing the statistical differences between the scores attained by group 1 participants in the three testing phases.

Table 5. 9: All active recall tests (-DDL +CAT/ group 1)

		Chi-square	df	p. value
Treatment/ group 1 (N= 33)	Pre-active	60.813	2	.000
	Post-active			
	Delayed-active			

The test clearly indicates that there was a statistically significant difference in the scores across the pre, post and delayed post-tests for active knowledge of the target non-congruent collocations, $\chi^2(2, N= 33) = 60.813$, $P < .05$. Additionally, checking of the median values showed a considerable escalation in active knowledge scores from pre-treatment ($MD = 2.00$) to post-treatment ($MD = 21.00$). The median value of the delayed post-treatment test scores ($MD = 18.00$) dropped slightly in comparison to that of the post-test, while still higher than the median value of the pre-test scores. To determine the statistical differences in active knowledge scores between the different testing time points, three post-hoc tests involving individual Wilcoxon signed-rank tests with a Bonferroni adjusted alpha value of .025 were run.

Table 5. 10: Pre/ post-tests of active recall (-DDL +CAT/ group 1)

		<i>z</i>	<i>p. value</i>	<i>r</i>
Treatment/ group 1 (<i>N</i> = 33)	Pre-active	-5.016	.000	.9
	Post-active			

A comparison between the pre-test for active collocational knowledge and the immediate post-test, which was made using a Wilcoxon signed-rank test, showed a statistically significant increase in the students' active knowledge of non-congruent collocations after the treatment, $z = -5.015$, $p < .025$. The development in the participants' active knowledge is evident in the median scores of the post-test ($MD = 21.00$, $SD = 5.263$) compared to the median scores achieved in the pre-test ($MD = 2.00$, $SD = 2.463$). The magnitude of this difference is high ($r = .9$). This suggests a considerable positive change in the participants' active collocational knowledge as a result of the treatment.

Table 5. 11: Pre/ delayed post-tests of active recall (-DDL + CAT/ group 1)

		<i>z</i>	<i>P. value</i>	<i>r</i>
Treatment/ group 1 (<i>N</i> = 33)	Pre-active	-5.015	.000	.9
	delayed-active			

The results from the second statistical test show a statistically significant difference between the scores obtained in the pre-test and the delayed post-test for the active collocational knowledge $z = -5.015$, $p < .025$. The median scores of the delayed post-test ($MD = 18.00$, $SD = 5.676$) are higher than the mean scores of the pre-test ($MD = 2.00$, $SD = 2.463$). The difference proves to be not only statistically significant, but also of a large size ($r = .9$).

Table 5. 12: Post/ delayed post-test of active recall (-DDL + CAT/ group 1)

		<i>z</i>	<i>P. value</i>	<i>r</i>
Treatment/ group 1 (<i>N</i> = 33)	Post-active	-4.531	.000	.9
	delayed-active			

The above Wilcoxon signed-rank test results confirm the alternative hypothesis⁵⁶ i.e. there is a significant statistical difference between the median scores of the post-test and

⁵⁶ The alternative hypothesis is a counterpart to the null hypothesis i.e. it predicts that there is a correlation between variables in the population or that there is a difference between the mean of populations.

delayed post-test for active collocational knowledge $z = -4.531, p < .025$. The magnitude of this difference was found to be large, indicating a considerable drop in the participants' scores of the non-congruent collocations' active knowledge. However, the statistical differences and effect size, found when the participants' delayed post-test results were compared to their prior knowledge in the pre-test, cannot be neglected. They provide evidence that in comparison to their prior active knowledge of the target collocations, the participants still attained prominent gains.

5.2.2 Effect of (+DDL + CAT/ group 2) treatment on collocational knowledge

This section will present analysis of the data gathered from group 2 in order to examine the effect of (+DDL +CAT) treatment on the participants' passive and active knowledge of the target non-congruent collocations in comparison to their entry-level knowledge. The table below shows descriptive statistics obtained from all of the three testing stages for both passive and active levels of knowledge.

Table 5. 13: Descriptive statistics (+DDL +CAT/ group 2)

group 2 (N= 32)	Passive recall Max. = 30	Active recall Max. = 30
Pre-test	16.66 (16.50) SD 2.209	2.59 (2.00) SD 2.525
Post-test	28.72 (29.00) SD 1.508	22.72 (23.00) SD 3.612
Delayed post-test	28.28 (29.00) SD 1.727	20.63 (20.50) SD 4.361

The descriptive statistics, once again, shows variances between the scores of the passive knowledge tests and the active knowledge tests at the entry level. The participants' prior passive knowledge of the target collocations is clearly higher than their active knowledge ($M = 16.66$) ($MD = 16.50$) vs. ($M = 2.59$) ($MD = 2.00$). This also applies to the participants' passive knowledge test scores immediately after the treatment and in a delayed post-test.

5.2.2.1 Effect of (+DDL +CAT/ group 2) treatment on passive knowledge

The table below provides a holistic view of the statistical differences between the scores achieved by participants in group 2 in the three tests of passive knowledge of target collocations.

Table 5. 14: All passive recall tests (+DDL +CAT/ group 2)

		Chi-square	df	P. value
	Pre-passive			
Treatment/ group 2 (N= 32)	Post-passive	53.965	2	.000
	Delayed-passive			

This Friedman test shows that there is a significant statistical difference between the results yielded in the pre, post and delayed testing phases $\chi^2 (2, N= 32) = 53.96, p < .05$. The median values of the delayed post-test ($MD= 29.00$) ($SD= 1.727$) and the immediate post-test ($MD= 29.00$) ($SD= 1.508$) did not vary. However, the differences of both median values compared to the participants' entry level in the pre-test ($MD= 16.50$) ($SD= 2.209$) is considerable. This suggests a strong positive effect of the (+DDL +CAT) treatment on the participants' passive collocational knowledge.

Table 5. 15: Pre/ post-tests of passive recall (+DDL +CAT/ group 2)

		z	p. value	r
Treatment/ group 2 (N= 32)	Pre-passive			
	Post-passive	-4.947	.000	.9

As shown in the table above, the passive knowledge of the participants in this experimental group at an entry level was quite high, as represented by the pre-test scores ($MD= 16.66$) ($SD= 2.209$). Nonetheless, considerable development in their passive knowledge was attained in the post-test as indicated by the median scores ($MD= 29.00$) ($SD= 1.508$). The results of a Wilcoxon signed-rank test also show that ($z= -4.947, p < .025$), which confirms the alternative hypothesis i.e. there is a significant statistical difference between the participants' performance in the pre and post-test. This difference is clearly in favour of the post-test results. The calculated magnitude of this effect turned to be high ($r= .9$).

Table 5. 16: Pre/ delayed post-tests of passive recall (+DDL +CAT/ group 2)

		<i>z</i>	<i>p. value</i>	<i>r</i>
Treatment/ group 2 (N= 32)	Pre-passive	-4.950	.000	.9
	Delayed-passive			

When another Wilcoxon signed-rank test was run, the results showed that significant statistical differences continue to exist in a delayed post-test ($MD= 29.00$) ($SD= 1.727$) in comparison to the pre-test results ($MD= 16.50$) ($SD= 2.209$) ($z= -4.950$, $P< .025$). The effect size also continues to be high ($r= .8$). This indicates a positive impact on the participants' passive collocational knowledge as a result of the treatment.

Table 5. 17: Post/ delayed post-tests of passive recall (+DDL +CAT/ group 2)

		<i>z</i>	<i>p. value</i>	<i>r</i>
Treatment/ group 2 (N= 32)	Post-passive	-1.558	.119	.2
	Delayed-passive			

In order to validate the claimed lasting effect of the treatment in a delayed post-test, a Wilcoxon signed-rank test was run between the post-test and the delayed post-test for passive collocational knowledge. The statistical test results show that there is no statistical discrepancy between the two test scores ($z= -1.558$, $p> .025$). Additionally, when the magnitude of the statistical difference was calculated, it was found to be quite small. Thus, this researcher could safely claim a lasting effect of the treatment on the participants' passive collocational knowledge.

5.2.2.2 Effect of (+DDL + CAT/ group 2) treatment on active knowledge

Once again due to the violation of the normality of distribution assumption, a Friedman test was carried out to check for progressive and significant differences (if any) in the target non-congruent active knowledge across the three testing phases attained by participants in group 2.

Table 5. 18: All active recall tests (+DDL +CAT/ group 2)

		Chi-square	df	<i>p. value</i>
Treatment/ group 2 (N= 32)	Pre-active	54.774	2	.000
	Post-active			
	Delayed-active			

The results of the test reveal a statistically significant difference between the participants' scores in the three testing phases, $\chi^2 (2, N= 32) = 54.77, p < .005$. This shows a significantly low entry level for participants' active collocational knowledge ($M= 2.59$) ($SD= 2.525$), and a progress attained in the post-test as indicated by the median value ($MD= 23.00, SD= 3.612$). It also shows a continued high level of performance in the delayed post-test ($MD= 20.50, SD= 4.361$) in comparison to the performance in the pre-test. However, one can still notice the drop in the mean scores of the delayed post-test in comparison to the immediate post-test mean scores. To assess the significance of the differences between the pairs of groups and to calculate its size, follow-up statistical tests were run.

Table 5. 19: Pre/ post-tests of active recall (+DDL +CAT/ group 2)

		<i>z</i>	<i>p. value</i>	<i>r</i>
Treatment/ group 2 (<i>N</i> = 32)	Pre-active	-4.941	.000	.9
	Post-active			

The results obtained from a Wilcoxon signed-rank test show that the statistical difference between the results of the pre-test and the immediate post-test is significant ($z= -4.941, p= <.025$). The effect size between these two tests is large ($r= .9$).

Table 5. 20: Pre/ delayed post-tests of active recall (+DDL +CAT/ group 2)

		<i>z</i>	<i>p. value</i>	<i>r</i>
Treatment/ group 2 (<i>N</i> = 32)	Pre-active	-4.941	.000	.9
	Delayed-active			

The second Wilcoxon signed-rank test results, comparing the pre-test and the delayed post-test, also reveal a significant statistical difference between the test scores ($z= -4.941, p < .025$). The magnitude of the difference is high as well ($r= .9$). This is a clear indication of the effective impact of the treatment on the participants' active knowledge of the target collocations.

Table 5. 21: Post/ delayed post-tests of active recall (+DDL +CAT/ group 2)

		<i>t</i>	<i>p. value</i>	<i>r</i>
Treatment/ group 2 (<i>N</i> = 32)	Post-active	4.587	.000	.4
	Delayed-active			

The statistical significance of the differences between the post-test and the delayed post-test's mean scores was checked by running a paired-sample t-test since the scores were normally distributed. The test results in the table above show that, statistically, there is actually a significant difference between the scores of active collocational knowledge in the two testing time points. Interestingly, the magnitude of the difference in the means was also found to be large ($r = .4$).⁵⁷ This suggests that active recall of the target collocations in the delayed post-test, though significant, is not as strong and evident as in the immediate post-test.

5.2.3 Effect of (+DDL -CAT/ group 3) treatment on collocational knowledge

This section will look at the impact of the data-driven/ corpus-based tasks which do not involve comparative analysis and translation on the passive and active knowledge of the target collocations in group 3. The table below presents the initial descriptive statistics. Inferential statistics are presented in the next sections.

Table 5. 22: Descriptive statistics (+DDL -CAT/ group 3)

group 3 (<i>N</i> = 32)	Passive recall	Active recall
	Max. = 30	Max. = 30
Pre-test	16.88 (17.00) <i>SD</i> 2.498	2.69 (1.50) <i>SD</i> 2.753
Post-test	26.47 (28.00) <i>SD</i> 3.473	16.34 (15.50) <i>SD</i> 5.277
Delayed post-test	25.91 (27.00) <i>SD</i> 3.354	13.63 (13.00) <i>SD</i> 4.361

It can be seen from the table above that the participants' passive and active collocational knowledge in the three testing phases is different. In the three testing phases of passive

⁵⁷ Effect size for paired-sample t-test was interpreted according to Cohen (1988) criteria of .1= small effect, .06= medium effect, .14= large effect.

collocational knowledge, the participants achieved median scores of ($MD= 17.00$, $MD= 28.00$, $MD= 27.00$) consecutively. In the active collocational knowledge test on the other hand, they attained median scores of ($MD= 1.50$, $MD= 15.50$, $MD= 13.00$). Though not at the same progression rate, both the mean scores of the passive and of the active knowledge of the target collocations have considerably increased in the immediate and delayed post-tests. The following sections will look at more detailed inferential tests to investigate statistical differences in the participants' performances in the three testing time points.

5.2.3.1 Effect of (+DDL -CAT/ group 3) treatment on passive knowledge

As in the previous sections, a Friedman test was initially carried out to examine the differences in the participants' passive knowledge of the target non-congruent collocations. The test rendered the following results.

Table 5. 23: All passive recall tests (+DDL -CAT/ group 3)

		Chi-square	df	P. value
Treatment/ group 3 (N= 32)	Pre-passive			
	Post-passive	53.600	2	.000
	Delayed-passive			

The statistical test shows significant statistical differences between the participants' passive knowledge of the target collocations in the three testing phases, $\chi^2 (2, N= 32) = 53.60$, $p < .005$. Inspection of the median values showed that there is progress in the participants' performance in the post-test ($MD= 28$, $SD= 3.473$). The progress continues to be seen in the delayed post-test, however, it is slightly lower than that of the post-test as indicated by their median value ($MD= 27$, $SD= 3.354$).

Table 5. 24: Pre/ post-tests of passive recall (+DDL -CAT/ group 3)

		Z	P. value	r
Treatment/ group 3 (N= 32)	Pre-passive			
	Post-passive	-4.952	.000	.9

As shown in the table above, there is indeed a significant statistical difference between the two test results as indicated by the Wilcoxon signed-rank test ($z= -4.952$, $p < .025$). The participants in this treatment group have attained a median score of ($MD= 16.88$, $SD= 2.498$) in their pre-testing phase, exhibiting relatively high entry-level passive

knowledge of the target non-congruent collocations. However, noticeable progress in their mean score can still be seen in the mean score of their immediate post-testing phase ($M= 26.47, SD= 3.473$). Upon calculation, this difference in mean scores between the two testing times was found to be of a large size ($r= .9$). This suggests a positive effect of the treatment on the participants' passive collocational knowledge.

Table 5. 25: Pre/ delayed post-tests of passive recall (+DDL -CAT/ group 3)

		Z	P. value	r
Treatment/ group 3 (N= 32)	Pre-passive	-4.950	.000	.8
	Delayed-passive			

The significant statistical difference in the median scores found between the pre and post-tests was maintained in the comparison conducted between the pre-test and delayed post-test ($z= -4.950, p< .025$). The median scores of the passive knowledge of the target collocations attained by the participants increased from the pre-treatment test ($MD= 17.00$) to the delayed post-treatment test ($MD= 27.00$).

Table 5. 26: Post/ delayed post-tests of passive recall (+DDL -CAT/ group 3)

		Z	P. value	r
Treatment/ group 3 (N= 32)	Post-passive	-1.935	.053	.3
	Delayed-passive			

A third Wilcoxon signed-rank test was run to examine the differences between the two post-treatment tests and to calculate the size of the decrease in the mean scores which were observed in the descriptive statistics. The test revealed significant statistical differences between the mean scores of the two tests ($z= -1.935, p= .053> .025$). Additionally, the calculated size of the difference in the means was medium ($r= .3$).

5.2.3.2 Effect of (+DDL -CAT/ group 3) treatment on active knowledge

A Friedman test was carried out to compare the pre-test, post-test and delayed post-test scores for active knowledge of the target collocations. The following table shows the results of the comparison.

Table 5. 27: All active recall tests (+DDL -CAT/ group 3)

		Chi-square	df	P. value
Treatment/ group 3 (N= 32)	Pre-active	60.816	2	.000
	Post-active			
	Delayed-active			

There was a statistically significant difference in the median scores between the entry-level knowledge in the pre-test, which was low ($MD= 2.69$, $SD= 2.753$), and the participants' knowledge immediately and three weeks after the treatment, which is evident in the post-test median scores ($MD= 16.43$, $SD= 5.277$) ($MD= 13.63$, $SD= 4.361$) $\chi^2 (2, N= 32) = 60.816$, $p < .05$. More tests were run to identify the significance of the difference between the mean scores of the participants in the three tests and the size of the difference (if any).

Table 5. 28: Pre/ post-tests of active recall (+DDL -CAT/ group 3)

		Mean	Std. Deviation	Z	P. value	r
Treatment/ group 3 (N= 32)	Pre-active	2.69	2.753			
	Post-active	16.34	5.277	-4.940	.000	.9

A Wilcoxon signed-rank test (above) indicated that significant statistical differences exist between the pre and post-tests of active collocational knowledge ($z= -4.940$, $p < .025$). The difference in mean scores between the two tests was found to be large ($r= .9$).

Table 5. 29: Pre/ delayed post-tests of active recall (+DDL -CAT/ group 3)

		Z	P. value	r
Treatment/ group 3 (N= 32)	Pre-active	-4.943	.000	.9
	Delayed-active			

A second statistical test confirmed the positive effect of the treatment, showing significant differences between the pre-test and the delayed post-test ($P < .025$). The participants maintained high scores in the delayed post-test for active knowledge ($M= 14.31$, $SD= 4.361$) in comparison to their entry level scores in the pre-test ($M= 2.69$, $SD= 2.753$). The difference in the means between the two tests has was large ($r= .9$).

Table 5. 30: Post/ delayed post-tests of active recall (+DDL -CAT/ group 3)

		<i>t</i>	df	<i>P. value</i>	<i>r</i>
Treatment/ group 3 (<i>N</i> = 32)	Post-active	7.124	31	.000	.6
	Delayed-active				

The paired sample t-test results in the table above show that a significant statistical difference does exist between the immediate post-test and the delayed post-test ($p = .000 < .025$). Additionally, when the size of this difference was calculated, it turned to be very significant as well ($r = .6$).⁵⁸ This indicates that knowledge of the target collocations in delayed post-test was not as good as it was in the immediate post-test.

5.3 Effect of no-treatment: control group

Despite the fact that the control group did not receive any collocational instruction in this research, examining any development, or lack thereof, is crucial. Since the current researcher has no control over the input all participants in all groups were exposed to as a result of majoring in English, results obtained from this group would support or refute any claims made by the researcher on the effectiveness of the explicit instruction in the experimental treatments. Hence, this section will look at the scores obtained by the participants in the control group (under no-treatment condition). The table below presents the descriptive statistics of the passive and active recall test scores.

Table 5. 31: Descriptive statistics of the control group

No treatment/ control group (<i>N</i> = 32)	Passive recall	Active recall
	Max. = 30	Max. = 30
Pre-test	16.78 (17.00) <i>SD</i> 3.098	2.50 (2.00) <i>SD</i> 2.214
Post-test	18.41 (18.00) <i>SD</i> 3.991	4.59 (4.00) <i>SD</i> 3.555
Delayed post-test	18.38 (18.00) <i>SD</i> 3.892	4.31 (4.00) <i>SD</i> 3.459

⁵⁸ Using Cohen's (1988, 1992) criteria of .01= small effect, .06= medium effect, .14= large effect.

The descriptive statistics shown in the table above reveals that the participants' passive and active collocational knowledge in the three testing phases is different. In the pre, post and delayed-post testing phases of the passive collocational knowledge, the participants achieved mean scores of ($M= 16.78$, $M= 18.41$, $M= 18.38$) consecutively. In the active collocational knowledge test on the other hand, they achieved mean scores of ($M= 2.50$, $M= 4.59$, $M= 4.31$). The mean scores on both passive and active recall tests have increased in the immediate and delayed post-tests, though very slightly. The following sections will present the statistical differences in the participants' performances in the three testing stages in more detail.

5.3.1 The control group performance on passive recall tests

As revealed in the test of normality of distribution, the data from the pre, post and delayed post-test were normally distributed. Hence, a one-way repeated measure ANOVA was carried out to examine the differences in the participants' passive knowledge of the target non-congruent collocations in the three testing stages. The test showed the following results.

Table 5. 32: All passive recall tests of the control group

		Wilk's lambda	f	Sig.	r.
Control group ($N= 32$)	Pre-passive				
	Post-passive	.576	11.047	.000	.4
	Delayed-passive				

The statistical test revealed that there is a significant statistical difference between the results yielded in the pre, post and delayed testing phases, Wilk's lambda= .576, $F(2, 30)= 11.043$, $p < .0005$, multivariate eta squared $r= .424$. The mean values of the immediate post-test ($M= 16.78$) ($SD= 3.09$) and the delayed post-test ($M= 18.38$) ($SD= 3.892$) did not differ significantly. However, the difference of both mean values compared to the participants' entry level in the pre-test ($M= 16.78$) ($SD= 3.098$) is quite visible. As indicated by the eta-squared value, the size of the differences is large. Although the results from this test are indicative of some statistical significance among the three sets of scores, they do not reveal how these scores differ. Hence, a series of post hoc paired-sample t-tests were carried out.

Table 5. 33: Paired-sample t-tests of passive recall tests (control group)

Control group (<i>N</i> = 32)		<i>t</i>	df	<i>p</i> . value	<i>r</i>
	Pre-passive				
	Post-passive	-4.131	31	.000	.4
	Pre-passive				
	Delayed-passive	-4.666	31	.000	.4
	Post-passive				
	Delayed-passive	.115	31	.909	.0

The post hoc tests showed that there was a statistically significant increase in the learners' receptive/ passive knowledge of collocations between the pre-test ($M= 16.78$) ($SD= 3.098$) and the post-test ($M= 18.41$) ($SD= 3.991$), $t(31) = -4.131, p < .0005$. The calculated eta squared $r = .4$ indicates a large effect size. Similarly, the comparison carried out between the pre-test ($M= 16.78$) ($SD= 3.098$) and delayed post-test ($M= 18.38$) ($SD= 3.892$) showed a statistically significant increase in the learners' scores for passive recall $t(31) = -4.666, p < .0005$. The effect size was also found to be large. Additionally, the statistical tests showed that the decrease in the scores of the delayed post-test ($M= 18.38$) ($SD= 3.892$) in comparison to the immediate post-test ($M= 18.41$) ($SD= 3.991$) was statistically insignificant $t(31) = .115, p < .0005$, eta squared $r = .0$.

5.3.2 The control group performance on active recall tests

The scores of the active recall tests obtained from the control group's participants were not normally distributed according to the normality of distribution test. Accordingly, the non-parametric Friedman test was run to examine the differences in the learners' performance across the three testing times.

Table 5. 34: All active recall tests of the control group

		Chi-square	df	<i>p</i> . value
Control group (<i>N</i> = 32)	Pre-active			
	Post-active	24.641	2	.000
	Delayed-active			

The results of the above test suggest that there are significant differences in the active recall scores across the three testing times. This is indicated by a Sig. level of .000, which means a p value of $< .0005$. Comparing the mean ranks of the three sets of scores shows an increase in the scores of the active recall scores from the pre-test ($MD= 2.50$) to the

post-test ($MD= 4.59$). The mean rank decreased in the delayed post testing phase ($MD= 4.31$) in comparison to the immediate post-test, but not when compared to the pre-test. This statistical test could not determine, however, whether this increase or decrease was statistically significant or not. Hence, the following series of Wilcoxon signed-rank tests were run between pairs of scores set.

Table 5. 35: Wilcoxon signed-rank tests of active recall tests (control group)

Control group ($N= 32$)	z	p . value	r
Pre-active			
Post-active	-3.707	.000	.6
Pre-active			
Delayed-active	-3.500	.000	.6
Post-active			
Delayed-active	-.842	.400	.1

As shown in the table above, a significant statistical difference was found between the median scores of the pre-test ($MD= 2.00$, $SD= 2.214$) and the post-test ($MD= 4.00$, $SD= 3.555$), $z= -3.707$, $p < .025$. The calculated eta squared $r= .6$ showed a large effect size. The significant statistical difference with large effect size was retained in the delayed post-test ($MD= 4.00$, $SD= 3.459$) in comparison to the pre-test $z= -3.500$, $p < .025$, $r= .6$. The drop in median scores between the two post-tests was statistically insignificant and the size of it was small $z= -.842$, $p > .025$, $r= .1$.

5.4 Summary

The statistical procedures (Friedman test, Wilcoxon signed-rank test, one-way repeated measure ANOVA, and paired-sample t-test) were utilised to compare the sets of scores obtained from each of the experimental groups at three testing points. The statistical tests investigated the effectiveness of each of the FonFs treatments as well as the effect of no treatment on the learners' collocational knowledge, by examining the statistical differences between the participants' scores in the different testing stages. The results obtained from these statistical tests show three main results. First, all treatments proved to have a positive impact of the participants' passive and active knowledge of the target collocations as evident in the progress they made in the immediate post-testing phase in comparison to their entry-level knowledge. This is also evident in the participants retaining passive and active knowledge in the delayed-testing phase three weeks after the

immediate post-tests. Second, the results obtained from the control group which received no treatment also showed significant changes in the participants' performance on passive and active recall tests. Finally, on all the tests, the scores of passive recall are higher than the corresponding scores of active recall, which is not surprising. Vocabulary learning is an incremental process and learners usually acquire passive knowledge of a word before they acquire its active knowledge (Laufer 1998; Laufer and Goldstein, 2004; Laufer, 2008a; Webb 2005, see also section X above).

Effect size is a crucial and useful way for quantifying the effectiveness of a particular intervention (Coe, 2002), and was calculated for within-group comparisons. Within each experimental treatment, the calculated effect sizes proved to be large whenever a comparison was made between the pre-test and post-test mean scores for both the passive and active knowledge of collocation. Given the fact that there was a decrease in the delayed-test mean scores for both passive and active knowledge of the target collocations, the size of this difference needed to be calculated. For passive knowledge of the target collocations, it was found that the magnitude of the decrease in the mean scores was medium in group 1 (-DDL +CAT), small in group 2 (+DDL +CAT) and large in group 3 (+DDL -CAT). For active knowledge of the target collocations, on the other hand, it was found that the size of the difference in mean scores between the post-tests was large for all groups. However, when the pre-test and delayed post-test were compared in both levels of knowledge, the derived magnitudes of differences in the mean scores were found to be large. Thus, it can be concluded that each treatment in itself had a positive impact on the participants' immediate passive and active knowledge of the target collocations as well as a positive lasting impact on their passive and active knowledge even three weeks later, though they differ in the size of effectiveness.

The following section will be allocated for depicting the differences between the experimental treatments (i.e. groups) regarding their effectiveness and their impact on the participants' acquisition of the target non-congruent lexical collocations on passive and active levels of knowledge.

5.5 Effect of the treatments: between-group comparisons

This section will be allocated to presenting the findings regarding the effect of the three treatments (-DDL +CAT/ +DDL +CAT/ +DDL -CAT), and the effect of no treatment (control group) on the participants' passive and active knowledge of the target non-congruent collocations. This will be done by comparing the results of the translation test scores for both levels of knowledge between the groups. Once again, the normality tests' results were used to decide on the appropriate statistical test to compare between the groups. The parametric statistical test ANOVA and the non-parametric test Kruskal-Wallis were utilised. ANOVA aims at assessing the significance of the differences in the means between more than two groups; Kruskal-Wallis is its non-parametric alternative. The tests were utilised to compare between the mean/ median scores of the four groups at their entry level knowledge, and later to compare between the experimental groups and the control group in the post and delayed-post testing phases. An independent samples t-test was utilised when no violation of the normal distribution assumption or homogeneity of variance was found, to compare between pairs of groups. Alternatively, the non-parametric Mann-Whitney U test was used for the same purpose.

5.5.1 Entry level

Participants' passive and active knowledge of the target non-congruent lexical collocations was elicited by means of collocation translation tests in three testing phases; a pre-test, a post-test and a delayed post-test. Each test comprised two parts, passive knowledge and active knowledge of the target non-congruent collocations. This subsection presents the participants' performances in the pre-test. The pre-test aimed to determine participants' entry knowledge of the target collocations at both passive and active levels. Table (5.36) below shows the descriptive statistics of the pre-test scores.

Table 5. 36: Descriptive statistics of the pre-test scores of all groups

Parameter	Pre-passive	Pre-active
	Max. = 30	Max. = 30
group 1 (N= 33)	16.03 (17.00) SD 3.459	2.55 (2.00) SD 2.463
group 2 (N= 32)	16.66 (16.50) SD 2.209	2.59 (2.00) SD 2.525

group 3 (N= 32)	16.88 (17.00) SD 2.498	2.69 (1.50) SD 2.753
Control group (N= 32)	16.78 (17.00) SD 3.098	2.50 (2.00) SD 2.214

By examining the mean scores attained by the three experimental groups and the control group in both levels of knowledge (shown in table 5.36), it is noticeable that they are rather close in their values. However, two tests were carried out to examine the statistical difference between the groups: one-way analysis of variance (ANOVA) and a Kruskal Wallis test.

Table 5. 37: ANOVA's test results of pre-tests of passive recall between all groups

	Sum of squares	df	Mean square	f	p. value	r
Between groups	14.238	3	4.746			
Within groups	1025.157	125	8.201	.579	.630	.01
Total	1039.395	128				

As indicated in table (5.37) above, there was no significant statistical difference in *passive* knowledge of the target collocations among the experimental groups (GP1, $M= 16.03$, $SD= 3.459$), (GP2, $M= 16.66$, $SD= 2.209$), (GP3, $M= 16.88$, $SD= 2.498$) and the control group (CG, $M= 16.78$, $SD= 3.098$) ($p= .871 > .001$). Moreover, the actual difference in size between the mean scores of the groups is extremely small ($r= .01$).⁵⁹

Table 5. 38: Kruskal-Wallis test of pre-tests of the active recall between all groups

Parameter	Groups/ treatments	Median	Chi-square	df	p. value
Pre-active	Exp. group 1	2.00	.026	3	.999
	Exp. group 2	2.00			
	Exp. group 3	1.50			
	Control group	2.00			

⁵⁹ This is according to Cohen's (1988) criteria of .1= small effect, .6= medium effect, .14= large effect.

As for the participants' active knowledge of the target collocations, results in table (5.38) indicate that there was also no significant difference between the experimental groups (GP1, $N= 33$), (GP2, $N= 32$), (GP3, $N= 32$), and the control group (CG, $N= 32$), $\chi^2 (N= 129) = .026, p > .05$ ⁶⁰. These results from both passive and active collocational knowledge of the participants in all four groups show no statistically significant differences in the entry level between the groups.

5.5.2 Participants performance on post-tests: All groups

This section gives a comprehensive look at the differences in the participants' performance in the post-tests and delayed post-tests of passive and active knowledge of the target non-congruent collocations across all four groups i.e. three experimental groups and the control group. Later, a more detailed and closer look at the differences between the groups will be presented.

5.5.2.1 Participants performance on post-tests of passive collocational knowledge

This section will look at the differences in the participants' performances in the post-tests and delayed post-tests for passive knowledge of the target non-congruent collocations. Table (5.39) presents the descriptive statistics of the tests for all groups.

Table 5. 39: Descriptive statistics of the post- tests of passive recall of all groups

Parameter	Post-passive	Delayed-passive
	Max. = 30	Max. = 30
group 1 ($N= 33$)	27.06 (28.00) <i>SD</i> 3.211	26.48 (27.00) <i>SD</i> 3.374
group 2 ($N= 32$)	28.72 (29.00) <i>SD</i> 1.508	28.28 (29.00) <i>SD</i> 1.727
group 3 ($N= 32$)	26.47 (28.00) <i>SD</i> 3.473	25.91 (27.00) <i>SD</i> 3.354
Control group ($N= 32$)	18.41 (18.00) <i>SD</i> 3.991	18.00 (2.00) <i>SD</i> 3.892

⁶⁰ It is worth noting that like the Friedman test, there is no particular formula to calculate the effect size of the difference between the groups. Hence, the effect size will be calculated when comparing between pairs of groups.

The mean scores attained in the post-tests as well as the delayed post-tests show that the participants in the (+DDL +CAT) condition achieved the highest mean scores in both tests, followed by the (-DDL +CAT) condition, then the (+DDL -CAT). The control group which received no treatment expectedly achieved the lowest scores. It is widely acknowledged that descriptive statistics do not allow for any general conclusions to be drawn that would go beyond the sample. Hence, to ensure that no significant statistical differences exist between the groups, and to draw generalizable results, further analysis was conducted and will be described in the following sections.

Two Kruskal-Wallis tests were run to examine the statistical differences between the four groups regarding their scores in the post-tests and delayed post-tests for passive knowledge of the target collocations. It rendered the following results.

Table 5. 40: Post-tests of the passive collocational knowledge between all groups

Parameter	Groups/ treatments	Chi-square	df	<i>p</i> . value
Post-passive	Exp. group 1	67.337	3	.000
	Exp. group 2			
	Exp. group 3			
	Control group			
Delayed-passive	Exp. group 1	64.998	3	.000
	Exp. group 2			
	Exp. group 3			
	Control group			

The first statistical test revealed a statistically significant difference in the participants' passive knowledge of the target non-congruent collocations across the three experimental groups (GP1, *MD*= 28), (GP2, *MD*= 29), (GP3, *MD*= 28), and the control group (CG, *MD*= 18), $\chi^2 (N= 129) = 67.33, p < .05$. Similarly, the second Kruskal-Wallis test showed a significant statistical difference in the participants' passive collocational knowledge three weeks after the treatment/ no treatment conditions (GP1, *MD*= 27), (GP2, *MD*= 29), (GP3, *MD*= 27), and the control group (CG, *MD*= 18), $\chi^2 (N= 129) = 67.33, p = .000 < .05$. Group 2 (+DDL +CAT) recorded a higher median score than the other two experimental groups, and expectedly higher than the control group in both testing phases.

5.5.2.2 Participants' performance on post-tests of active collocational knowledge

Before investigating the statistical differences between the groups, it is essential to present informative and summarized sets of the numerical data gathered in the post and delayed post-tests. Table (5.41) below shows descriptive statistics of the data.

Table 5. 41: Descriptive statistics of the post-tests of active recall of all groups

Parameter	Post-active	Delayed-active
	Max. = 30	Max. = 30
group 1 (N= 33)	19.85 (21.00) <i>SD</i> 5.263	17.03 (18.00) <i>SD</i> 5.676
group 2 (N= 32)	22.72 (23.00) <i>SD</i> 3.621	20.63 (20.50) <i>SD</i> 4.361
group 3 (N= 32)	16.34 (15.00) <i>SD</i> 5.277	16.63 (13.00) <i>SD</i> 4.361
Control group (N= 32)	04.59 (4.00) <i>SD</i> 3.555	4.31 (4.00) <i>SD</i> 3.459

The table above shows that the participants in experimental group 1, 2 and 3 have made considerable progress in both post- testing phases for active knowledge of the target non-congruent collocations in comparison to their knowledge at the entry level and in comparison to the control group. However, the mean scores of experimental group 2 are higher than those of experimental group 1 and experimental group 3 in the post-tests and in the delayed post-tests. The control group, on the other hand, have hardly made any progress in both testing phases.

To compare the scores of active knowledge of the target collocations achieved by the participants in the post-tests and delayed post-tests across all groups, two Kruskal-Wallis tests were run.⁶¹

⁶¹ Instead of ANOVA, a Kruskal-Wallis test was run to compare the delayed post-tests scores for active knowledge due to violation of the homogeneity of variance assumption i.e. the significance value for Levene's test was not greater than .05.

Table 5. 42: Post-tests of the active collocational knowledge between all groups

Parameter	Groups/ treatments	Chi-square	df	<i>p</i> . value
Post-active	Exp. group 1	81.534	3	.000
	Exp. group 2			
	Exp. group 3			
	Control group			
Delayed-active	Exp. group 1	80.228	3	.000
	Exp. group 2			
	Exp. group 3			
	Control group			

Both statistical tests revealed significant differences in the participants' active knowledge of the target non-congruent collocations across the four groups. In the post-testing phase, the three experimental groups achieved greater median values than the control group (as shown in table 5.41), thus reaching a statistically significant alpha value $\chi^2 (N= 129) = 81.534, p < .05$ (table 5.42). Likewise, in the delayed post-testing phase the difference across the median values of the three experimental groups and the control group reached a statistically significant level $\chi^2 (N= 129) = 80.228, p < .05$.

Obtaining statistically significant results from a Kruskal-Wallis test, though holistic, does not allow the current researcher to determine which of the groups are statistically significantly different from one another. Therefore, different post-hoc tests were performed to look in more detail at statistical differences and for patterns that had not been identified a priori. This investigation took two approaches discussed in the following sections; the effect of CAT and the effect of DDL.

5.5.3 Effect of CAT

Two approaches were exploited in teaching the target non-congruent collocations; a corpus-based comparative analysis and translation approach (group 2) and a more traditional teacher-fronted non-corpus-based comparative analysis and translation approach (group 1). In this section, the current researcher aims at looking into the effectiveness of the two approaches at both passive and active levels of collocational knowledge, first in comparison to each other, then in comparison to a control group which received no treatment.

5.5.3.1. Effect of (-DDL +CAT VS. +DDL +CAT) on passive knowledge of collocations

An independent-samples t-test was carried out to compare the scores of collocational passive knowledge attained by the participants in each group prior to the treatments and to double check the non-significant statistical difference found between all groups (see section 4.4.1). The results in the table below confirm that there was no significant statistical difference between the -DDL +CAT condition (G1, $M= 16.03$, $SD= 3.459$) and the +DDL +CAT condition (G2, $M= 16.66$, $SD= 2.209$), $t (.866)$, $p= .390 > .05$. The non-existence of significant statistical differences between the two groups was confirmed by the very small difference size in the mean scores ($r= .01$).

Table 5. 43: Pre-tests of passive collocational knowledge (group 1 & 2)

Parameter	Groups/ treatments	<i>t</i>	df	Sig. (2-tailed)	<i>r</i>
Pre-passive	Exp. group 1	.866	63	.390	.01
	Exp. group 2				

When the same test for passive collocational knowledge was administered immediately after the treatments in each group, different scores were observed. A Kruskal-Wallis test revealed a statistically significant difference in the students’ passive knowledge of the target non-congruent collocations across the two experimental groups and the control group (group 1, $N= 33$, group 2= 32, group 3= 32, control group $N=32$), $\chi^2 (2, N= 97) = 60.317$, $p < .05$. Group 2 (+DDL +CAT) recorded a higher median score ($MD= 29$) than Group 1 (-DDL +CAT) which recorded a median value of 28. Both experimental groups 1 and 2 recorded higher median scores than the control group ($MD= 18$). However, one still cannot determine which of the groups are statistically significantly different, or what the actual size of any existing statistical difference is. Therefore, some follow-up Mann-Whitney U tests were run to compare between pairs of groups.

Table 5. 44: Immediate post-tests of passive collocational knowledge (group 1&2)

Parameter	Groups/ treatments	<i>z</i>	<i>U</i>	Sig. (2-tailed)	<i>r</i>
Post-passive	Exp. group 1	2.363	352.500	.018	.3
	Exp. group 2				

A Mann-Whitney U test was run to compare between the post-test scores of participants in experimental group 1 and 2 and revealed a significant statistical difference in passive knowledge scores between the two groups (GP 2, $MD = 29$, $SD= 1.508$), (GP 1, $MD = 28$, $SD= 3.211$), $p= .01 < .025$. However, the magnitude of this difference was found to be medium ($r= .3$).

Table 5. 45: Immediate post-tests of passive collocational knowledge (group 1& control group/ 2& control group)

Parameter	Groups/ treatments	z	U	Sig. (2-tailed)	r
Post-passive	Exp. group 1 Control group	-6.187	58.000	.000	.8
Post-passive	Exp. group 2 Control group	-6.855	5.000	.000	.9

More statistical Mann-Whitney U tests were run eventually to compare each of the experimental groups 1 and 2 to the control group. Thus, the effect of each treatment versus no treatment on passive knowledge of the target non-congruent collocations was examined. The two statistical tests revealed that there were indeed significant statistical differences in the post-test results between each experimental group in comparison to the control group (both $P < .025$). The magnitude of the differences was also found to be significantly large in both statistical tests (GP1 vs. CG, $r = .8$) (GP2 vs. CG, $r = .9$).

The significant statistical difference in the median values of passive collocational knowledge achieved by the participants in experimental group 1, experimental group 2 and the control group was retained in a delayed post-test three weeks after the treatments of each group. Another Kruskal-Wallis test comparing the delayed post-test scores across the three groups showed a statistically significant difference in the students' passive knowledge scores of target non-congruent collocations across the three groups $\chi^2 (2, N= 97) = 57.978$, $p < .05$. Once again Group 2 (+DDL +CAT) recorded the highest median score ($MD= 29$) compared to Group 1 (-DDL +CAT), which recorded a median value of ($MD= 27$) and to the control group ($MD= 18$). This was followed by a post-hoc Mann-Whitney U test to compare the scores of the two experimental groups 1 and 2.

Table 5. 46: Delayed post-tests of passive collocational knowledge (group 1 & 2)

Parameter	Groups/ treatments	<i>z</i>	<i>U</i>	Sig. (2-tailed)	<i>r</i>
Delayed-passive	Exp. group 1	2.316	354.000	.021	.3
	Exp. group 2				

The comparison between the participants' scores in the delayed post-test also showed that the difference in the median values of the two groups had reached a significant statistical level (GP2, *MD*= 29, *SD* = 1.727), (GP1, *MD*= 27, *SD*= 3. 824), $p = .021 < .025$. Interestingly, the magnitude of this statistical difference was similar to that found between the two groups in the post-tests. It was found to be still within the medium range ($r = .3$). This possibly indicates that passive knowledge of the non-congruent target collocations was better retained by participants in group 2 under (+DDL +CAT) treatment.

Table 5. 47: Delayed post-tests of passive collocational knowledge (group 1& control group/ 2& control group)

Parameter	Groups/ treatments	<i>z</i>	<i>U</i>	Sig. (2-tailed)	<i>r</i>
Delayed-passive	Exp. group 1	-6.044	68.500	.000	.7
	Control group				
Delayed-passive	Exp. group 2	-6.749	11.500	.000	.8
	Control group				

Additionally, a comparison of the delayed post-test score of experimental group 1 (GP1, *MD*= 27, *SD*= 3.374) and experimental group 2 (GP2, *MD*= 29, *SD*= 1.727) with the control group (CP, *MD*= 18, *SD*= 3.892) revealed that there were also significant statistical differences in the median values of these groups compared to the control group (GP1, $z = -6.044$ / GP2, $z = -6.749$), (both $p < .025$), with the actual size of the difference being large for both comparisons (GP1 vs. CG, $r = .7$) (GP2 vs. CG, $r = .8$).

5.5.3.2. Effect of (-DDL +CAT VS. +DDL -CAT) treatments on collocational active knowledge

Similar to the results shown previously in the all group comparison, a Mann-Whitney test also showed that there was no significant statistical difference between experimental group 1 ($MD= 2, N= 33$) and experimental group 2 ($MD= 2, N= 32$), $p > .025$. The extremely small calculated effect size shows the near non-existence of differences between the two groups' scores for active knowledge of the target non-congruent collocations prior to the treatments.

Table 5. 48: Pre-tests of active collocational knowledge (group 1 & 2)

Parameter	Groups/ treatments	z	U	Sig. (2-tailed)	r
Pre-active	Exp. group 1	-.073	522.500	.942	.009
	Exp. group 2				

To establish a comparison between the two experimental groups and the control group for the participants' performance in the post-treatment tests for active knowledge, a Kruskal-Wallis test was run. The test revealed a statistically significant difference across the three groups $\chi^2 (2, N= 97) = 64.514, p < .025$. However, follow-up tests were required to examine the differences between each pair of groups.

Table 5. 49: Immediate post-tests of active collocational knowledge (group 1 & 2)

Parameter	Groups/ treatments	z	U	Sig. (2-tailed)	r
Post-active	Exp. group 1	-2.159	364.000	.031	.3
	Exp. group 2				

The table above presents the results of a Mann-Whitney test. It shows a significant statistical difference in the active knowledge test scores after the treatments between group 1 ($MD= 21, SD= 5.263$) and group 2 ($MD= 23, SD= 3.612$), $z= -2.159, p= .31 < .05$. The magnitude of this difference is medium ($r= .3$).

Table 5. 50: Immediate post-tests of active collocational knowledge (group 1& control group/ 2& control group)

Parameter	Groups/ treatments	<i>z</i>	<i>U</i>	Sig. (2-tailed)	<i>r</i>
Post-active	Exp. group 1	-6.768	13.000	.000	.8
	Control group				
Post-active	Exp. group 2	-6.886	.000	.000	.9
	Control group				

Two more Mann-Whitney tests were run to examine the effect of each CAT treatment in comparison to no-treatment in the control group in a post-testing phase. The first test was run between experimental group 1 and the control group and showed very significant statistical differences in the participants' active collocational knowledge in the post-test (GP1, *MD*= 21, *SD*= 5.263), (CP, *MD*= 4, *SD*= 3.555), $z = -6.768$, $p < .05$. The magnitude of the difference between the two groups was very large ($r = .8$). Similarly, when the scores of the post-test for active knowledge were compared between experimental group 2 and the control group, significant statistical differences were found (GP 2, *MD*= 23, *SD*= 3.612), (CP, *MD*= 4, *SD*= 3.555), $z = -6.886$, $p < .05$. The size of this difference was also found to be large ($r = .8$).

A Kruskal-Wallis test⁶² was run to check whether the statistical differences between the three groups were maintained in the delayed post-test. The test revealed that there was a statistically significant difference at the $p < .05$ level in active collocational knowledge for the three groups, $\chi^2 (2, N = 97) = 62.419$, $p = .000$. Group 2 attained the highest median score among the three groups (GP2, *MD*= 20.50), (GP 1, *MD*= 18) with the control group attaining the lowest median value (CP, *MD*= 4). The comparative results obtained from the Kruskal-Wallis test regarding the differences between the three groups could be described as holistic; however, more detailed results were needed to examine the differences between the pair of groups. Hence, an independent-samples t-test was run to compare between the delayed post-test scores of group 1 and 2.

⁶² Despite the data of the delayed tests for the three groups being normally distributed, a Kruskal-Wallis test was used because of the violation of the homogeneity of variance assumption.

Table 5. 51: Delayed post-tests of active collocational knowledge (group 1 & 2)

Parameter	Groups/ treatments	<i>t</i>	df	Sig. (2-tailed)	<i>r</i>
Delayed-active	Exp. group 1	-2.857	63	.006	.11
	Exp. group 2				

The test revealed that despite reaching a statistically significant difference $t = -2.857$, $p < .025$, the actual difference in mean scores between experimental group 1 and experimental group 2 was medium. The effect size, calculated using eta squared, was (.11).⁶³

Table 5. 52: Delayed post-tests of active collocational knowledge (group 1& control group/ 2& control group)

Parameter	Groups/ treatments	<i>t</i>	df	Sig. (2-tailed)	<i>r</i>
Delayed-active	Exp. group 1	10.945	53.155	.000	.7
	Control group				
Delayed-active	Exp. group 2	16.579	62	.000	.8
	Control group				

The difference in the delayed post-test scores for active collocational knowledge is statistically significant in the comparison run between the control group (CG, $M = 4.31$, $SD = 3.459$) and both experimental group 1 (GP1, $M = 17.03$, $SD = 5.676$), $t = 10.94$ and experimental group 2 (GP2, $MD = 20$, $SD = 4.361$), $t = 16.579$, (both $p < .025$). The size of the difference between each experimental group and the control group was found to be very large ($r = .7 / .8$).

5.5.4 Effect of DDL

Two types of DDL approaches were used in teaching the target non-congruent collocations. The first was a corpus-based approach which involved bilingual corpus (English/ Arabic) consultations and contrastive analysis and translation tasks (group 2). The second was a corpus-based approach which involved monolingual corpus (English) consultations with fill-in-the-blank and MC tasks (group 3). This section will be allocated to the presentation of the comparison of results between the two approaches in order to make claims about the most and the least effective treatment among the two. Each

⁶³ This is according to Cohen's (1988) criteria of .01= small effect, .06= medium effect, .14= large effect.

teaching approach/ treatment will be then compared to the control group which received no treatment.

5.5.4.1 Effect of (+DDL +CAT) VS. (+DDL -CAT) on passive collocational knowledge

An independent-sample t-test was run to compare the mean scores achieved in the pre-test for passive knowledge of the target non-congruent collocations by experimental group 2 and experimental group 3.

Table 5. 53: Pre-tests of passive collocational knowledge (group 2 & 3)

Parameter	Groups/ treatments	<i>t</i>	df	Sig. (2-tailed)	<i>r</i>
pre-passive	Exp. group 2	-.371	62	.712	.002
	Exp. group 3				

The test confirms that there was no significant statistical difference in the mean scores between the -DDL +CAT condition (GP2, $M= 16.66$, $SD= 2.209$) and the +DDL +CAT condition (GP3, $M= 16.88$, $SD= 2.498$); $t(64) = .371$, $p=.712 > .025$. The difference in the mean scores between the two groups had an insignificant and extremely small size ($r= .002$).

The same test was administered to all participants immediately after the last treatment in each group. To investigate the changes in the mean scores achieved primarily as a result of treatment/ no treatment effects, a Kruskal-Wallis test was run. The test revealed significant statistical differences in median values across the three groups $\chi^2(2, N= 96) = 60.227$, $p < .05$. The median value of experimental group 2 was the highest (GP2, $MD= 29$, $SD= 1.508$) of the three (GP3, $MD= 28$, $SD= 3.473$), (CG, $MD= 18$, $SD= 3.991$). This test was followed by a series of post-hoc statistical tests to examine the differences between each pair of groups.

Table 5. 54: Immediate post-tests of passive collocational knowledge (group 2 & 3)

Parameter	Groups/ treatments	<i>z</i>	<i>U</i>	Sig. (2-tailed)	<i>r</i>
Post-passive	Exp. group 2	-6.768	13.000	.000	.8
	Exp. group 3				

A Mann-Whitney U test (see table above) revealed significant statistical differences when the median values of the scores of the passive collocational tests of experimental group 2 ($MD= 29$) and 3 ($MD= 28$) were compared $z= -6.768$, $p < .025$. The actual magnitude of this difference was found to be very large ($r= .8$).

Table 5. 55: Immediate post-tests of passive collocational knowledge (group 2& control group/ 3& control group)

Parameter	Groups/ treatments	z	U	Sig. (2-tailed)	r
Post-passive	Exp. group 2	-6.855	5.000	.000	.9
	Control group				
Post-passive	Exp. group 3	-6.036	64.000	.000	.8
	Control group				

Two Mann-Whitney U tests revealed significant statistical differences in the passive knowledge of the target non-congruent collocations of the control group in comparison to each of the two experimental groups (GP2, $z= -6.855$; GP3, $z= -6.036$), $p < .025$ for both groups. The differences between the two experimental groups and the control group were not only statistically significant, but also significant in their sizes (GP2 vs. CG, $r= .9$ / GP3 vs. CG, $r= .8$).

As revealed by a Kruskal-Wallis test, the differences in scores between the two experimental groups and the control group continue to be statistically significant in the delayed post-testing phase $\chi^2 (2, N= 96) = 58.292$, $p < .05$. The three groups attained median values of (GP2, $MD= 29$, $SD= 1.727$), (GP3, $MD= 27$, $SD= 3.354$) and (CG, $MD= 4$, $SD= 3.862$), with experimental group 2 achieving once again the highest median value of the three groups. However, more tests were needed to determine how significantly or insignificantly different each pair of groups were.

Table 5. 56: Delayed post-tests of passive collocational knowledge (group 2 & 3)

Parameter	Groups/ treatments	z	U	Sig. (2-tailed)	r
Delayed- passive	Exp. group 2	-6.768	13.000	.000	.8
	Exp. group 3				

When a Mann-Whitney test was carried out, it revealed a significant statistical difference in the scores of passive knowledge of the target collocations between experimental groups 2 ($MD= 29$) and 3 ($MD= 27$), $z= -6.768$, $p< .025$. Similar to the post-test, the difference between the groups in the scores of passive knowledge of the target collocations was large ($r=.8$).

Table 5. 57: Delayed post-tests of passive collocational knowledge (group 2&control group/ 3&control group)

Parameter	Groups/ treatments	z	U	Sig. (2-tailed)	r
Delayed- passive	Exp. group 2	-6.749	11.500	.000	.8
	Control group				
Delayed- passive	Exp. group 3	-5.832	79.000	.000	.7
	Control group				

The two statistical tests run to compare each experimental group's scores with those of the control group expectedly revealed significant statistical differences (GP2, $z= -6.749$; GP3, $z= -5.832$). The magnitude of the difference in both comparisons was very large (GP2, $r= .8$; GP3, $r= .7$).

5.5.4.2 Effect of (+DDL +CAT VS. +DDL -CAT) on active collocational knowledge

A Mann-Whitney U test showed that there was no significant statistical difference between the -DDL +CAT condition (GP2, $MD= 2$, $N= 33$) and the -DDL +CAT condition (GP3, $MD= 1.5$, $N= 32$), $p= .875 > .025$. This is also evident in the very small effect size which shows the near non-existence of differences between the two groups' scores for active knowledge of the target non-congruent collocation before the treatments.

Table 5. 58: Pre-tests of active collocational knowledge (group 2 & 3)

Parameter	Groups/ treatments	z	U	Sig.(2-tailed)	r
Pre-active	Exp. group 2	-.157	500.500	.875	.01
	Exp. group 3				

To compare between the two experimental groups and the control group for the participants' performance in the post-treatment tests for active knowledge, a Kruskal-Wallis test was carried out. The test provided evidence of statistically significant

differences between the three groups $\chi^2(2, N=97) = 70.995, p < .05$. Nonetheless, follow-up tests were required to examine the differences between each pair of groups.

Table 5. 59: Immediate post-tests of active collocational knowledge (group 2 & 3)

Parameter	Groups/ treatments	<i>t</i>	df	Sig.(2-tailed)	<i>r</i>
Post-active	Exp. group 2	5.639	54.821	.000	.33
	Exp. group 3				

The table above presents the results of a paired-samples t-test. It shows a significant statistical difference in the active knowledge test scores after the treatments between group 2 ($M= 22.72, SD= 3.612$) and group 3 ($M= 16.34, SD= 5.277$), $t= 5.639, p < .025$. The magnitude of this difference was found to be very large ($r= .3$).

Table 5. 60: Immediate post-tests of active collocational knowledge (group 2& control group/ 3& control group)

Parameter	Groups/ treatments	<i>z</i>	<i>U</i>	Sig.(2-tailed)	<i>r</i>
Post-active	Exp. group 2	-6.886	.000	.000	.9
	Control group				
Post-active	Exp. group 3	-6.598	21.500	.000	.8
	Control group				

To examine the effect of each DDL treatment in comparison to no-treatment in the control group in a post testing phase, two more Mann-Whitney tests were run. The first test run between experimental group 2 and the control group showed very significant statistical differences in the participants' active collocational knowledge in the post-test (GP2, $MD= 23, SD= 3.612$), (CP, $MD= 4, SD= 3.555$), $z= -6.886, p < .05$. The magnitude of the difference between the two groups was very large ($r= .8$). Similarly, when the scores of the post-test for active knowledge were compared between experimental group 2 and the control group, significant statistical differences were found (GP 3, $MD= 15.50, SD= 5.277$), (CP, $MD= 4, SD= 3.555$), $z= -6.598, p < .05$. The size of this difference was also found to be large ($r= .8$).

A one-way ANOVA was run to check whether the statistical differences between the three groups were maintained in the delayed post-test. The test revealed that there was a

statistically significant difference at the $p < .05$ level in active collocational knowledge for the three groups, $F(2, N=97) = 128.593, p = .000$. Group 2 attained the highest mean score among the three groups (GP2, $M = 20.63$), (GP3, $M = 13.63$) with the control group attaining the lowest mean value (CP, $M = 4.31$). The actual difference in mean scores between the groups was very large ($r = .26$).⁶⁴ Having found significant statistical differences between the three groups, the current researcher carried out post-hoc comparisons using the Tukey HSD test. It yielded results as follows.

Table 5. 61: Delayed post-tests of active collocational knowledge (group 1, 2& control group)

Parameter	Groups/ treatments	Groups/ treatments	Mean difference	<i>p</i> . value
Delayed-active	Exp. group 2	Exp. group 3	7.000	.000
		Control group	16.313	.000
Delayed-active	Exp. group 3	Exp. group 2	7.000	.000
		Control group	9.313	.000
Delayed-active	Control group	Exp. group 2	16.313	.000
		Exp. group 3	9.313	.000

The Tukey HSD test indicated that the mean score for both experimental groups (GP2, $M = 20.63, SD = 4.361$, GP3, $M = 13.63, SD = 4.361$) were significantly different from the control group (CG, $M = 4.31, SD = 3.459$). Most importantly, the mean score of experimental group 2 also differed significantly from experimental group 3. It is worth mentioning here that independent-sample t-tests were also run just for the purpose of providing the necessary data to calculate the magnitude of the statistical differences found between the pair of groups. The size of the difference between the mean scores of experimental group 2 and experimental group 3 was found to be large ($r = .40$). Similarly, the size of the differences in mean scores between the control group and experimental group 2 and group 3 were large (GP2, $r = .8$; GP3, $r = .6$).

5.6 Summary

The second part of this chapter showed between-group comparisons. Statistical tests (i.e. ANOVA/ Kruskal-Wallis/ Mann-Whitney/ independent-samples t-test) were used for that purpose. Initially, the current researcher carried out comprehensive tests comparing all three experimental groups and the control group to investigate whether or not there were

⁶⁴ This is according to Cohen's (1988) criteria of .01= small effect, .06= medium effect, .14= large effect.

any statistical differences between the groups prior to the treatments. The results obtained from the statistical tests revealed no statistical differences between the four groups in their passive and active knowledge of the target non-congruent collocations at the outset of the study. After establishing the no-difference baseline for comparison, the results of the post-tests and delayed post-test scores were compared across the four groups. The findings were as follows.

- There were significant statistical differences between all four groups in the post-testing phase for passive knowledge of the target non-congruent collocations.
- There were significant statistical differences between all four groups in the post-testing phase for active knowledge of the target non-congruent collocations.
- The significant statistical differences were retained in the delayed post-tests for passive knowledge of the target non-congruent collocations.
- The significant statistical differences were also retained in the delayed post-tests for active knowledge of the target non-congruent collocations.

The next step was then to determine which group is significantly different from the others as a result of a given treatment, and to calculate the size of any existing differences. Therefore, the current researcher decided to conduct further statistical analysis in two main categories: the effect of CAT and the effect of DDL. The effect of CAT confirms the first hypothesis. The pre, post and delayed-tests scores of both passive and active knowledge of the target non-congruent collocations obtained by the participants in experimental group 1 (-DDL +CAT) were compared to those of experimental group 2 (+DDL +CAT). Each of the two groups was then compared to the control group. The results of the statistical test were as follows:

- There were no significant statistical differences between participants' scores in experimental group 1 and experimental group 2 in the pre-tests for passive and active knowledge of the target non-congruent collocations. The difference size was also insignificant.
- There were significant statistical differences between participants' scores in experimental group 1 and experimental group 2 in the post-tests for passive and active knowledge of the target non-congruent collocations. The difference size was found to be medium.
- There were also significant statistical differences between participants' scores in experimental group 1 and experimental group 2 in the delayed post-tests for

passive and active knowledge of the target non-congruent collocations. The difference size remained within the medium range.

- There were significant statistical differences between participants' scores in experimental group 1 and experimental group 2 compared to those of the control group in the post and delayed post-tests for passive and active knowledge of the target non-congruent collocations. The difference size was found to be very large in each comparison.

The effect of DDL confirms the second hypothesis. The pre, post and delayed-test scores of both passive and active knowledge of the target non-congruent collocations obtained by the participants in experimental group 2 (+DDL +CAT) were compared to those of experimental group 3 (+DDL -CAT). Each of the two groups was then compared to the control group. The results were the following:

- There were no significant statistical differences between participants' scores in experimental group 2 and experimental group 3 in the pre-tests for passive and active knowledge of the target non-congruent collocations. The difference size was extremely small.
- There were significant statistical differences between participants' scores in experimental group 2 and experimental group 3 in the post-tests for passive and active knowledge of the target non-congruent collocations. The difference size was found to be significantly large.
- There were also significant statistical differences between participants' scores in experimental group 2 and experimental group 3 in the delayed post-tests for passive and active knowledge of the target non-congruent collocations. The difference size was within the medium size range.
- There were significant statistical differences between participants' scores in experimental group 2 and experimental group 3 in comparison to those of the control group in the post and delayed post-tests for passive and active knowledge of the target non-congruent collocations. The difference size was found to be very large in each comparison.

This thorough analysis of the data gathered from the three experimental groups under (-DDL +CAT), (+DDL +CAT) and (+DDL -CAT) treatments have led to the following conclusions:

- The passive and active knowledge of the target non-congruent collocations was found to progress better under the (+DDL +CAT) treatment than under the (-DDL +CAT) treatment or the (+DDL -CAT) treatment.
- The medium magnitude of the difference between the two (+CAT) treatments may suggest that the participants have benefited from (+DDL +CAT) in comparison to (-DDL +CAT), but not overly so.
- The large magnitude of the difference between the two (+DDL) treatments may suggest that the participants have greatly benefited from (+DDL +CAT) in comparison to (+DDL -CAT).
- Although no direct comparison was made between the (-DDL +CAT) treatment and the (+DDL -CAT) treatment, the magnitude of the significant statistical difference found between each of these treatments and the (+DDL +CAT) treatment could imply that (+CAT) treatments have led to better acquisition of the target non-congruent collocations than (-CAT).

The next chapter will be dedicated to explaining and discussing the results in light of the ‘noticing’, ‘pushed output’, and ‘task-induced involvement load’ hypotheses, and the influence that L1 exerts on the acquisition of L2 vocabulary.

Chapter 6: Discussion

This study investigated whether incorporating a corpus-assisted contrastive analysis and translation approach (+DDL +CAT) into an EFL classroom context would make a significant difference in learning non-congruent adjective/ noun collocations in comparison to a non-corpus-assisted contrastive analysis and translation approach (-DDL +CAT), and a corpus-assisted non-contrastive analysis and translation approach (+DDL -CAT). The results revealed that the (+DDL +CAT) group scored significantly higher than the two other groups on all four tests, both on the two identical immediate tests (passive recall of collocations, active recall of collocations) and on the two identical delayed post-tests (passive recall of collocations, active recall of collocations). The group that did not receive any treatment learned significantly fewer collocations. As proposed earlier in chapter 3, the effectiveness of corpus-assisted cross-linguistic form-focused instruction could be explained by the hypotheses of ‘noticing’, ‘task-induced involvement load’, ‘pushed output’, and could further be supported by findings that show the pervasive influence of L1 in vocabulary acquisition and processing. Hence, this chapter will discuss the results from a theoretical perspective and in relation to previous research where possible. The chapter will be divided into two sections: the first section is a general discussion of the participants’ knowledge of collocations, while the second one is structured according to the research hypotheses and sub-hypotheses.

6.1 Section one: participants’ knowledge of collocations

This section comprises general findings regarding the learners’ performance on the passive recall test (English/ Arabic translation) and active recall test (Arabic/ English translation) of collocations prior to and after the treatments.

6.1.1 Pre-treatments knowledge of collocations

Before discussing actual learning gains, it needs to be noted that the participants’ passive knowledge of the target collocations in all groups was expectedly high even before they were subject to any teaching condition, as shown in the pre-test results. The approximate mean percentage of the scores attained by G1 (-DDL +CAT), G2 (+DDL +CAT), G3 (+DDL -CAT) and the control group were 53, 55, 56 and 56 percent respectively.

The high scores could be attributed to several possible causes. Firstly, the adjective and noun constituents of each target collocation as well as the target collocations they made up were highly frequent (see chapter 4, section 4.1.1 for details). Since it is acknowledged that frequency plays an important role in vocabulary learning (Schmitt 2010), it may not

have been difficult for the learners to understand the meaning of such highly frequent collocations, and to provide their meaning in their mother tongue in an appropriate way. Secondly, given the mean scores achieved in the VLT (K2 and K3), the participants are likely to have the vocabulary size to understand the constituent words of the target collocations which belong to the (K1) and (K2) levels. Thirdly, many collocations could be easily comprehended due to their semantic transparency (Laufer & Girsai, 2008a, b; Laufer & Waldman, 2011; Peters, 2012). For example, Peters (2015, p. 2) states that “Dutch-speaking learners of English will easily understand the semantically transparent but incongruent collocation *to make an effort*. However, they might not notice that English has the verb *to make* in this collocation and might use *to do an effort* because they rely on their L1. Thus, the particular difficulty associated with collocational knowledge lies mainly in their production rather than their reception, as evident in the literature (see chapter 2, section 2.7 for details). The fourth possible explanation, which relates to the previous three, is the nature of the designed passive recall tests. The target collocations were contextualised i.e. provided in a full sentence, which may have facilitated the learners’ comprehension through guessing the meaning from the context. Another likely explanation is that the target collocations might either have been part of the participants’ potential vocabulary (i.e. the FL words that the learners have not come across before, but which they could nonetheless understand when first encountered), or part of their passive real vocabulary (i.e. the FL words that the learners have learned/encountered at some point in the learning process, and which they can understand) (cf. Berman et al., 1968, cited in Palmberg, 1987).

In a striking contrast to the high scores of collocational passive knowledge, the participants’ active knowledge of the target non-congruent collocations was very low in all four groups as indicated by the results of the active recall pre-test. The approximate mean percentage of the scores attained by G1 (-DDL +CAT), G2 (+DDL +CAT), G3 (+DDL -CAT) and the control group were 9, 9, 9 and 8 percent respectively.

Despite the fact that the main aim of the active-recall pre-test was mainly to elicit the participants’ knowledge of the target non-congruent collocations, rather than to assess their overall collocational knowledge, the results can still shed some light on the learners’ problems in producing collocations. This applies especially those collocations which do not have literal or commonly used equivalents in the students’ mother tongue (Arabic). The low scores achieved by the participants in the active recall pre-test, as opposed to the passive recall pre-test, are very much in line with the research on EFL elicited productive

collocational knowledge (e.g. Brashi, 2009; Farghal & Obiedat, 1995; Noor & Adubaib, 2011). Bahns and Eldaw (1993) also found that their EFL learners' collocational knowledge fell behind that of general vocabulary. The discrepancy between the scores achieved in passive and active knowledge of collocations indicates a gap between the learners' knowledge of collocations and that of meanings and forms of individual words.

As noted in chapter 2 (section 2.4), word knowledge is multidimensional and entails much more than word meanings alone. Simple comprehension of form-meaning links does not mean 'knowing' a word, and much less being able to use the word, which entails e.g. being able to provide its collocates (Nation, 2001; Palmberg, 1986; Richards, 1976). It has been suggested (e.g. Laufer, 1997; Schmitt, 1998) that different aspects of word knowledge may not develop at the same rate, and that using words productively is one aspect of word knowledge that may not be considered as an easy task (Nation, 2001). Such a gap in the learners' knowledge may stem from multiple aspects which have plagued the EFL context: (1) the poverty of input that allows for the multiple encounters necessary for incidental learning of collocations; (2) the misconception of teaching and learning words as discrete units and overlooking their syntagmatic relations; (3) the negligence of direct teaching of collocations; (4) lexical transfer and L1 influence. Consequently, these aspects affect the way in which collocations are processed, represented in the bilingual mental lexicon and eventually produced.

In relation to the first aspect, Webb et al. (2013) note that "despite the frequency of the individual items that make up collocations, most collocations do not occur very often; they are always less frequent than the most frequent word within the collocations" (p. 94). They also suggest that even when a particular collocation is encountered, there would be greater intervals between the encounters. The lower frequency of encounters with collocations as opposed to their constituents increases the likelihood that knowledge obtained through each encounter might be forgotten and that knowledge of the parts is likely to be greater than that of the collocations (Bahns & Eldaw, 1993). As suggested above, this lack of sufficient collocation encounters can be attributed to poverty of input in terms of both quality and quantity in the SL/ FL learning classrooms (Jiang, 2000). The poverty of input makes it extremely hard for L2 learners to extract and create lemmatic specifications about a word (e.g. collocates), and to integrate this information with the word's other specifications, as suggested by Jiang (2000).

Teaching and learning words as discrete units and neglecting direct teaching of collocations is common practice in the Saudi EFL context, as suggested by Al-Sugayyer

(2006) and Alhawsawi (2013), and as indicated in the small-scale exploratory study conducted at the outset of this research. A major consequence of these aspects is the likelihood that learners employ what Sinclair (1987, 1991) refers to as the ‘open-choice principle’ (see chapter 2 section 2.6 for details). The ‘open-choice principle’ and the ‘idiom principle’ were mainly intended to explain how meaning is conveyed in texts (as described in chapter 2, section 2.6). However, these principles are not only a model for interpretation, but also a model for language production. Sinclair (1987, 1991) argues that normal text (collocations in this case) would not be produced simply by operating the ‘open-choice principle’ as collocations do not occur at random, thus, the idiom principle is needed. This failure of the learners to produce the target collocations despite their high frequency, and the possible application of the ‘open-choice principle’, might indicate weak or non-existent associative links between the constituent words in the learners’ mental lexicon. This is because frequent collocation is taken to indicate the presence of “semi pre-constructed phrases that constitute single choices” for the language user (Sinclair, 1987 p. 320), or the presence of “a psychological association between words” (Hoey, 2005 p. 5). Accordingly, it is these “semi-preconstructed phrases” or “psychological associations between words” which second-language learners need to acquire appropriately if they are to attain full mastery of collocations’ use (Durrant & Schmitt, 2009).

Lexical transfer, particularly negative transfer, and L1 influence might also be a significant contribution in producing erroneous or non-native-like collocations by the participants in this research. There is a heavy emphasis on L1 influence in almost all of the reviewed research on EFL collocational knowledge that used different investigation methods (i.e. corpus based, elicitation tests, and psycholinguistic measures, see chapter 2, section 2.7). Moreover, Jiang (2000, 2002, 2004) argues that L2 learners may rely on the L1 system to learn and process new lexis in L2 due to the presence of the established L1 lexical system. According to Jiang’s (2000) psycholinguistic model of vocabulary acquisition, when less advanced EFL learners, like those in this research, learn lexis, they are believed to experience a unique process of form–meaning mapping i.e. ‘L1 lemma copying’. The resulting lexical use is then called ‘L1 lemma mediation’. If this model is to be taken into consideration, the process or stage of lemma mediation would at least partially explain the erroneous collocation produced by the learners. In the case of non-congruent collocations (which do not have exact translation equivalents), L2 collocation

processing through L1 lemma information mediation would be a case of unsuccessful negative transfer that results in erroneous combinations.

6.1.2 Post-treatments knowledge of collocations

Despite all the potential reasons for the high scores attained in the passive-recall test, the current researcher would still argue that the participants' knowledge of the target collocations has developed under all FonFs treatments. The rigorous analysis of the participants' scores attained within each experimental group showed significant increase in their passive/ receptive knowledge of collocations after the treatments as opposed to their entry level. This analysis of the scores of each individual group enabled the researcher to calculate the size of the statistical significance so as to claim that the teaching conditions work and work effectively, as suggested by Coe (2002). It also allowed for the claim that the statistical significance has theoretical and practical importance (Dornyei, 2007). The mean percentage of the scores attained by G1 (the non-corpus-assisted contrastive group, -DDL +CAT) on the passive recall immediate post-test was 90 percent and it was 88 percent on the delayed post-test. In both cases, the size of the increase was significantly large. In addition, the mean percentage attained by G2 (the corpus-assisted contrastive group, +DDL +CAT) on passive recall in the immediate and delayed post-tests were 96 and 94 percent respectively with a significantly large effect size as well. The size of the statistical significance was also large in G3 (the corpus-assisted non-contrastive group (+DDL -CAT) as the participants achieved mean percentages of 88 percent in the immediate and 86 percent in the delayed passive recall tests. Interestingly, the participants in the control group achieved mean percentages of 61 in the immediate post-test, which they then retained in the delayed test, thus showing a slight increase in their passive knowledge of the target collocations. The size of this increase is noticeable in comparison to the participants' entry level.

This effect of the three FonFs treatments is even more evident in the active recall of the target non-congruent collocations. Compared with the low level of the active collocational knowledge as indicated by the pre-test scores, participants in the three experimental groups recalled significantly larger numbers of the target collocations after the FonFs treatments as evidenced by their scores in the immediate post-tests. The mean percentage of the retained collocations by the learners within each group was 90 percent in group 1 (-DDL +CAT), 95 percent in group 2 (+DDL +CAT), 88 percent in group 3 (+DDL -CAT) and 15 percent in the control group. This high mean percentage was retained in the delayed post-tests three weeks after the FonFs treatments. It was 88 percent

in group 1, 94 percent in group 2, 86 percent in group 3, and 14 percent in the control group.

The positive effect of the three FonFs treatment on the participants' passive and active recall of the target non-congruent collocations, as opposed to the no-treatment condition, supports Laufer's (2005a, b) Planned Lexical Instruction (PLI) hypothesis. In this hypothesis, Laufer argues that the major source of L2 vocabulary knowledge is likely to be word-focused classroom instruction. While not rejecting the importance of reading and repeated exposure to vocabulary learning, researchers such as Webb & Kagimoto (2009) suggest that learning collocations incidentally may entail a rare occurrence and slow process due to the limited number of opportunities to encounter the same collocation twice. Additionally, Laufer (2001, 2003, 2005a,b, 2006, 2010) among many other researchers argues that the amounts of acquired words are usually greater in FFI conditions than non-FFI ones. This is evident in many comparative studies in relation to single words (e.g. Ellis & He, 1999; File & Adams, 2010; Knight, 1994; Laufer, 2000, 2003; Luppescu & Day, 1993; Paribakht & Wesche, 1997; Sonbul & Schmitt, 2010). However, this is also true in relation to learning collocations. In the relatively few empirical studies addressing the acquisition of collocations under purely FFI and under FFI and other non-FFI instruction (e.g. Laufer & Girsai, 2008a, b; Peters, 2014, 2015; Sonbul, 2012; Webb & Kagimoto, 2009, 2011; Szudarski, 2012), there is broad agreement that explicit vocabulary activities, in which collocations are the central focus of attention, seem to be an effective means of making initial form-meaning links in the mental lexicon.

The increase in the control group's passive and active collocational knowledge is worth acknowledging and explaining here. One justification of this increase could be attributed to general learning and the lengthy duration of the experiment. The target collocations are highly frequent and it is possible that learners were exposed to the collocations in the teacher's talk or in the course books that were used in their modules as this researcher could not control for such a factor. Another more likely reason for the results of the control group may perhaps be the exposure to the target items in the administered tests, which should not be neglected when interpreting the results. Such learning gains in the control group are not completely unexpected in comparison to similar empirical research. For example, in Szudarski's (2012) six-week study on acquisition of frequent delexical verb-noun collocations, the results of the control group under no-treatment condition

indicated that the learners' collocational knowledge improved significantly between the pre-test and the post-test on two productive tests and one receptive test.

Although these findings come as no surprise, they still add to and confirm the literature on the effect of FFI and specially FonFs on learning collocations. However, the particular investigation driving this research is what makes the difference i.e. which type of the three FonFs tasks would render better acquisition of non-congruent adjective/noun collocations. Accordingly, the second section of this chapter will present the effectiveness of these FonFs teaching conditions in learning the target non-congruent collocations passively and actively as compared to each other and as addressed in the research hypotheses.

6.2 Discussion of the research hypotheses and sub-hypotheses

This section will be allocated to discussing the findings of the main research hypotheses and the sub-hypotheses related to them

6.2.1 Hypothesis 1 and sub-hypotheses

The first hypothesis entails a comparison of two approaches to conducting contrastive analysis: corpus-assisted CAT and teacher-fronted CAT. As indicated in section (3.4.2.1), contrastive analysis in this research is defined in cognitive terms and entails explicit cross-linguistic comparison between the learners' L1 and L2. In the corpus-assisted CAT group (group 2), the contrastive analysis was carried autonomously and explicitly through translation tasks and meta-linguistic input i.e. parallel, juxtaposed English/ Arabic corpus data. On the other hand, group 1 received teacher-fronted CAT in which the contrastive analysis was implicitly initiated by the translation tasks and then explicitly carried out by the teacher. As already discussed in section 3.6, the first hypothesis and sub-hypotheses are thus as follows:

H1. The corpus-assisted CAT condition will lead to the learning of a significantly larger number (if any) of adj. /noun collocations than the non-corpus-assisted CAT condition.

- a) The corpus-assisted CAT condition will lead to the passive recall of a significantly larger number of adj. /noun collocations than the non-corpus assisted CAT condition.
- b) The corpus-assisted CAT condition will lead to the active recall of a significantly larger number of adj. /noun collocations than the non-corpus assisted CAT condition.

- c) The differences between the conditions in active and passive recall will be retained in a delayed post-test.

As detailed in chapter 5, results from this research have shown that the corpus-assisted contrastive analysis and translation approach does indeed lead to the learning of more non-congruent adjective/ noun collocations than the non-corpus assisted contrastive analysis and translation both receptively and productively. It also revealed that the corpus-assisted CAT led to retention of more collocations than the non-corpus assisted CAT.

Comparing these results with the results of other studies which have investigated FFI effectiveness in learning collocations is quite difficult. As suggested by Laufer and Girsai (2008a), different empirical studies utilise different numbers of target items, different numbers of form-focused activities and different ways to measure learning. The duration and the execution method of the intervention also differs across the different studies. Although the non-corpus CAT approach could, in one way or another, be considered as an adoption of Laufer and Girsai's (2008a, b) CAT approach, the results from this research could only relate to and confirm the overall effectiveness of CAT in learning non-congruent collocations. Due to the lengthy duration of the CAT intervention in this study compared to that in Laufer and Girsai's, the mean percentage of the retained active collocations are higher in both the immediate and delayed post-tests (80 and 79 percent) than Laufer and Girsai's (51 percent).⁶⁵ It is, however, difficult to compare the results of passive knowledge of collocations attained by the CAT groups to those of Laufer and Girsai's, since the researchers did not test their participants' pre-passive knowledge of collocations. Laufer and Girsai (2008a, b) argued that because many collocations are semantically transparent and thus easily comprehended, administering a test of passive knowledge of collocations was not necessary. Therefore, the percentage of gains in the passive collocation knowledge attained in the post-testing phases was not quite clear. Additionally, the researchers have only reported the statistically significant superiority of the CAT condition in comparison to the other conditions (i.e. non-contrastive FFI and MFI), but not the size of the significance, nor the size of increase within the CAT group. This adds to the difficulty of comparing the results attained in my research to those of Laufer and Girsai's.

⁶⁵ This is despite the fact that the current researcher did not control for zero knowledge of active collocational knowledge prior to the intervention.

6.2.2 Hypothesis two and sub-hypotheses

The second hypothesis compares the learning outcome under two DDL conditions i.e. the corpus-based CAT condition and a corpus-assisted condition that does not involve contrastive analysis and translation. Typically, research on collocation learning and teaching incorporating a corpus as a learning/ teaching tool makes use of various receptive and productive vocabulary learning tasks in the intervention stage (e.g. Chan & Liou, 2005; Sun & Wang, 2003; Tsai, 2011), but not translation tasks. The main aim of these studies was to examine the effectiveness of DDL as a pedagogical approach for collocation learning and teaching. They affirmed the usefulness and effectiveness of corpus consultation and the variety of tasks in learning collocations (including non-congruent ones in Chan & Liou, 2005). However, they did not investigate the effectiveness of the different types of tasks on the acquisition of the collocations. Thus, the inclusion of a corpus-assisted approach with no CAT condition was important in order to claim effectiveness of the proposed approach i.e. corpus-assisted CAT. What distinguishes the corpus-assisted non-CAT condition from previous research is the utilisation of printed corpus data rather than what is called “full hands-on DDL” (see chapter 3, section 3.4.1.3 for more information). In contrast to the implicitly contrastive analytical approach embedded in the use of bilingual parallel English/ Arabic translation tasks in the proposed condition (+DDL +CAT), the participants in the (+DDL –CAT) condition explicitly consulted corpus data and carried out receptive (MC) tasks and productive (gap-filling) tasks.

The researcher’s objectives were formulated into the following research hypothesis and sub-hypotheses:

H2. The corpus-assisted CAT condition will lead to the learning of a significantly larger number (if any) of adj. /noun collocations than the corpus-assisted non-CAT condition.

- a) The contrastive analysis and translation conditions (both) will lead to the passive recall of a significantly larger number of adj. /noun collocations than the non-contrastive and translation tasks.
- b) The corpus-assisted CAT condition will lead to the active recall of a significantly larger number of adj. /noun collocations than the corpus-assisted non-CAT condition.
- c) The differences between the conditions in active and passive recall will be retained in a delayed post-test.

The detailed description of the statistical analysis of the two groups' data in chapter 5 revealed that the participants in the corpus-assisted CAT group learned significantly more collocations as evidenced in both passive and active recall test results in comparison to those learned by participants in the corpus-assisted non-CAT group. This knowledge was retained significantly more by participants under the corpus-assisted CAT condition than those under the corpus-assisted non-CAT condition.

Despite the growing interest in corpus applications for language learning, and the significant potential of corpus-based resources and tools to highlight linguistic regularities such as collocations, few empirical studies have investigated corpus-assisted collocations learning. Similar to what has been suggested in section 6.2.1, it is quite hard to compare the results obtained from this study with the other empirical studies utilising a DDL approach. This is not only because of the different numbers and types of target collocations, the different tasks assigned to the participants and/ or the different ways used to measure learning, but also because of the different way in which the corpus consultation was carried out in these studies. While the present study utilised printed corpus data for the learners' consultations during the tasks, the earlier studies employed a full hands-on consultation i.e. learners accessed an online corpus with concordancing tools to complete the tasks. It is not at all clear whether such differences in corpus consultation would have any effect on the learning outcome of the DDL approach. Thus, it is safer to relate the results from the current research to the overall effect of a DDL approach on the learning of individual words and collocations. The learning effect of the two DDL conditions in this study confirms the results from previous research. For example, Sun and Wang (2003) found that web-based concordancing with online exercises resulted in significant gains in learning the target grammatical collocations. Similarly, Chan and Liou (2005) revealed that web-based concordancing has a positive impact on learning different types of verb/ noun collocations, including non-congruent ones. Moreover, Tsai's (2011) multidimensional investigation revealed that online corpus consultations had a positive impact on EFL learners' receptive, controlled productive and free productive collocational knowledge. The study also suggested that the elicited data on the learners' thought process provided evidence that a data-driven approach to collocation learning engaged them in profound cognitive processing of the collocations during both corpus consultation and task completion.

The next sections will be allocated to discussing the results and evaluating the three conditions from a theoretical perspective.

Noticeability

As noted earlier in this section and in chapter 3, the superiority of corpus-assisted CAT could be explained by the hypothesis of ‘noticing’. The noticing hypothesis, which provided the theoretical framework for FFI, suggests that noticing is necessary to convert input into intake that is essential for learning. According to Schmidt (1990, 2001), there are several factors that determine noticeability of a given form; perceptual salience, frequency and task demands. In relation to this research, three types of noticeability determinants for learning the target non-congruent collocations were more available for learners in the corpus-assisted CAT group than the non-corpus assisted CAT and the corpus-assisted non-CAT groups. The three conditions will be evaluated below in terms of the existence (+), non-existence (-) or emphasized existence (++) of these noticeability determinants.

Perceptual salience, in general, is concerned with how prominent a form (e.g. collocations) is in input. The more a form stands out in the input, the more likely it is to be noticed. Moreover, perceptually salient forms have a greater chance of impinging on awareness. According to this general notion of perceptual salience, it could be argued that the three conditions do actually provide learners with the salience necessary for noticing. The salience in the (+DDL –CAT) condition comes primarily from the fact that this is a focus-on-forms (FonFs) condition in which the target collocations are discretely addressed in the MC and gap-filling tasks. Another source of salience is provided through the corpus data in the form of concordance lines which have been proven to make linguistic regularities such as collocations perceptually salient for L2 learners. However, perceptual salience is one source of salience required for noticing. James (1996) proposed a contrast-dependent salience of form as a crucial type of perceptual salience. This contrast-dependent salience was only available for both +CAT groups, but not for the -CAT condition, at least ‘perceptually’.

According to the Revised Hierarchical Model (RHM) of the bilingual mental lexicon, the contrast (translation) may have mentally occurred during task completion. The RHM proposes that during early stages of SLA, words in the L2 are linked to their translation equivalents. This is also supported by Jiang’s (2000) psycholinguistic model of L2 vocabulary acquisition in an instructional setting. The model suggests that in the processing of an L2 vocabulary item, a learner initially maps it to its L1 translation, after which it goes through an L1 lemma mediation process for use, before it eventually reaches an L2 integration stage. The results of the active recall post-tests in the corpus-assisted

non-CAT group may give indications to the mental representations established as a result of the condition. The learners were able to translate the collocations into English, even though they were not exposed to the translations in the corpus data or in the tasks associated with it. Hence, this evidence supports the arguments made in the two models to some extent. Perceptual salience available to learners under the corpus-assisted non-CAT condition is thus evaluated as (+).

Conversely, the learners in both +CAT groups were obliged to notice the target items through cross-linguistic instruction. In the non-corpus CAT treatment, the items became salient as input when learners were taught the corresponding L1 forms and received information about the particular difficulties resulting from L1-L2 differences. The participants in the corpus-assisted CAT group were given contrastive meta-linguistic input and tasks, and were instructed to notice the differences of the target combinations between Arabic and English. In this general sense, the two CAT groups may seem to have engaged in equivalent levels of noticing, but it is not quite so simple.

Learners in the corpus-assisted CAT group were urged to observe and discover cross-linguistic differences between the target collocations in their L1 and L2 while carrying out the translation tasks. As informed by the connectionist model of SLA (see section 3.5.1 for an overview), the learning of the target collocations under corpus-assisted CAT may have occurred as a result of a reconstructing and strengthening of the mental links between the node words and their collocates. Additionally, the perceptual salience of collocations provided by consultations of bilingual parallel corpus data may have raised learners' cross-linguistic awareness, which may have made it easier for them to notice the fossilised characteristics of their interlanguage. This awareness then may have facilitated the learners' production of the target non-congruent collocations, in particular given their impressive gain in that aspect of knowledge. This is also true for the non-corpus CAT group as contrast-dependent perceptual salience was also provided for this group. Through translation tasks and the explicit explanation of the differences between the target collocations in the two languages, the learners may have established durable mental links between collocates. Thus, the evaluation of the two CAT conditions on contrast-dependent perceptual salience is (++).

It seems that perceptual salience as provided to the -CAT group was effective enough for learners to establish strong mental representations and links both between the collocations in both languages, as well as between the node words and their collocates. However, the contrast-dependent perceptual salience available to learners in the +CAT groups seems

to have led to stronger and more durable associations, as indicated by the learners' scores. The meta-linguistic data in the corpus-assisted CAT, the teacher comparative analysis in the other CAT group, and the translation tasks may have emphasized the learners' awareness towards their more or less fossilised knowledge of the target collocations, which may have initiated a process of knowledge restructuring.

Task demands is the second determinant of noticeability that relates to this research. Task demands, according to Schmidt (1990), offer one of the essential arguments that what is learned is what is noticed. Both the corpus-assisted CAT and the corpus-assisted non-CAT conditions entail corpus-data consultations i.e. a DDL approach to learning the target collocations. The keywords for the differences between the three collocation teaching and learning conditions are *discover* (in corpus-assisted CAT and corpus-assisted non-CAT) and *provided* (non-corpus CAT). In fact, these keywords constitute the rationale behind DDL as a pedagogical approach of autonomous learning. When comparing between the two corpus-assisted conditions (i.e. +DDL), this researcher would argue that it is the overall objective of the corpus-consultation task that determines the level of the task demands. While learners in the CAT group consulted parallel bilingual corpus data during their L2/ L1 and L1/ L2 translation tasks, the non-CAT group consulted a monolingual corpus data during MC and gap-filling tasks. Hence, given the overall objective of the corpus consultation, it can be argued that the corpus-assisted CAT condition is more demanding than the non-CAT one.⁶⁶ The evaluation of the task demands determinant of noticeability is (+) for the corpus-assisted non-CAT condition and (++) for the corpus-assisted CAT condition.

On the other hand, the function of discovery learning associated with DDL is more prominent in the comparison between the (+DDL and -DDL) CAT conditions. According to Schmidt (1991), certain tasks may, through their characteristics, make certain language forms salient and thus noticeable. While learners in the corpus-assisted CAT condition *discover* the differences between collocations in their L1 and the L2 and carry the contrast autonomously, cross-linguistic salience and analysis is *provided* to the learners in the non-corpus assisted CAT. The two CAT conditions are nonetheless similar when it comes to the demands of the translation tasks. Hence, the overall evaluation of the task demands as a noticeability determinant is higher in the corpus-assisted CAT (++) and lower in the teacher-fronted CAT (+). Accordingly, the learning effect of the target collocations, both receptively and productively, as well as the superiority of results attained by the corpus-

⁶⁶ This will be discussed in more detail below in relation to the involvement load hypothesis.

assisted CAT group could be attributed to enforced noticing through more intense task demand.

The final and most significant determinant of noticeability in relation to corpus-assisted CAT is *frequency*. Frequency of occurrence of a linguistic form is believed to enhance the likelihood of that form being noticed in input, and thus being learned. Frequency of occurrence and profusion of language instances is at the heart of the DDL approach to language learning. Multiple exposures to the same target collocation in the parallel corpus data and the monolingual corpus data could be a crucial contributor to collocation learning as evident by the learners' scores in the post-test of passive and active recall of collocations. The corpus data may also have provided the learners with the flood of input recommended for vocabulary retention (e.g. Laufer, 2005a, b; Thurston & Candlin, 1998), and with the wealth of resources for a principled 'recycling' of previously learned collocations as recommended by other researchers (e.g. Nation, 1990; Schmitt & Schmitt, 1995; Schmitt, 2008). This frequency of occurrence is believed to be equal for both corpus-assisted conditions in this research. In both conditions, the frequency of each of the target collocations in the monolingual and bilingual corpus data was set to five occurrences. As noted earlier in this research, the learning process as a result of frequency could be accounted for within the connectionist framework to SLA. Learning in the connectionist model occurs due to associative processes between elements and creating links between them. These links become stronger as these associations (i.e. collocations) keep recurring (Mitchell & Myles, 2004). Hence, the significant results achieved by learners in the corpus-assisted groups (both +DDL) could be explained by their ability to either form new connections between the node words and their collocates, or their ability to restructure and strengthen existing connections through intensive exposure to an abundance of corpus-data. Accordingly, both corpus-assisted conditions receive an evaluation of (++) frequency.

It is worth noting that the teacher fronted/ non-corpus assisted CAT condition in this research lacks this aspect of frequency that is provided by the corpus data within the DDL approach, but it entails the type of frequency which is typically involved in the general FFI approach. This means that the collocations appeared several times throughout the teaching phase. Hence, the evaluation of this noticeability determinant for the teacher-fronted CAT is (+), while it is (++) for the corpus-assisted ones.

To sum up, this researcher argues that with regard to the three factors which determine the noticeability of a linguistic form (the target collocations in this case) and their

evaluation mechanism of (+), (-) and (++) explained above, corpus-assisted CAT is a more effective condition than the non-corpus assisted CAT and the corpus-assisted non-CAT conditions in enhancing noticing which results in learning of non-congruent collocations. This alone could explain the superiority of results attained by the proposed approach.

Involvement load

The superiority of the results attained under corpus-assisted CAT could be explained by the ‘Involvement Load’ hypothesis as discussed earlier (in chapter 3, section 3.5.2). In the ‘task-induced involvement load’ hypothesis proposed by Laufer and Hulstijn (2001), the retention of previously unfamiliar words is conditional upon the amount of involvement of learners while processing these words. Involvement is operationalized by tasks designed to differ in three dimensions: need, search and evaluation. The need dimension constitutes the motivational element of involvement, while the search and evaluation dimensions are the cognitive ones responsible for the level of information processing. This proposed hypothesis attempted to operationalize different concepts in relation to good retention. The most significant of these is the ‘depth of processing hypothesis’ i.e. the deeper a piece of information is processed, the more likely it is to be committed to long term memory (Craik & Lockhart, 1972). Accordingly, the researchers emphasized that words processed with a higher involvement load will be better retained than those with a lower involvement load. Additionally, tasks designed with a higher involvement load will better facilitate vocabulary retention than those with a lower involvement load. In order to evaluate the involvement load in the three conditions in this research, similar criteria of (+), (-) or (++) are utilised to refer to the existence, non-existence or emphasized existence of the involvement dimensions.

In the teacher-fronted/ non-corpus assisted CAT condition, the learners were required to translate sentences with the target words from L2 to L1 and from L1 to L2. The ‘need’ element was present since the learners had to focus on the target collocations to complete the task. The element of ‘search’ was present in both translation tasks as the translation into L1 required a search for meaning, whereas translation into L2 required a search for form. However, to attain the element of search, the learners in this group either inferred meanings of the target lexical items from context, or asked the teacher for help with the form.⁶⁷ Typically, a sentence can be translated in more than one way. The final choice of

⁶⁷ Similar to Laufer and Girsai’s (2008a and 2008b) studies.

the translation must be made after an evaluation of several translation equivalents. After having attempted their own translation, students received a contrastive cross-linguistic explanation of the target lexical items from the teacher. As a result, it could be concluded that although the elements of search and evaluations are present in both translation tasks (L2/ L1 and L1/ L2), the search component in particular is believed to be of moderate level. It is worth noting that this evaluation of teacher-fronted CAT is compatible with Laufer and Girsai's (2008a, b) studies. The researchers evaluated the load involvement of CAT condition as +need +search +evaluation for the L2/ L1 translation task, and +need, +search, ++evaluation for the L1/ L2 translation task.

In corpus-assisted CAT, the need component is present in the same way as in the other non-corpus condition i.e. the learners will have to understand the target collocations in order to translate them into L1, or to produce them when translating into L2 as required by the task. However, when it comes to the other dimensions of involvement load, the learners in the corpus-assisted CAT group were involved in an autonomous search of the target non-congruent collocations in the parallel corpus data. Additionally, they were presumably involved in evaluation of the appropriateness of both the meaning and form of the collocations, as well as evaluation of the similarities or differences between the collocations in the two languages. Thus, the involvement load dimensions of research and evaluation are exceptionally high under this condition. However, research and evaluation differ between the translation tasks within corpus-assisted CAT. For the L2/ L1 translation task the load is +need ++search +evaluation, and for the L1/ L2 translation task the load entails +need, ++search, +++evaluation.

Similar to the corpus-assisted CAT the 'need' component is present in the corpus-assisted non-CAT condition. In carrying out the MC tasks, learners need to understand the context and the different available adjectives in order to choose the target adjective and match it with the node noun. In order to complete the gap-filling task, an appropriate adjective has to be provided to produce the target collocations. In order to fulfil the needs or requirements of the tasks, the learners have to engage in an autonomous search using the monolingual corpus data. This search may either confirm a learner's intuition about or knowledge of a target collocation, or establish previously non-existing knowledge and association. Thus, the search dimension in the corpus-assisted non-CAT condition is as heavily emphasized as it is in the CAT condition. It is worth noting that the evaluation of the involvement load in this condition is thus far similar to Laufer and Girsai's (2008a,

2008b) MC meaning recognition task. However, it differs in the evaluation dimension. Laufer and Girsai decided that there was no evaluation involved in their MC task as it had no context. However in the current research, the evaluation dimension is present, as learners had to evaluate, with the help of the corpus data, their choice of adjective for the context provided. On the other hand, the gap-filling task received the same evaluation on the three involvement load dimensions as the same task in Laufer and Girsai's studies. In their task, the learners needed the target word, searched for it in a provided word bank and had to make a decision (evaluation) as to which word fitted the context. In this study, the corpus data was used in the search for the adjective and evaluation of its suitability for the given context. Thus, the involvement load of the corpus-assisted non-CAT condition in both tasks is +need +search +evaluation.

Accordingly, one might claim that the task design in corpus-assisted CAT entails a higher involvement load than in the non-corpus assisted CAT and the corpus-assisted CAT conditions. Thus, corpus-assisted CAT might have engaged the learners in deeper processing of the target non-congruent collocations and committing them to the long term memory. The different involvement loads of the three conditions are reflected in the results of passive and active recall post-tests. As indicated earlier in chapter 5, the corpus-assisted CAT group better retained their knowledge of the target collocations than the other two groups in both immediate and delayed post-tests.

Pushed output

Although there used to be general agreement among researchers that 'comprehensible input' is both a crucial and a sufficient element of SLA, most now maintain that, though essential, it is not sufficient. ESL and EFL learners also need what Swain (1993, 1995, 2005) called "comprehensible output". In the 1980s, the word 'output' was used to refer to the 'outcome', or 'product' of the device utilised for language learning, and it was synonymous with what the learners have learned. In the following decades and with growing interest in Swain's 'output hypothesis', there has been a shift in the concept of output from a 'product' to the 'process' of learning. The 'output' hypothesis proposes that producing language by negotiating meaning, either through speaking or writing, will allow for language learning to occur (Swain, 1995). According to Swain, negotiating meaning has to incorporate the notion of being 'pushed' toward the delivery of precise, coherent, and appropriate language. This act of pushing, she claims, drives learners to make more efforts and to "stretch" their interlanguage resources, which in turn forces

them to process language more deeply, and helps them to move beyond their current stage of language development. Besides providing opportunity for meaningful practice of one's linguistic resources, Swain (1995) proposed three functions for pushed output: (1) the noticing/ triggering role or consciousness-raising function; (2) the meta-linguistic function or the reflective role; and (3) the hypothesis-testing function (Swain, 1993; Swain, 1995; Swain, 2005; Swain & Lapkin, 1995). The activity of producing the target language may thus push learners to become aware of gaps and problems in their current second-language system. When they encounter difficulties in producing the target language, the output tasks provide them with opportunities to reflect on and analyse these problems, and engage them in cognitive processes which may play a role in second-language learning. They do so even when external feedback is unavailable (Swain, 1995). Based on the functionality of the pushed output hypothesis, the gap-filling task in the corpus-assisted non-CAT condition, though a task of active/ productive knowledge, is a form-recall task and might not be classified as a pushed output task. Thus, the evaluation of collocation learning under this hypothesis is restricted to the two +CAT conditions.

As noted earlier in chapter 3 (section 3.5.3), translation into L2 tasks exemplify pushed output tasks. In fact, Laufer and Girsai (2008a, b) affirmed that they are perfect pushed output tasks that require stretching one's linguistic resources. The participants in the two CAT groups showed a significant increase in passive and active collocational knowledge which could be explained in relation to the 'pushed output' hypothesis and the depth of processing involved in producing precise translations in the L2. The superiority of the results attained by the corpus-assisted CAT group in the two post-testing phases on both passive and active levels of knowledge could be explained by the three functions of 'output' by Swain (1995). These functions are used to evaluate the degree to which the learners were 'pushed' in the process of learning the target collocations. Once again the criteria of (+), (-) and (++) were utilised to indicate the degree to which the function was achieved. Since the translation into L1 entails comprehension but no production of the L2, this task will not be considered as a pushed-output task. While attempting to produce a good translation of the Arabic sentences with the target collocations into English, participants in both CAT groups could not avoid problematic words or structures, since they were predetermined by the source language. Thus, they might have noticed that they did not know how to convey the meaning and produce the form that they were supposed to deliver. This might have prompted them to consciously recognise their inability to produce the target non-congruent collocations. Thus, it could be assumed that the

noticing/ triggering function is similar for both CAT groups i.e. both are (+) noticing/ triggering.

In producing translations of the target collocations, the participants in both groups were obliged to try out a translation, which in this case might be influenced by their L1. This 'trial run' is what Swain (1995, 2005) calls the hypothesis-testing function. The current researcher believes that the learning in the two CAT groups occurs at different levels for both the hypothesis-testing function and the meta-linguistic/ reflective function. Corpus-assisted CAT requires a higher level of processing during the translation task. This assumption is drawn from research investigating the acquisition of other linguistic forms than vocabulary and/ or with mainly oral pushed output tasks. This is because within the vocabulary acquisition domain, the pushed output tasks are mainly oral and they are compared to modified input (e.g. He and Ellis, 1999; de la Fuente, 2002), with the exception of Laufer and Girsai's (2008a, b) study. Despite the researchers' use of translation as pushed output tasks, the comparison of the task was made with other active/ productive form-focused task (text fill-in task), and with a meaning-focused input-based task.

Empirical studies conducted in relation to oral pushed output in communication classes identify two types of output tasks: one-way and two-way tasks (Ellis, 2003). Mackey (2012) describes one-way tasks as non-reciprocal tasks in which the learner carries out most of the talking (i.e. output) and is responsible for conveying the information to successfully complete the task. Conversely, in two-way tasks, participants exchange the information. In doing so, learners may use the method of trial and error to test their production. Consequently, the learners' production may generate responses or feedback which provides them with information about the comprehensibility or well-formedness of their output. Feedback and responses could take different forms including confirmation checks or implicit and explicit corrections. This feedback could motivate the learners to modify or reprocess their output, resulting in learning (Swain, 1993, 1995, 2005). Long (1980) suggests that two-way tasks promote more modifications of output than one-way tasks. However, there is no consensus in the literature regarding this issue. Ellis (2003) argues that this conclusion reached is premature given that other existing studies do not bear out the same conclusion. Other researchers (e.g. Iwashita, 1999; Shehadeh, 1999; among others) revealed that one-way tasks, such as picture description, create more

opportunities for learners to produce modified pushed output than opinion exchange and other two-way tasks.

This researcher tentatively suggests that in the two CAT tasks, the teacher-fronted CAT could be categorised as a two-way task. The reason for this is that in carrying out the L1/L2 translation task, the participants could ask the teacher for help whenever they needed in order to produce the translation. The translation produced in this stage is more of a 'trial run' associated with the hypothesis-testing function. Then, when the learners received corrective and contrastive feedback from the teacher, modification of output and the metalinguistic/ reflective function of the task took place. This type of externally induced reflection mediated the learning of the target collocations, and engaged the learners in a deep cognitive processing of them. Therefore, the evaluation of the teacher-fronted CAT could be seen as (+) noticing, (+) hypothesis testing and (+) reflective. On the other hand, corpus-assisted CAT could be identified as a one-way task. While carrying out the translation task and after noticing their linguistic gap, learners in this group would have either produced a written 'trial' translation of a target collocation, or mentally considered one. Then, they would have consulted the corpus data to test their hypothesis and modified or maintained their output. Thus, they had the opportunity to reflect on their knowledge of the target collocations in comparison to their intuitions⁶⁸ and attained what might be described as implicit feedback. After the completion of the task, the learners received an explicit confirmation check which may have given them a second opportunity for reflection. Accordingly, the evaluation of corpus-assisted CAT could be (+) noticing, (++) hypothesis testing and (++) reflective. It is the current researcher's assumption that it is through this self-initiated hypothesis testing, modification of output and reflection that the learners in the corpus-assisted CAT group achieved more significant learning results than those in the teacher-fronted CAT group. The significance of the learning outcome of the collocations (in the immediate and delayed post-test results) under corpus-assisted CAT suggests that the pushed output may have triggered deeper and more elaborate processing of the form, which led the students to establish a more durable memory trace.

Having justified the significance of the results achieved by the participants in the corpus-assisted CAT group, and having established an evaluation of the three teaching/learning

⁶⁸ Which could have been influenced by their L1.

conditions based on the theoretical framework of this research, the next section will provide a brief explanation for the medium effect size between the two +CAT groups.

6.2.3 Why medium effect size?

As reported earlier in the comparison between the CAT conditions, the magnitude of the statistical significance found between the pre-test and both post-tests of passive and active recall was of medium size. There are two explanations for the size of actual difference in results between the two CAT conditions. First, the gains in collocational knowledge achieved by participants in both groups could be explained and supported by the same hypotheses discussed in the previous section. However, it is the different degrees of determinants, components and functions within each hypothesis that have made the difference. The second and most likely explanation is related to the EFL learners' dependency on teachers as a teaching/ learning resource and the degree of autonomy involved in DDL and corpus consultation. Borg and Al-Busaidi (2012), among many others, have suggested that "students come to the university with limited study skills, and with an over-dependence on the teacher for their learning" (p. 8). On the other hand, cutting out the middleman in language teaching and learning is what distinguished the DDL approach as proposed by Johns (1994). Although the parallel corpus consultation in this research did not entirely 'cut out the middleman', the teacher has taken on a new role as a director. This change in the teacher's role entails a degree of autonomy that the learners might not be used to in their learning approach, which may have affected their scores.

6.3 Summary and final remarks on the proposed CAT condition

In-depth discussion of the acquisition of non-congruent collocations under the three FonFs conditions has been presented in this chapter. In comparison to the teacher-fronted CAT and non-CAT conditions, the corpus-assisted contrastive analysis and translation condition was found to be the best condition under which the acquisition of non-congruent collocations occurred. The superior results for the acquisition/ retention of the target collocations (receptively and productively) of the corpus-assisted contrastive analysis and translation group were primarily supported by the hypotheses of 'noticing', 'involvement load' and 'pushed output'. Through providing the learners in the corpus-assisted contrastive analysis and translation condition with a high involvement load, and through submitting them to a pushed output task, it is believed that the learners have been engaged in deeper processing of the target non-congruent collocations. This learning/ teaching condition may also have helped the EFL learners to establish strong mental connections

between the node words and their collocates, and restructured previous or existing knowledge or intuition about the target collocations.

What is interesting in the comparison between the three conditions in this research, is related to the notion of noticing, particularly to the contrast-dependent perceptual salience as a determinant of noticeability in both CAT conditions. Many researchers in the field of SLA acknowledge that conscious understanding and awareness of the target language systems is essential if learners are to produce correct forms and use them appropriately (Schmidt, 1990). Although both ‘consciousness’ and ‘awareness’ of an L2 system involve ‘noticing’, James (1996) argues that they should be viewed as complementary, but not synonymous. To James (*ibid*), among other researchers, awareness raising involves focusing attention on and having insight into already possessed linguistic knowledge. Consciousness raising on the other hand, is more relevant for language learners who are not yet in control of these formal repertoires and consistent intuitions. James (1996) adds that consciousness raising means drawing learners’ attention to those properties which they must learn, but might find difficult to learn. Despite the fact that awareness and consciousness are distinctive in the way outlined by James (*ibid*), he argues that they could effectively coincide, as he states:

“what interests me even more though, is the effect of their coincidence, that is in what ways the learner’s newly-raised LA [language awareness] can give him insight into the FL, and, conversely, how newly-raised consciousness of properties of the FL might impinge on his MT [mother tongue]” (p. 142).

James argues that in the early stages of FL learning and in order to overpower interlingual transfer, the learners will have to depend meta-cognitively on a combination of high language awareness of their native language and high consciousness raising of the foreign language. CA and translation tasks are one way of attaining this goal. In light of this, the current researcher would argue that the presumably enhanced noticing through autonomous CA (i.e. corpus consultation and cross- linguistic comparisons in the translation tasks) would be a hugely useful form of language awareness with consciousness-raising necessary for learning non-congruent collocations.

Chapter 7: Conclusion

This chapter starts with a brief summary of the main findings and contribution of this research in sections 7.1 and 7.2. The research's limitations, a critique of the methodological design along with recommendations for future research are presented in section 7.3. Section 7.4 comprises implications for FL pedagogy and for practitioners who are interested in the areas covered by this study.

7.1 Brief summary of the study

In recent years, there has been a growing interest in collocations in the field of EFL teaching and learning processes. Researchers acknowledge the importance of learning collocations in the foreign-language classroom for many valid reasons. Collocations serve a number of communicational and interactional functions. They are pervasive in language, so they allow more fluency in language output. Appropriate use and interpretation of collocations by L2 learners is an indication of their linguistic proficiency. However, it is acknowledged in the literature that EFL learners' collocational knowledge is lacking, and that the use of collocations is problematic for them. According to most studies investigating the collocational knowledge of EFL learners, many collocation errors are induced by L1 influence (i.e. interlingual). Learners tend to rely on L1 transfer, disregard restrictions on word combinations, and overuse the same typical collocations. According to Selinker and Lakshmanan (1992), interlingual errors are very likely to fossilise, where fossilisation is understood as the persistence of non-target forms in the interlanguage. Non-congruent collocations, i.e. those collocations that do not have word for word equivalents in the learners' L1, tend to be more problematic in their production than congruent ones. This adds to the learning burden of collocations (Peters, 2015). In order to provide the necessary salience for the learning of such problematic collocations and to raise learners' awareness of the difficulty they present, it was recommended that EFL classrooms should be supplemented by form-focused instruction which singles out the target items and has learners practice them out of an authentic context. The researchers suggested that particular emphasis should be directed at production and cross-linguistic comparison. As far as this researcher's knowledge goes, Laufer and Girsai's (2008a) study is one a few studies which examined the effects of different FFI conditions on learning non-congruent collocations, and the only one examining the effect of a contrastive/ cross-linguistic condition. Their study looked at the effect of a contrastive analysis and translation form-focused condition on the learning of non-congruent

collocations and individual words, in comparison with a meaning-focused condition. The researchers replicated their study to include another FFI task that did not involve cross-linguistic emphasis or contrasts. They concluded that the contrastive condition outperformed both non-contrastive FFI as well as MFI. On the other hand, the effectiveness of a DDL approach and corpus resources in learning collocations are empirically proven. However, they have not been employed to allow EFL learners to carry out contrastive analysis.

This research filled this methodological gap in the literature and proved the hypothesised effectiveness of the corpus-assisted contrastive analysis and translation approach both over Laufer and Girsai's teacher-fronted contrastive analysis and translation approach and over corpus-assisted non-contrastive FFI conditions. The results from the immediate and delayed post-tests demonstrated that a corpus-assisted CAT approach is beneficial for developing EFL learners' receptive and controlled productive knowledge of lexical non-congruent collocations. This approach has heightened and raised learners' consciousness of the ubiquitous phenomenon of collocations in L2, and raised the learners' awareness of the cross-linguistics differences between the collocations in the L1 and L2. The increase in learners' collocational knowledge may be attributable to the intense cognitive processing of the target non-congruent collocations induced by the corpus-assisted CAT condition, as supported and explained by the 'noticing', 'involvement load' and 'pushed output' hypotheses.

7.2 contribution of the study

The present study incorporated a combinatory approach to investigate how EFL learners could best learn non-congruent collocations. It extended Laufer and Girsai's studies to examine further the effect of contrastive analysis and translation as a pedagogical approach to learning collocations. However, it is distinctive and contributive in several important ways. First, this research explores the effect of a CAT approach on the acquisition of collocations over an extended period of time, which contributes to the validation of this approach. Second, the research provided a significant evidence in support of a pedagogical shift towards an autonomous and more effective contrastive approach through the incorporation of DDL and corpus resources. Another distinctive and contributive feature in this research is the evaluation and assessment of different aspects of the three FonFs conditions in relation to the three hypotheses that are used to explain the gains in learners' collocational knowledge in a way that was not employed in other studies. Finally, this research studied Arabic English language learners to whom

non-congruent collocations are both different and more complex, due to the complexity and richness of the Arabic language (see section 4.4.1.2 for an overview). One outcome of this context is the generation of a list of non-congruent adjective/noun collocations with the Arabic language. The collocability of the word combinations was checked statistically according to the frequency-based approach, as well as phraseologically. The generated list of non-congruent collocations can be a valuable and useful resources for future researchers, and the systematic approach to evaluating non-congruency and obtaining the list could be adopted by Arab EFL researchers in other contexts.

7.3 Limitations and directions for future research

The following sub-sections critically reflect on the limitations of the current research and point to some possible ways forward for future research.

7.3.1 The target non-congruent collocations

This research examined the effect of different conditions on the acquisition of a list of 30 frequent non-congruent adjective/ noun collocations. As described in detail in chapter 4 section 4.4.1, the extraction of the collocations was done in three stages. First, a list of statistically highly frequent collocations was extracted from the BNC. The mutual information (MI) scores test was utilised as a statistical measurement to quantify the degree of attraction between the pairs of words (adjectives/ nouns), based on a comparison between expected and observed co-occurrence frequency. An MI score of 3 or higher indicates significant and strong collocations. Then, the generated list was filtered into a shorter list of non-congruent collocations according to four judges of Arabic native speakers. Finally, this shorter list was submitted to a phraseological assessment where native speakers of English were asked to recall the collocations in a pilot gap-filling test. Only the collocations recalled by the majority of the English native speakers were chosen for the purpose of this study. Nevertheless, the list comprised four free or less restricted collocations according to Howarth's (1998) phraseological classification of collocations, although from a statistical perspective they could be still classified as frequent, strong collocates. The statistical measurement of t-scores could have been used to indicate certainty of collocations along with the MI scores. Additionally, this research only focused on the effect of corpus-assisted CAT on learning frequent non-congruent collocations. According to Daller and Xue (2009), lexical sophistication which indicates knowledge of infrequent words is most closely connected to FL students' academic success. Thus, for further research, it would also be interesting to examine the effect of a

corpus-assisted CAT on the acquisition of restricted and relatively less frequent collocations with less frequent node words by more advanced EFL learners.

7.3.2 Corpus resources

DDL is a pedagogic continuum ranging from product to process (Batstone 1995). It has the advantage of a product approach because L2 learners are presented with multiple exposures to specific aspects of language within genuine contexts. At the same time, it has a process approach towards learning since DDL promotes L2 learners' autonomy and self-discovery when learning. This research emphasized this significance and effectiveness of corpora and corpus resources, especially parallel bilingual corpora, for providing the EFL learners with the necessary salience and multiple exposures needed for collocation noticing and intake. Additionally, bilingual data which present the collocations in juxtaposition enabled the learners to attain cross-linguistic awareness. Cross-linguistic awareness is a vital asset in establishing the right lexical links. These links may not be established through the inherent salience of the target language form, but can be established through contrast-dependent salience. Despite these advantages of the DDL approach, this researcher would argue that this approach has been neglected in the Arab EFL context in general, and in the Saudi EFL context in particular. This is evident in the scarcity of research incorporating the DDL approach in this context. Corpus resources, especially large bilingual English/ Arabic corpus and concordancer tools which process Arabic data, have also been neglected to a great extent. The obvious value and strength of the findings of this research should encourage researchers in the Arab (Saudi) context to further investigate the effect of the DDL approach. The current researcher would recommend two avenues of investigation in particular: 1) the effect of a corpus-assisted contrastive approach on the acquisition of other types of collocations, and 2) the effect of a corpus-assisted contrastive approach on populations with a different level of English language proficiency. Arab applied linguists are also urged to enrich this promising field with large, general and multiple genres of parallel English/ Arabic corpora. It has been proven in this research that this type of corpus is very effective in learning non-congruent collocations.

7.3.3 Measurements of collocational knowledge

For this research, collocational knowledge was elicited through passive recall tests (English/ Arabic translation) and active recall tests (Arabic/ English translation) at the pre, post and delayed-post treatment stages. To distinguish between passive and active knowledge, this researcher followed Nation (2001), Laufer et al. (2004) and Laufer and

Girsai (2008a, 2008b). The ability to supply the word form constitutes active knowledge and the ability to supply the word meaning constitutes passive knowledge. It could be argued that the test instruments were similar to the practice tasks of both CAT groups, thus skewing the results. However, as in Laufer and Girsai's (2008a, 2008b) studies, the participants were required to translate *sentences* in the intervention translation tasks, while in the tests they were required to translate the target *collocations* only. Additionally, in the passive recall test the participants saw the target collocations in context to help them recall their meanings, and this was the same as during the intervention stage. It could thus be argued that the test requirement was not totally unfamiliar to them. However, evidence drawn from the RHM and from Jiang's (2000) psycholinguistic model of L2 vocabulary acquisition in an instructional setting would argue against this. Both models suggest that adolescent and adult learners do use their L1 to mediate their L2 learning of lexis (see sections 3.4.2.2 and 3.4.2.3 in chapter 3 for more details). This is also supported by the elicited data in this research, as the learners in the non-CAT condition were able to translate the target collocations even though no translation was practised in class. The similarity between practice and test has therefore not necessarily influenced the outcome. Nonetheless, the current researcher would admit that including another test of active knowledge, such as gap filling, would potentially have rendered a better insight into the non-CAT learners' collocational knowledge. This can be taken into consideration in further research. It would also be interesting to investigate the effect of the corpus-assisted CAT approach on the free production of collocations in e.g. essay writing tasks or translation of longer texts.

7.3.4 Further reflections

Due to time and space limitations and other practical issues, only the learning product of the proposed corpus-assisted contrastive analysis and translation condition and comparative conditions were examined. For further research, it would be very interesting to examine other dimensions of corpus-assisted CAT. Examples include:

- Investigating the learning processes i.e. the thinking processes underlying the observable translation and corpus consultation behaviour in order to understand what contributes to the restructuring of collocational knowledge. This might be achieved through collecting mentalistic data i.e. concurrent think-aloud verbal protocols.
- Learners' perceptions and attitudes towards corpus-assisted CAT approach is another dimension worth investigating in future research. This could be carried

out through quantitative measures (questionnaires), qualitative measures (interviews) or both.

7.4 Pedagogical implications

This study has borne out the efficacy of corpus-assisted CAT on developing receptive and productive collocational knowledge and raising cross-linguistic collocational awareness, and revealed a possible way in which corpus-assisted CAT mechanisms might differ from teacher-fronted CAT and non-CAT conditions. The findings have important implications for FL pedagogy as to teaching collocations (7.4.1), developing collocational knowledge (7.4.2), scaffolding contrastive FFI and raising cross-linguistic collocational awareness, and scaffolding corpus-assisted collocation learning (7.4.3).

7.4.1 Developing collocational knowledge

It seems that vocabulary pedagogy and learning in this particular EFL context (i.e. Saudi) is rather problematic. As verified by the learners' pre-test performance, their collocational knowledge is lacking despite their possible definitional knowledge of component words. This was attributed to the teaching and learning of vocabulary as discrete units, and the negligence of teaching and learning collocations as part of word knowledge in this particular EFL context.⁶⁹ As a result, learners might only have limited awareness and knowledge of collocations. They might also have the misconception that relations between words are determined exclusively by semantics, and that they can rely on the open-choice principle and on their L1 in their language production. By doing so, they might neglect the pragmatic co-occurrence of words and violate the co-occurrence restrictions. This has implications for vocabulary pedagogy in the Saudi FL context and in other FL contexts with similar teaching and learning situations. Learners should be sensitised to the multidimensionality of lexical knowledge, and be made aware of less marked aspects such as collocations and word associations. Teachers, on the other hand, should redirect their attention beyond the boundary of lexis to lexical chunks in order to help their learners become successful communicators and accurate language users.

7.4.2 Scaffolding contrastive FFI and raising cross-linguistic collocational awareness

In the literature, non-congruent collocations have proven to be problematic in terms of their production as well as processing. This puts them, particularly, at risk of being influenced by problematic FL vocabulary instructional settings. Thus, more effort should be devoted to the teaching and practice of these lexical items, and to raising learners'

⁶⁹ As indicted by the exploratory study at the outset of this research.

cross-linguistic awareness of them. As shown in this research, the learners' knowledge of this type of collocations significantly improved as a result of contrastive FFI. This gives rise to the most significant and crucial pedagogical implication of the current research. Lexical contrastive analysis (LCA) should be revived as a pedagogical approach to learning and teaching vocabulary in the bilingual FL classrooms. Before elaborating on this implication, this researcher will reiterate that although the notion of lexical contrastive analysis proposed in this research originates in Lado's (1957) 'Contrastive Analysis Hypothesis', it relates more to James' (1980) cognitive view of CA. In this research, LCA is defined in terms of the learners' cognitive ability to compare and contrast the lexical items in the L1 and L2 through the means of translation tasks and meta-linguistic corpus data.

Traditional CA was conducted by applied linguists in their ivory towers rather than by learners in classrooms. The drawbacks expected from L1 interference in FL learning was perceived to be unavoidable, and the often equal and opposite advantages of facilitative positive L1 transfer were overlooked. According to James (1996), all this has shifted ever since CA took on a cognitive nature, where the learners are not only in charge of their own learning destiny, but also have the explicit goal of including cultural understanding along with accuracy. This language learning pedagogy is concerned with linguistic comparisons and contrast-dependent methods (Widdowson, 1992; James, 1996, 2005). Implementation of CA in language classrooms could target familiar and common linguistic forms in the L1 and L2, such as collocations, before proceeding to the contrastive, unfamiliar and different forms in the two languages. However, it is believed that basic early contrasts can be motivating, since these add some intellectual content to FL learning, and encourage a habit of caution in learners. As suggested earlier in this research, James (1988, 1996) perceives translation tasks as one way of conducting CA and promoting the cultural study of language as well as a useful form of language awareness with consciousness raising. Implementing LCA pedagogy through L1-based tasks and resources might thus have a significant positive impact on learning collocations in the context of EFL learning. The current researcher would also urge for more research on the effect of LCA in the FL bilingual classroom, and would echo Laufer and Girsas' (2008a) justifications for such scarcity of contrastive FFI studies as they state:

“One reason for the relative paucity of contrastive FFI studies may be that researchers of a Second Language context did not find L1-based tasks relevant to the teaching of multilingual classes. However, contrastive FFI is of relevance in a Foreign Language context and is of theoretical interest to the Language Learning field” (p. 317).

7.4.3 Scaffolding corpus-assisted vocabulary learning

As the participants' performance data attested to the effectiveness of DDL, it is important to consider the ways in which corpus resources can better serve vocabulary-learning purposes in the EFL context. Corpus data could be useful tools for learners to exploit when carrying out vocabulary-learning tasks. For example, textbook vocabulary activities could be supplemented by printed corpus data in order to provide EFL learners with several advantages inherent to the DDL approach. Such advantages generally include: (1) promoting more autonomous learning within the classroom setting so that the teacher's role is not jeopardized; (2) providing the learners with exposure to a wealth of genuine language which they can authenticate by relating it to their learning objectives. In addition to these general and well-established advantages, bilingual corpus-data in particular could be very effective in implementing cross-linguistic and contrastive form-focused vocabulary instruction in terms of carrying out translation tasks as well as other L1-based activities. This is especially significant in the Arab EFL context, because the Arabic language has always been in dire need of collocation dictionaries. The only two currently in existence⁷⁰ suffer from serious problems such as disregarding the Arabic legacy of collocational equivalents while translating the English terms (Galal, 2015). However, the fundamental difference between corpus resources and other bilingual resources such as dictionaries is that the former provide materials which both increase the EFL learners' exposure to linguistic forms and help them induce such language patterns across their L1 and L2. Corpus resources allow for consulting language in use (i.e. real language in bilingual corpora) for language usage (i.e. collocations). This inevitably involves more analysis on the part of the learner. As argued previously, vocabulary learning in the EFL context could most effectively benefit from monolingual and bilingual corpus-data supplements. However, this research would support the argument made by researchers (e.g. Frankenberg-Garcia, 2005; Nishina, 2008) that the use of parallel corpora is most effective and should come before monolingual corpora in the bilingual classroom.

⁷⁰ Dar El-Ilm's Dictionary of Collocations (DEDC) and Al-Hafiz Arabic Collocations Dictionary (AACD)

References

- Adolphs, S. and Durow, V. (2004) 'Social-cultural integration and the development of formulaic sequences', in Schmitt, N (ed.) *Formulaic sequences: acquisition, processing and use*. Amsterdam: John Benjamins, pp. 107-126.
- Al-Ajmi, H. (2003) 'Compiling an English-Arabic parallel text corpus', in Murata, M., Yamada, S. and Tono, Y. (eds.), *Proceedings of Asian Association for Lexicography*. Meikai University: Asialex, pp. 51-54.
- Albousaif, M. A. (2011) *Factors Determining Saudi Learners' Difficulties in Attaining EFL Vocabulary*. Unpublished PhD thesis. University of Newcastle.
- Alfallaj, F. S. (1998) *An evaluative study of the teaching of English as a foreign language at the Junior College of Technology, Buraidah, Saudi Arabia*. Unpublished PhD thesis. Indiana University.
- Al-Guayyed, E. H. (1997) 'The future of English teaching in Saudi Arabia', *Knowledge*, 53(1), pp. 62-69.
- Alhammadi, F. (1998) *Common grammatical errors in the written production of second and fourth year female students majoring in English at King Faisal University*. Unpublished MA thesis. King Faisal University.
- Alhawsawi, S. (2013) *Investigating Student Experiences of Learning English as a Foreign Language in a Preparatory Programme in a Saudi university*. Unpublished PhD thesis. University of Sussex.
- Al-Jarf, R. (2006) 'Making Connections in Vocabulary Instruction', *2nd ClaSIC Conference*. Singapore, 7-9 December. Available at: <http://repository.ksu.edu.sa/jspui/bitstream/123456789/7730/1/Making%20connections%20in%20vocabulary%20instruction.pdf>
- Allan, R. (2006) 'Data-Driven Learning and Vocabulary: Investigating the Use of Concordances with Advanced Learners of English', *Centre for Language and Communication Studies, Occasional Paper*, 66. Dublin: Trinity College Dublin
- Allwright, R. (1976) 'Language learning through communication practice', *ELT Documents, London, The British Council*, 3, pp.2-14.
- Allwright, R. (1984) 'Why don't learners learn what teachers teach? 'The Interaction Hypothesis' in Singleton, D. and Little, D. (eds.) *Language learning in formal and informal contexts*. Dublin: IRRAAL, pp 3-18.
- Almazrou, R. (1988) *An evaluative study of teaching English as a foreign language in secondary schools as perceived by Saudi EFL teachers*. Unpublished MA thesis. Southern Illinois University.
- Alqahtani, S. (2009) *The role of using reading stories technique on teaching English vocabulary for EFL Learners (elementary stage) in Saudi Arabia*. Unpublished MA thesis. Imam Muhammad Ibn Saud Islamic University.

- Al-Sugayyer, K. M. (2006) 'The Saudi student departs English beyond school premises', *Knowledge*, 123(1), pp. 118-121.
- Al-Sulaiti, L. (2004) 'Arabic corpora and corpus analysis tools', in Bel, B. and Marlien, I. (eds.), *Proceedings of TALN04: XI Conference sur le Traitement Automatique des Langues Naturelles*, 2, pp. 547-552. ATALA.
- Ammar, A. and Lightbown, P. M. (2005) 'Teaching marked linguistics structures—more about the acquisition of relative clauses by Arab learners of English', in Housen, A and Pierrard, M (eds) *Investigations in Instructed Second Language Acquisition*. Berlin: Mouton de Gruyter, pp. 167-98.
- Aston, G. (1995) 'Corpora in Language Pedagogy: Matching Theory and Practice', in Cook, G. H and Seidlhofer, B. (Eds.) *Principle and Practice in Applied Linguistics: Studies in Honour of H. G Widdowson*. Oxford: Oxford University Press, pp. 257-270.
- Bahns, J., and Eldaw, M. (1993) 'Should we teach EFL students collocations?', *System*, 21(1), pp. 101-114.
- Bahns, J., Burmeister, H., and Vogel, T. (1986) 'The pragmatics of formulas in L2 learner speech: Use and development', *Journal of Pragmatics*, 10, pp. 693–723.
- Barfield, A., and Gyllstad, H. (2009) 'Introduction: Researching L2 collocation knowledge and development', in Barfield, A and Gyllstad, H. (eds.) *Researching collocations in another language multiple interpretation*. London: Palgrave Macmillan, pp. 1-18.
- Barlow, M. (1996) 'Corpora for Theory and Practice', *International Journal of Corpus Linguistics*, 1(1), 1-37.
- Barnbrook, G. (1996) *Language and computers: A Practical Introduction to the Computer Analysis of Language*. Edinburgh: Edinburgh University Press.
- Bartsch, S. (2004) *Structural and Functional Properties of Collocations in English*. Tübingen: Gunter Narr Verlag.
- Batstone, R. (1995) 'Product and process: Grammar in the second language classroom', in Bygate, M., Tonkyn, A., and William, E. (eds.) *Grammar and the language teacher*. London: Prentice Hall, pp. 224-236.
- Beheydt, L. (1987) 'The semantization of vocabulary in foreign language learning', *System*, 15 (1), pp. 55- 67.
- Belpoliti, F., and Plascencia-Vela, A. (2013) 'Translation Techniques in the Spanish for Heritage Learners' Classroom: Promoting Lexical Development', in Tsagari, D. and Floros, G. (eds.) *Translation in Language Teaching and Assessment*. Newcastle upon Tyne: Cambridge Scholars Publishing, pp. 65-92.
- Benson, M., Benson, E., and Ilson, R. (1997) *The BBI combinatory dictionary of English: A guide to word combinations (revised edition)*. Amsterdam: John Benjamins.
- Benson, M., Benson, E. and Ilson, R. F. (1986) *Lexicographic Description of English*. Philadelphia: John Benjamins.

- Bernardini, S. (2002) 'Exploring New Directions for Discovery Learning', *Language and Computers*, 42(1), pp. 165-182.
- Bernardini, S. (2004) 'Corpora in the Classroom - an Overview and Some Reflections on Future Developments', in Sinclair, J. (Ed.) *How to Use Corpora in Language Teaching*. Amsterdam: John Benjamins, pp. 15-34.
- Biber, D. (2009) 'Corpus-Based and Corpus-driven Analyses of Language Variation and Use' *Oxford Handbooks Online*, available at: <http://www.oxfordhandbooks.com/view/10.1093/oxfordhb/9780199544004.001.0001/oxfordhb-9780199544004-e-008> (Accessed: 29 Mar. 2016).
- Bishop, B. (1998) 'A history of the Arabic language', *Linguistics*, 450[online]. Available at: <http://linguistics.byu.edu/classes/ling450ch/reports/arabic.html>
- Blaikie, N. (2000) *Designing Social Research*. Cambridge: Polity Press.
- Bolinger, D. (1976) 'Meaning and memory', *Forum Linguisticum*, 1, pp. 1-14.
- Boulton, A. (2009a) 'Data-driven learning: reasonable fears and rational reassurance', *Indian Journal of Applied Linguistics*, 35(1), pp.81-106.
- Boulton, A. (2009b) 'Data-Driven Learning: On Paper, In Practice', In Harris, T and Moreno Jaén, M. (eds.) *Corpus Linguistics in Language Teaching*. Bern: Peter Lang, pp.17-52.
- Boulton, A. (2009c) 'Testing the Limits of Data-Driven Learning: Language Proficiency and Training', *ReCALL*, 21(1), pp. 37-54.
- Boulton, A. (2010) 'Data-driven learning: taking the computer out of the Equation', *Language Learning, Wiley-Blackwell*, 60 (3), pp.534-572.
- Borg, S. and Al-Busaidi, S. (2012) 'Teachers' beliefs and practices regarding learner autonomy', *ELT Journal*, 66(3), pp.283-292.
- Brashi, A. (2009) 'Collocability as a problem in L2 production', *Reflection on English Language Teaching (National University of Singapore)*, 8 (1), pp. 21–34.
- Braun, S. (2005) 'From Pedagogically Relevant Corpora to Authentic Language Learning Contents', *ReCALL*, 17(1), pp. 47-64.
- Braun, V. and Clarke, V. (2006) 'Using thematic analysis in psychology', *Qualitative Research in Psychology*, 3 (2), pp. 77-101.
- Breyer, Y. (2006) My Concordancer: tailor-made software for language learners and teachers. In S. Braun, k. Kohn and J Mukherjee (eds.), *Corpus technology and language pedagogy: new resources, new tools, new methods. (English corpus linguistics, 3)*. Frankfurt: Peter Lang, pp. 157–176.
- Brown, D. (1974) 'Advanced vocabulary teaching: the problem of collocation', *RELC Journal*, 5, pp. 1-11.
- Brown, D. (2010) 'What Aspects of Vocabulary Knowledge Do Textbooks Give Attention to?', *Language Teaching Research*, 15(1), pp. 83–97.

- Brown, J. D., and Rodgers, T. S. (2002) *Doing second language research*. Oxford: Oxford University Press.
- Browne, C. (2002) To push or not to push: A vocabulary research question. Aoyama Gakuin University Press.
- Bryman A. (2012) *Social research methods*. Oxford: Oxford University Press.
- Carter, R. (1988) 'Vocabulary, cloze and discourse: An applied linguistic view', in R. Carter, R. and McCarthy, M. (eds.), *Vocabulary and language teaching*. Harlow: Longman, pp. 161-180.
- Carter, R. and McCarthy, M. (1995) 'Grammar and the Spoken Language', *Applied Linguistics*, 16(2), pp. 141-158.
- Chi Man-Lai, A., Wong Pui-Yiu, K. and Wong Chau-ping, M. (1994) 'Collocational problems amongst ESL learners: A corpus-based study', in Flowerdew, L. and Tong, A.K.K. *Entering text* (eds.) Hong Kong: Language Centre, Hong Kong University of Science and Technology, and Department of English, Guangzhou Institute of Foreign Languages, pp.157-165.
- Chan, T. P., and Liou, H. C. (2005) 'Effects of Web-based concordancing instruction on EFL students' learning of verb-noun collocations', *Computer Assisted Language Learning*, 18, pp. 231-251.
- Chambers, A. (2005) 'Integrating Corpus Consultation in Language Studies', *Language Learning & Technology*, 9(2), pp. 111-125.
- Chomsky, N. (1965) *Aspects of the theory of syntax*. Cambridge, MA: MIT Press.
- Clear, J. (1996) 'From Firth principles: computational tools for the study of collocation', in Baker, M., Francis, G. and Tognini-Bonelli, E. (eds.) *Text and Technology: In honour of John Sinclair*. John Benjamins Publishing, pp. 271- 292.
- Clough, P & Nutbrown, C. (2002) *Justifying Enquiry: a student's guide to methodology*. London: SAGE.
- Cobb, T. (1999) 'Breadth and Depth of Lexical Acquisition with Hands-on Concordancing', *Computer Assisted Language Learning*, 12(4), pp. 345-360.
- Coe, R. (2002) 'It's the effect size, stupid: What effect size is and why it is important' *Paper presented at the Annual Conference of the British Educational Research Association*, University of Exeter, England, 12-14 September. Available at: <http://www.leeds.ac.uk/educol/documents/00002182.htm> (Accessed: 10 March 2014).
- Cohen, J. (1988) *Statistical power analysis: A computer program*. London/New York: Routledge.
- Cohen, J. (1992) 'A power primer', *Psychological bulletin*, 112(1), pp.155-159.
- Cohen, L., Manion, L. and Morrison, K. (2007) *Research Methods in Education*. London: Routledge.
- Corder, S. P. (1967) 'The significance of learners' errors', *IRAL*, 5, pp.161-170.

- Corder, G. W. and Foreman, D.I. (2011) 'Nonparametric Statistics for Non-Statisticians: A Step-by-Step Approach', New Jersey: John Wiley & Sons.
- Corson, D.J. (1995) *Using English Words*. Dordrecht: Kluwer Academic Publishers.
- Cowie, A. P. (1981) 'The treatment of collocations and idioms in learners' dictionaries', *Applied Linguistics*, 2(3), pp. 223-235.
- Cowie, A. P. (1992) 'Multiword lexical units and communicative language teaching', in Arnaud, P. and Bejoint, H. (eds.) *Vocabulary and applied linguistics*. London: Macmillan Academic and Professional LTD, pp. 216-331.
- Cowie, A. P. (1994) 'Phraseology', in Asher, R. E. and Simpson, J.M.Y. (eds.), *The encyclopaedia of language and linguistics*, Vol. 6. Oxford: Pergamon Press, pp. 3168-3171.
- Cowie, A. P. (1998) 'Introduction', in Cowie, A.P. (ed.) *Phraseology, Theory, Analysis, and Applications*. Oxford: Clarendon, pp. 1-20.
- Craik, F.I.M., and Lockhart, R.S. (1972) 'Level of processing: A framework for memory research', *Journal of Verbal Learning and Verbal Behaviour*, 11, pp. 671- 484.
- Creswell, J. W. (2014) *Research Design Qualitative, Quantitative, and Mixed Methods Approaches*. Thousand Oaks, CA: SAGE Publications
- Crotty, M. (2003) *Foundations of Social Research: Meaning and Perspective in the Research Process*. London: Sage.
- Daller, H., Milton, J., and Treffers-Daller, J. (2007) 'Editors' introduction: conventions, terminology and an overview of the book', In Daller, H. Milton, J. and J. Treffers-Daller, J. (eds.). *Modelling and Assessing Vocabulary Knowledge* Cambridge: Cambridge University Press, pp. 1-32.
- Daller, H. M., and Xue, H. (2009) 'Vocabulary Knowledge and Academic Success: a Study of Chinese Students in UK Higher Education', in In Richards, B., Daller, H. M., Malvern, M., Meara, P., and Treffers-Daller, J. (eds.), *Vocabulary Studies in first and second language acquisition. The interface between theory and applications*. Houndmills Basingstoke: Palgrave Macmillan, pp. 179-193.
- De la Fuente, M J. (2002) 'Negotiation and oral acquisition of L2 vocabulary. The roles of input and output in receptive and productive acquisition of words', *Studies in Second Language Acquisition*, 24, pp. 81-112.
- Dörnyei, Z (2007) *Research methods in applied linguistics: quantitative, qualitative, and mixed methodologies*. Oxford: Oxford University Press.
- Duff, P. and Early, M. (1996) 'Problematics of classroom research across sociopolitical contexts', in Gass, S. and Schachter, J. (eds.), *Second language classroom research: Issues and opportunities*. Hillsdale, NJ: Lawrence Erlbaum, pp.1-30.
- Durrant, P. (2008) *High frequency collocations and second language learning*. Unpublished PhD thesis: University of Nottingham.

- Durrant, P., and Schmitt, N. (2009) 'To what extent do native and non-native writers make use of collocations?', *IRAL - International Review of Applied Linguistics in Language Teaching*, 47, pp. 157–177.
- Ellis, N. C. (1996) 'Sequencing in SLA: Phonological memory, chunking, and points of order', *Studies in Second Language Acquisition*, 18, pp. 91-126.
- Ellis, N. C. (1997) 'Vocabulary acquisition: Word structure, collocation, word-class, and meaning', in Schmitt, N., and McCarthy, M. (eds.) *Vocabulary: Description, acquisition and pedagogy*. Cambridge: Cambridge University Press, pp.122-139.
- Ellis, N. C. (1998) 'Emergentism, Connectionism and Language Learning', *Language Learning*, 48(4), pp. 631-664.
- Ellis, N. C. (1999) 'Cognitive approaches to SLA', *Annual Review of Applied Linguistics*, 19, pp. 22-42.
- Ellis, N. C. (2002) 'Reflections on frequency effects in language processing', *Studies in Second Language Acquisition*, 24, pp. 297–339.
- Ellis, N. C. (2003) 'Constructions, chunking, and connectionism: The emergence of second language structure', in Doughty, C. & Long M. H. (eds.) *The Handbook of Second Language Acquisition*. Malden, MA: Blackwell, pp. 63-103.
- Ellis, N. C. (2006) 'Selective attention and transfer Phenomena in L2 acquisition: Contingency, cue Competition, salience, interference, overshadowing, blocking, and perceptual learning', *Applied Linguistics*, 27(2), pp.164-194.
- Ellis, N. C. (2008) 'Words and their usage: Commentary on the special issue on the bilingual mental lexicon', *The Mental Lexicon*, 3(3), pp. 375–385.
- 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(4), pp. 375-396.
- Ellis, R. (2001a) 'Introduction: Investigating form-focused instruction. *Language Learning*', 51, pp. 1-40.
- Ellis, R. (2001b) 'Focusing on form: Towards a research agenda', in Renandya, W., and Sunga, N. (eds.) *Language curriculum and instruction in multicultural societies*. Singapore: SEAMEO, pp. 123-144.
- Ellis, R. (2003) *Task-based language learning and teaching*. Oxford: Oxford University Press.
- Ellis, R., Basturkmen, H., and Loewen, S. (2002) 'Doing focus on form', *System*, 30, pp. 419- 432.
- Ellis, R., and He, X. (1999) 'The roles of modified input and output in the incidental acquisition of word meanings', *Studies in Second Language Acquisition*, 21, pp. 285-301.
- Erman, B., and Warren, B. (2000) 'The idiom principal and the open- choice principle', *Text*, 20, 1, pp 29-62.

- Evert, S. (2004) *The Statistics of Word Cooccurrences: Word Pairs and Collocations*. PhD thesis. Institut für maschinelle Sprachverarbeitung, University of Stuttgart [online]. Available at <http://elib.uni-stuttgart.de/opus/volltexte/2005/2371/pdf/Evert2005phd.pdf> (Accessed: 17 April, 2013).
- Evert, S. (2008) 'Corpora and collocations', in Lüdeling, A., and Kytö, M. (eds.) *Corpus linguistics: An international handbook* (extended manuscript of Chapter 58). Berlin: Mouton de Gruyter, pp. 1-53. Available at http://www.stefan-evert.de/PUB/Evert2007HSK_extended_manuscript.pdf (Accessed: 17 April, 2013).
- Faerch, C., Haastrup, K. and Phillipson, R. (1984) *Learner Language and Language Learning*. Clevedon: Multilingual Matters.
- Fan, M. (2009) 'An exploratory study of collocational use by ESL students-A task based approach', *System*, 37, pp. 110-123.
- Farghal, M., and Obiedat, H. (1995) 'Collocations: A neglected variable in EFL', *International Review of Applied Linguistics*, 33, pp. 313-331.
- Farr, F. (2008) 'Evaluating the use of corpus-based instruction in a language teacher education context: perspectives from the users', *Language Awareness*, 17(1), pp. 25-43.
- File, K.A., and Adams, R. (2010) 'Should vocabulary instruction be integrated or isolated?', *TESOL Quarterly*, 44, pp.222-249.
- Firth, J. R. (1957/1968) 'A synopsis of linguistic theory 1930-55', in F. R. Palmer (ed.), *Selected papers of J.R. Firth, 1952-59*. Harlow: Longmans, pp. 168-205.
- Flowerdew, J. (1996) 'Concordancing in Language Learning', in Pennington, M.C. (ed.) *The Power of CALL*. Houston, Texas: Athelstan, pp. 97-113.
- Flyvbjerg, B. (2001) *Making social science matter: Why social enquiry fails and how it can succeed again*. Cambridge: Cambridge University Press.
- Foster, P. (2001) 'Rules and routines: A consideration of their role in the task-based language production of native and non-native speakers', in Bygate, M. Skehan, P., and Swain, M. (eds.) *Language tasks: Teaching, learning and testing*. Harlow: Longman, pp. 75-93.
- Frankenberg-Garcia, A. (2005) 'Pedagogical uses of monolingual and parallel concordances', *ELT journal*, 59(3), pp.189-198.
- Fred L., Perry, Jr. (2011) *Research in Applied Linguistics: Becoming a Discerning Consumer*. New York: Routledge.
- Galal, M. (2015) 'Lost in Collocation: When Arabic Collocation Dictionaries Lack Collocations', *English Language and Literature Studies*, 5(2), pp. 18-38.
- Gass, S. (1989) 'Language universals and second language acquisition', *Language learning*, 39, pp. 497-534.

- Gass, S., and Selinker, J. (2008) *Second language acquisition: An introductory course*. London: Routledge.
- Gavioli, L. (2002) 'Some thoughts on the Problem of Representing ESP through small corpora', in Kettemann, B., and Marko, G. (eds.) *Teaching and Learning by Doing Corpus Analysis*. Amsterdam/ New York: Rodopi.
- Goundareva, I. (2011) 'Effect of translation practice on vocabulary acquisition in L2 Spanish', *Working Papers of the Linguistics Circle of the University of Victoria*, 21, pp. 145-154.
- Granger, S. (1998) 'Prefabricated patterns in advanced EFL writing: collocations and formulae', in Cowie, A.P. (ed.) *Phraseology: theory, analysis, and applications*. Oxford: Oxford University Press, pp. 145-160.
- Greenbaum, S. (1988) 'Some verb-intensifier collocations in American and British English', in Greenbaum, S. (ed.) *Good English and the grammarian*. London: Longman, pp. 113-124.
- Grissom, R.J. and Kim, J.J. (2005) *Effect Sizes for Research: Univariate and Multivariate Applications*. Mahwah, N.J. : Lawrence Erlbaum Associates.
- Grix, J. (2004) *The foundations of research*. London: Palgrave Macmillan.
- Groom, N. (2009) 'Effects of second language immersion on second language collocational development', in Barfield, A and Gyllstad, H. (eds.) *Researching collocations in another language: Multiple interpretations*. Basingstoke: Palgrave Macmillan, pp. 21-33.
- Guba, E. G., & Lincoln, Y. S. (1998) 'Competing paradigms in qualitative research', in Denzin, N.K. and Lincoln, Y.S. (eds.), *The landscape of qualitative research: Theories and issues*. Thousand Oaks, CA: Sage, pp. 195-220.
- Gwet, K.L. (2012) *Handbook of inter-rater reliability: the definitive guide to measuring the extent of agreement among multiple raters (4th edition)*. Gaithersburg, MD: Advanced Analytics Press.
- Gyllstad, H. (2007) *Testing English collocations: Developing receptive tests for use with advanced Swedish learners*. PhD thesis. Lund University [online]. Available at: <http://lup.lub.lu.se/luur/download?func=downloadFile&recordId=599011&fileId=2172422> (Accessed: 1 June 2014)
- Haastrup, K., and Henriksen, B. (1998) 'Vocabulary acquisition: From partial to precise comprehension', in Haastrup, K. and Viberg, A. (eds.), *Travaux de l'Institut de Linguistique de Lund: Vol. 38. Perspectives on lexical acquisition in a second language*. Lund, Sweden: Lund University Press, pp. 97-114.
- Hadlich, R. L. (1965) 'Lexical Contrastive Analysis', *The Modern Language Journal*, 49, (7), pp. 426-429.
- Halliday, M. A. K. (1966) 'Lexis as a linguistic level', in Bazell, C. E J. C. Catford, J.C. Halliday, M.A.K., and Robins, R.H. (eds.) *In Memory of J. R. Firth*. London: Longmans, pp. 148- 162.

- Hasanuzzaman, H. (2013) 'Arabic Language: Characteristics and Importance', *the Echo*, 1(3), pp. 11-16.
- Hasselgren, A. (1994) 'Lexical teddy bears and advanced learners: A study into the ways Norwegian students cope with English vocabulary', *International Journal of Applied Linguistics*, 4(2), pp. 237-258.
- Haynes, M. and Baker, I. (1993) 'American and Chinese readers learning from lexical familiarization in English texts', in T. Huckin, T., Haynes, M., and Coady, T. (eds.) *Second Language Reading and Vocabulary Acquisition*. Norwood, NJ: Ablex, pp. 153-178.
- Henriksen, B. (1999) 'Three dimensions of vocabulary development', *Studies in Second Language Acquisition*, 21, pp. 303-317.
- Herbst, T. (1996) 'What are collocations: Sandy beaches or false teeth?', *English Studies*, 77(4), pp. 379-393.
- Hill, M., and Laufer, B. (2003) 'Type of task, time-on-task and electronic dictionaries in incidental vocabulary acquisition', *IRAL*, 41, pp.87–106.
- Hoey, M. (1991) *Patterns of lexis in text*. Oxford: Oxford University Press.
- Hoey, M. (2005) *Lexical priming: A new theory of words and language*. London: Routledge.
- Horst, M., Cobb, T. and Meara, P. (1998) 'Beyond a Clockwork Orange: Acquiring a second language vocabulary through reading', *Reading in a Foreign Language*, 11, pp. 207–223.
- Horst, M., Cobb, T. and Nicolae, I. (2005) 'Expanding academic vocabulary with an interactive on-line database', *Language Learning and Technology*, 9 [online]. Available at: <http://llt.msu.edu/vol9num2/horst/> (Accessed: 19 August 2014).
- Housen, A. and Pierrard, M. (2005) 'Investigating instructed second language acquisition', in Housen, A and Pierrard, M. (eds.), *Investigations in instructed second language acquisition*. Berlin: Walter de Gruyter, pp.1-27.
- Howarth, A. P. (1998a) 'Phraseology and second language proficiency', *Applied Linguistics*, 19(1), pp. 24-44.
- Huckin, T, and Bloch, J. (1993) 'Strategies for inferring word meaning in context: A cognitive model', In Huckin, T., Haynes, M. and Coady, J. (eds.) *Second Language Reading and Vocabulary Acquisition*. Norwood, NJ: Ablex, pp. 153-178.
- Hunston, S. (2002) *Corpora in Applied Linguistics*. Cambridge: Cambridge University Press.
- Hussein, R. F. (1990) 'Collocations: The missing link in vocabulary acquisition amongst EFL learners', *Papers and Studies in Contrastive Linguistics*, 26, pp. 123-136.
- Hussein, R. F. (1998) 'Collocations revisited', *King Saud University Journal, Language & Translation*, 10, pp. 39-47.

- Iwashita, N. (1999) 'Tasks and learners' output in nonnative-nonnative interaction', in Kanno, K. (ed.) *The acquisition of Japanese as a second language*. Amsterdam: John Benjamins' Publishing Company, pp. 31-53.
- Jaen, M. M. (2010) 'Teaching collocations: a corpus-based approach', in B. B. Fortuno, B.B., Cubillo, M.C. and Gea-Valor, M.L. (eds.) *Exploring Corpus-Based Research in English Language Teaching*. UNIVERSITAT JAUME I: Universidad Jaume I. Servicio De Comunicación Y Publicaciones, pp.13-22.
- James, C. (1980) *Contrastive Analysis*. Harlow: Longman.
- James, C. (1996) 'A cross-linguistic approach to language awareness. Special issue: Cross linguistic approaches to language awareness', *Language Awareness*, 5, pp.138-148.
- James, C. (2005) 'Contrastive analysis and the language learner', in Allerton, D.J., Tschichold, C., and Wieser, J. (eds.) *Linguistics, Language Teaching and Language Learning*. Basel: Schwabe, pp.1-20.
- Jarvis, S. (2009) 'lexical transfer', in Pavlenko, A. (ed.) *The Bilingual Mental Lexicon Interdisciplinary Approaches*. Cromwell Press Ltd, pp. 99-124.
- Jarvis, S. and Pavlenko, A. (2008) *Crosslinguistic Influence in Language and Cognition*. London/New York: Routledge.
- Jiang, N. (2000) 'Lexical representation and development in a second language', *Applied Linguistics*, 21, pp. 47-77.
- Jiang, N. (2002) 'Form-meaning mapping in vocabulary acquisition in a second language', *Studies in Second Language Acquisition*, 24, pp. 617-637.
- Jiang, N. (2004) 'Semantic transfer and its implications for vocabulary teaching in a second language', *The Modern Language Journal*, 24, pp. 416-432.
- Joe, A. (1995) 'Text-based tasks and incidental vocabulary learning: A case study', *Second Language Research*, 11, pp. 159-177.
- Johns, T. (1986) 'Micro-concord: a language-learner's research tool', *System*, 14(2), pp. 151-162.
- Johns, T. (1991) 'Should You Be Persuaded: Two Samples of Data-Driven Learning Materials', in Johns, T. and King, P. (eds.) *Classroom Concordancing*. Birmingham: Centre for English Language Studies, University of Birmingham, pp. 1-13.
- Johns, T. (1994) 'From Printout to Handout: Grammar and Vocabulary Teaching in the Context of Data-Driven Learning', in Odlin, T. (ed.) *Perspectives on Pedagogical Grammar*. Cambridge: Cambridge University Press, pp. 293- 313.
- Johnson, B., and Christensen, L. (2004) *Educational research: Quantitative, qualitative, and mixed approaches* (2nd ed.). Needham Heights, MA: Allyn & Bacon.
- Kennedy, C. and Miceli, T. (2010) 'Corpus-Assisted Creative Writing: Introducing Intermediate Italian Learners to a Corpus as a Reference Resource', *Language Learning and Technology*, 14(1), pp. 28-44.

- Kinncar, P. R. and Gray, C.D. (2012) *IBM SPSS statistics 19 made simple*. Hove, East Sussex; New York: Psychology Press.
- Knight, S.M. (1994) 'Dictionary use while reading: The effects on comprehension and vocabulary acquisition for students of different verbal abilities', *Modern Language Journal*, 78, pp. 285–299.
- Krashen, S. (1989) 'We acquire vocabulary and spelling by reading: additional evidence for the input hypothesis', *Modern Language Journal*, 73, pp. 445-464.
- Kroll, J.F., Hell, J.G., Tokowicz, N., and Green, D. W. (2010) 'The Revised Hierarchical Model: A critical review and assessment', *Biling (Camb Engl)*, 13(3), pp. 373-381.
- Kupferberg, I. (1999) 'The cognitive turn of contrastive analysis: Empirical evidence', *Language Awareness*, 8, pp. 210- 222.
- Kupferberg, I. and Olshtain, E. (1996) 'Explicit contrastive instruction facilitates the acquisition of difficult L2 forms. Special issue on Cross linguistic approaches to language awareness', *Language Awareness*, 5, pp. 149–165.
- Lado, R. (1957) *Linguistics across Cultures*. Ann Arbor: University of Michigan Press.
- Larson-Hall, J. (2010) *A guide to doing statistics in second language research using SPSS*. New York, NY: Routledge.
- Laufer, B. (1998) 'The development of passive and active vocabulary in a second language: same or different?', *Applied Linguistics*, 19 (2), pp.255-271.
- Laufer, B. (1990a) 'Why are some words more difficult than others? Some intralexical factors that affect the learning of words', *IRAL - International Review of Applied Linguistics in Language Teaching*, 28(4), pp. 293-308.
- Laufer, B. (1990b) 'Ease and difficulty in vocabulary learning: Some teaching implications', *Foreign Language Annuals*, 23, pp.147-155.
- Laufer, B. (2000) 'Electronic dictionaries and incidental vocabulary acquisition: Does technology make a difference?', in Heid, U., Evert, S., Lehmann, E., and Rohrer, C. (eds.) *EURALEX*. Stuttgart: Stuttgart University, pp. 849–854.
- Laufer, B. (2001) 'Reading, word-focused activities and incidental vocabulary acquisition in a second language', *Prospect*, 16 (3), pp. 44-54.
- Laufer, B. (2003) 'Vocabulary acquisition in a second language: Do learners really acquire most vocabulary by reading?', *Canadian Modern Language Review*, 59, pp. 565–585.
- Laufer, B. (2005a) 'Instructed second language vocabulary learning: The fault in the 'default hypothesis'', in Housen, A., and Pierrard, M. (eds.) *Investigations in instructed second language acquisition*. Berlin: Mouton de Gruyter, pp. 311-329.
- Laufer, B. (2005b) 'Focus on form in second language vocabulary acquisition', in Foster-Cohen, S.H., Garcia-Mayo, M.P. and Cenoz, J. (eds.) *EUROSLA Yearbook 5*. Amsterdam: John Benjamins, pp. 223-250.

- Laufer, B. (2006) 'Comparing Focus on Form and Focus on FormS in second language vocabulary learning', *The Canadian Modern Language Review*, 63, pp. 149-166.
- Laufer, B. (2010) 'Form focused instruction in second language vocabulary learning', in Chacón-Beltrán, R., Abello-Contesse, C., Torreblanca-López, M.M. and López-Jiménez, M.D. (eds.) *Further insights into non-native vocabulary teaching and learning*. Bristol, Buffalo, Toronto: Multilingual Matters, pp. 15-27.
- Laufer, B. and Girsai, N. (2008a) 'Form-focused instruction in second language vocabulary learning: a case for contrastive analysis and translation', *Applied Linguistics*, 29, pp.694- 716.
- Laufer, B. and Girsai, N. (2008b) 'The use of native language for improving second language vocabulary', in Stavans, A. and Kupferberg, I. (eds.) *Studies in Language and Language Education. Essays in Honor of Elite Olshtain*. Jerusalem:The Hebrew University Magnes Press, pp. 261-75.
- Laufer, B., and Goldstein, Z. (2004) 'Testing vocabulary knowledge: size, strength and computer adaptiveness', *Language Learning*, 54 (3), pp. 399-436.
- Laufer, B., and Hulstijn, J. (2001) 'Incidental vocabulary acquisition in a second language: the construct of task- induced involvement', *Applied Linguistics*, 22 (1), pp. 1-26.
- Laufer, B., and Waldman, T. (2011) 'Verb-noun collocations in second language writing: A corpus analysis of learners' English', *Language Learning*, 61(2), pp. 647-672.
- Leech, G. (1997) 'Teaching and language corpora: a convergence', in Wichmann, A., Fligelstone, S., McEnery, T., and Knowles, G. (eds.) *Teaching and Language Corpora*. London: Longman, pp. 1-23.
- Leech, N.L. and Onwuegbuzie, A.J. (2002) 'A Call for Greater Use of Nonparametric Statistics', *Paper presented at the annual-meeting of the Mid-South Educational Research Association*, Chattanooga, TN, November 7. Available at: <http://files.eric.ed.gov/fulltext/ED471346.pdf> (Accessed: 13 November 2013).
- Leech, G., Rayson, P., and Wilson, A. (2001) *Word frequencies in written and spoken English: Based on the British National Corpus*. London: Longman.
- Lewis, M. (2000) 'Learning in the lexical approach', in Lewis, M. (ed.) *Teaching collocation: Further developments in the lexical approach*. Hove, UK: Language Teaching Publications, pp. 155-184.
- Li, J., and Schmitt, N. (2009) 'The acquisition of lexical phrases in academic writing: A longitudinal case study', *Journal of Second Language Writing*, 18, pp. 85-102.
- Lightbown, P.M., and Spada, N. (2013) *How languages are learned*. Oxford: Oxford University Press.
- Lin, Y. P. (2002) *The effects of collocation instruction on English vocabulary developments of senior high students in Taiwan*. Unpublished MA Thesis: Kaohsiung University.
- Liu, C. P. (2000) 'A study of strategy use in producing lexical collocations', in Katchen, J. and Yiu-nam, L. (eds.) *Selected Papers from the Ninth International*

Symposium on English Teaching. Taipei, Taiwan: Crane Publishing Co., Ltd, pp. 481-492.

- Loewen, S. (2010) 'Focus on Form', in Hinkel, E. (ed.) *Handbook of research in second language teaching and learning. Vol II*. New York: Routledge, pp. 576-592.
- Long, M. (1981) 'Input, interaction and second language acquisition', in Winitz, H. (ed.), *Annals of the New York Academy of Sciences: Vol. XX. Native language and foreign language acquisition*. New York: New York Academy of Sciences, pp. 259-278.
- Long, M. (1991) 'Focus on form: A design feature in language teaching methodology', in de Bot, K., Ginsberg, R. and Kramsch, C. (Eds.) *Foreign language research in cross-cultural perspective*. Amsterdam: John Benjamins, pp. 39-52.
- Long, M., and Robinson, P. (1998) 'Focus on form: Theory, research and practice', in Doughty, C. and Williams, J. (eds.) *Focus on form in classroom second language acquisition*. Cambridge: Cambridge University Press, pp. 15-41.
- Lowie, W. and Seton, B.J. (2013) *Essential Statistics for Applied Linguistics*. Basingstoke: Palgrave MacMillan.
- Lupescu, S., and Day, R.R. (1993) 'Reading, dictionaries and vocabulary learning', *Language Learning*, 43, pp. 263- 287.
- Lynch, B. K. (1996) *Language Program Evaluation: Theory and practice*. Cambridge: Cambridge University press.
- Lyster, R., and Ranta, L. (1997) Corrective feedback and learner uptake: negotiation of form in communicative classrooms, *Studies in Second Language Acquisition*, 19, pp. 37-66.
- Machida, S. (2008) 'A Step Forward to Using Translation to Teach a Foreign/Second Language', *Electronic Journal of Foreign Language Teaching*, 5, Suppl. 1, pp. 140-155.
- Mack, L. (2010) 'The Philosophical Underpinnings of Educational Research', *Polyglossia*, 19, pp. 5-11.
- Mackey, A. (2012) *Input, interaction and corrective feedback in L2 classrooms*. Oxford: Oxford University Press.
- Mackey, A. and Gass, S. (2005) *Second Language Research: Methodology and Design*. Mahwah, NJ: Lawrence Erlbaum Associates.
- Mackenzie, N., and Knipe, S. (2006) 'Research dilemmas: Paradigms, methods and methodology', *Issues In Educational Research*, 16(2), pp. 193-205.
- Marton, W. (1977) 'Foreign vocabulary learning as problem No. 1 of language teaching at the advanced level', *Interlanguage Studies Bulletin*, 2, pp. 33-57.
- Maynard, M. (1994) 'Methods, practice and epistemology: the debate about feminism and research', in Maynard, M. and Jane Purvis, J. (eds.), *Researching women's lives from a feminist perspective*. London: Taylor and Francis, pp.10-27.
- McCarthy, M. (1990) *Vocabulary*. Oxford: Oxford University Press.

- McEnery, T. and Wilson, A. (1997) 'Teaching and Language Corpora (TALC)', *ReCALL*, 9(1), pp. 5-14.
- McEnery, T., and Wilson, A. (2001) *Corpus linguistics: An introduction*. Edinburgh: Edinburgh University Press.
- McEnery, T., Xiao, R. and Tono, Y. (2006) *Corpus-based Language Studies: An Advanced Resource Book*. London/New York: Routledge.
- McEnery, T. and Xiao, R. (2011) 'What corpora can offer in language teaching and learning', in Hinkel, E. (ed.) *Handbook of Research in Second Language Teaching and Learning*. vol. 2. London/New York: Routledge, pp. 364-380.
- Meara, P. (1980) 'Vocabulary acquisition: A neglected aspect of language learning', *Language Teaching*, 13, pp. 221-246.
- Meara, P. (1982) 'Word associations in a foreign language', *Linguistics Circular*, 11, pp. 29-38.
- Meara, P. (1990) 'A note on passive vocabulary', *Second Language Research*, 6, pp. 150-154.
- Meara, P. (1996). *The vocabulary knowledge framework*. Unpublished manuscript. Available at: <http://www.lognostics.co.uk/vlibrary/meara1996c.pdf> (Accessed: 21st April, 2009)
- Meara, P. (2009) *Connected Words: Word Associations and Second Language Vocabulary Acquisition*. Amsterdam: John Benjamins.
- Mertens, D. M. (1998) *Research methods in education and psychology: Integrating diversity with quantitative and qualitative approaches*. London: Sage.
- Messick, S. (1989) 'Validity', in Linn, R. L. (ed.) *Educational measurement* (3rd ed.). New York: Macmillan, pp. 13-103.
- Messick, S. (1994) 'The interplay of evidence and consequences in the validation of performance assessments', *Educational Researcher*, 23(2), pp. 13-23.
- Messick, S. (1995) 'Validity of psychological assessment. Validation of inferences from persons' responses and performances as scientific inquiry into score meaning', *American Psychologist*, 50, pp. 741-749.
- Meunier, F. (2002) 'The Pedagogical Value of Native and Learner Corpora in EFL Grammar Teaching', in Granger, S., Hung, J. and Petch-Tyson, S. (eds.) *Computer Learner Corpora, Second Language Acquisition and Foreign Language Teaching*. Amsterdam: John Benjamins, pp. 116.
- Milton, J. (2009) *Measuring Second Language Vocabulary Acquisition*. Bristol, UK; Buffalo N.Y.: Multilingual Matters.
- Mitchell, R. and Myles, F. (2004) *Second language learning theories*. London: Hodder Arnold.
- Moon, R. (1997) 'Vocabulary connections: Multi-word items in English', in N. Schmitt, N., and McCarthy, M. (eds.) *Vocabulary: Description, acquisition and pedagogy*. Cambridge: Cambridge University Press, pp. 40-63.

- Moon, R. (1998) *Fixed Expressions and Idioms in English: A Corpus-Based Approach*. Oxford: Oxford University Press.
- Mukherjee, J. (2006) 'Corpus Linguistics and Language Pedagogy: The State of the Art and Beyond', in Braun, S., Kohn K. and Mukherjee, J. (eds.) *Corpus Technology and Language Pedagogy: New Resources, New Tools, New Methods*. Frankfurt: Peter Lang, pp. 5-24.
- Nattinger, J. R. (1980) 'A lexical phrase grammar for ESL', *TESOL Quarterly* .14(3), pp. 337-344.
- Nattinger, J. R. (1988) 'Some current trends in vocabulary teaching', in Carter, R. and McCarty, M. (eds.) *Vocabulary and language teaching*. New York: Longman pp. 62-82.
- Nation, I. S. P. (1990) *Teaching and learning vocabulary*. New York: Newbury House.
- Nation, I. S. P. (2001) *Learning vocabulary in another language*. Cambridge: Cambridge University press.
- Nation, I.S.P., and Webb, S. (2011) *Researching and Analysing vocabulary*. Boston: Heinle Cengage Learning.
- Nesselhauf, N. (2003) 'The use of collocations by advanced learners of English and some implications for teaching', *Applied Linguistics*, 24(2), pp. 223-242.
- Nesselhauf, N. (2005) *Collocations in a learner corpus*. Amsterdam: John Benjamins.
- Nishina, Y. (2008) 'Lexical study based on parallel corpora, DDL and Moodle', in de Cássia Veiga Marriott, R. (ed.) *Handbook of Research on E-Learning Methodologies for Language Acquisition*. IGI Global, pp. 204-218.
- Norris, J., and Ortega, L. (2000) 'Effectiveness of L2 instruction: A research synthesis and quantitative meta-analysis', *Language Learning*, 50, pp. 417-528.
- Noor, H. and Adubaib, A. (2011). Strategies Used in Producing English Lexical Collocations by Saudi EFL Learners. Paper Presented at the 1st International Conference on Foreign Language Teaching and Applied Linguistics, Sarajevo. Available at: http://eprints.ibu.edu.ba/87/1/FLTAL%202011%20Proceed%20Book%201_p574-p595.pdf (Accessed: 20 December 2012).
- Norusis, M.J. (2006) *SPSS 15.0 Guide to Data Analysis*. Upper Saddle River, N.J.: Prentice Hall.
- Oppenheim, N. (2000) 'The importance of recurrent sequences for nonnative speaker fluency and cognition', in Riggenbach, H. (ed.) *Perspectives on fluency*. Ann Arbor, MI: University of Michigan Press, pp. 220-240.
- O'Sullivan, I. (2007) 'Enhancing a Process-Oriented Approach to Literacy and Language Learning: The Role of Corpus Consultation Literacy', *ReCALL*, 19(03), pp. 269-286.

- Özgül, B. and Abdülkadir, Ç (2012) 'Teaching vocabulary through collocations in EFL Classes: The case of Turkey', *International Journal of Research Studies in Language Learning*, 1(1): pp. 21-32.
- Pallant, J. (2007) *SPSS survival manual: A step-by-step guide to data analysis using SPSS version 15*. Maidenhead, Berkshire, England: McGraw-Hill Education.
- Palmberg, R. (1987) 'Patterns of vocabulary development in foreign languages learners', *Studies in Second Language Acquisition*, 9, pp. 202-221.
- Paribakht, T.S., and Wesche, M. (1997) 'Vocabulary enhancement activities and reading for meaning in second language vocabulary acquisition', in Coady, J. and Huckin, T. (eds.) *Second language vocabulary acquisition: A rationale for pedagogy*. Cambridge: Cambridge University Press, pp. 174- 200.
- Pawley, A. and Syder, F.H. (1983) 'Two puzzles for linguistic theory: Nativelike selection and nativelike fluency', in Richards, J.C. and Schmidt, R.W. (eds.) *Language and Communication*. London: Longman, pp. 191-225.
- Peters, A.M. (1983) *Units of language acquisition*. Cambridge: Cambridge University press.
- Peters, E. (2012) 'Learning German formulaic sequences: the effect of two attention-drawing techniques', *Language Learning Journal*, 40, pp. 65-79.
- Peters, E. (2014) 'The effects of repetition and time of post-test administration on EFL learners' form recall of single words and collocations', *Language Teaching Research*, 18(1), pp.75-94.
- Peters, E. (2015) 'The learning burden of collocations: The role of interlexical and intralexical factors', *Language Teaching Research*, pp. 1-26.
- Porte, G. K. (2002) *Appraising Research in Second Language Learning: A Practical Approach to Critical Analysis of Quantitative Research*. Amsterdam: John Benjamins.
- Razali, N.M and Wah, Y.B. (2011) 'Power comparisons of Shapiro-Wilk, Kolmogorov-Smirnov, Lilliefors and Anderson-Darling tests', *Journal of Statistical Modeling and Analytics*, 2 (1), PP. 21-33.
- Read, J. (1988) 'Measuring the vocabulary knowledge of second language Learners', *RELC Journal*, 19, pp. 12-25.
- Read, J. (2000) *Assessing vocabulary*. Cambridge: Cambridge University Press.
- Richards, J.C. (1976) 'The role of vocabulary teaching', *TESOL Quarterly*, 10 (1), pp. 77- 89.
- Ringbom, H. (1987) *The role of the first language in foreign language learning*. Clevedon, England: Multilingual Matters.
- Römer, U. (2009) Corpora and language teaching, in Ludeling, A. and Kytö, M. (eds.) *Corpus linguistics an international book volume 2*, pp. 112- 131.

- Salhi, H. (2010) 'Small Parallel Corpora in an English-Arabic Translation Classroom: No Need to Reinvent the Wheel in the Era of Globalization', in Shiyab, S.M., Rose, M.G., House, J. and Duval, J. (eds.) *Globalisation and Aspects of Translation*. Newcastle: Cambridge Scholars Publishing, pp. 53-67.
- Schmied, J. (2006) 'New Ways of Analysing ESL on the WWW with WebCorp and WebPhraseCount', in Renouf, A., and Kehoe, A. (eds.) *The Changing Face of Corpus Linguistics*. Amsterdam: Rodopi, pp. 309-324.
- Schmidt, R., and Frota, S. (1986) 'Developing basic conversational ability in a second language: A case study of an adult learner of Portuguese' in Day, R. R. (ed.) *Talking to learn: Conversation in second language acquisition*. Rowley, MA: Newbury House, pp. 237-324.
- Schmidt, R. (1990) 'The Role of Consciousness in Second Language Learning', *Applied Linguistics*, 11(2), pp. 129-158.
- Schmidt, R. (1993a) 'Awareness and second language acquisition', *Annual Review of Applied Linguistics*, 13, pp. 206-226.
- Schmidt, R. (1993b) 'Consciousness, learning and interlanguage pragmatics', in Kasper, G. and Blum-Kulka, S. (eds.) *Interlanguage pragmatics*. Oxford: Oxford University Press, pp. 21-42.
- Schmidt, R. (1994) 'Deconstructing consciousness: In search of useful definitions for Applied Linguistics', *AILA Review*, 11, pp. 11-26.
- Schmidt, R. (2001) 'Attention', in Robinson, P. (ed.) *Cognition and second language instruction*. Cambridge: Cambridge University Press, pp. 3-32.
- Schmitt, N. (1997) 'Vocabulary Learning Strategies', N. Schmitt & M. McCarthy (Eds.), *Vocabulary: Description, acquisition and pedagogy* (eds). Cambridge University Press, pp. 199-227.
- Schmitt, N. (2000) *Vocabulary in language teaching*. Cambridge: Cambridge University press.
- Schmitt, N. (2008) 'Instructed Second Language Vocabulary Learning', *Language Teaching Research*, 12(3), pp. 329-363.
- Schmitt, N. (2010) *Researching Vocabulary. A Vocabulary Research Manual*. Basingstoke: Palgrave Macmillan.
- Schmitt, N., and Carter, R. (2004) 'Formulaic sequences in action: An introduction', in Schmitt, N. (ed.) *Formulaic sequences: Acquisition, processing, and use*. Amsterdam: John Benjamins, pp. 1-22.
- Schmitt, N., and Schmitt, D.R. (1993) 'Identifying and Assessing Vocabulary Learning Strategies', *ELT Journal Volume*, 49 (2), pp. 133- 143.
- Schmitt, N., and Schmitt, D.R. (1995) 'Vocabulary notebooks: theoretical underpinnings and practical suggestions', *Thai TESOL Bulletin*, 5 (4), pp. 27-33.

- Schmitt, N., Schmitt, D., and Clapham, C. (2001) 'Developing and exploring the behaviour of two new versions of the Vocabulary Levels Test', *Language Testing*, 18, pp. 55-88.
- Selinker, L. (1992) *Rediscovering Interlanguage*. New York: Longman.
- Selinker, L., and Lakshmanan, U. (1992) 'Language transfer and fossilization: The multiple effects principle', in Gass, S. and Selinker, L (eds.), *Language transfer in language learning*. Amsterdam: Benjamins, pp. 197-216.
- Sheen, R. (1996) 'The advantage of exploiting contrastive analysis in teaching and learning a foreign language', *International Review of Applied Linguistics*, 24, pp. 183-97.
- Shehadeh, A. (1999) 'Insights into learner output', *English Teaching Forum*, 37(4), pp. 2-6.
- Sheskin, D. J. (2004) *Handbook of Parametric and Nonparametric Statistical Procedures, 3rd edition*. Boca Raton: Chapman & Hall/CRC.
- Shohamy, E. (1993) 'The power of tests: The impact of language tests on teaching and learning', *NFLC Occasional Paper*. Washington, DC: National Foreign Language Center.
- Sinclair, J. M. (1987) 'Collocation: a progress report', in Steele, R. and Threadgold, T. (eds.) *Language topics: Essays in honour of Michael Halliday* Vol. 2. Amsterdam: John Benjamins, pp. 319-331.
- Sinclair, J. (1991) *Corpus, concordance, collocation*. Oxford: Oxford University Press.
- Sinclair, J. (1996) 'Preliminary Recommendations on Corpus Typology', EAGLES Document EAG-TCWG-CTYP/P, <http://www.ilc.cnr.it/EAGLES96/corpusstyp/corpusstyp.html>
- Siyanova, A., and Schmitt, N. (2008) 'L2 learner production and processing of collocation: A multi-study perspective', *Canadian Modern Language Review*, 64(3), pp. 429-458.
- Soars, L., and Soars, J. (2012) *New Headway plus Intermediate (Special Edition)*. Oxford: Oxford University Press.
- Sonbul, S. (2012) *The Interface of Explicit/Implicit Lexical Knowledge: Acquisition and Processing of Collocations by Natives and Nonnatives*. Unpublished PhD thesis: University of Nottingham.
- Sonbul, S., and Schmitt, N. (2010) 'Direct teaching of vocabulary after reading: Is it worth the effort?', *ELT Journal*, 64(3), pp. 253-260.
- Souza Hodne, L. (2009) *Collocations and teaching: Investigating word combinations in two English textbooks for Norwegian upper secondary school students*. Unpublished PhD thesis: Toronto University.
- Stern, H. (1990) 'Analysis and experience as variables in second language pedagogy', in Harley, B., Allen, P. Cummins, J. and Swain, M. (eds.) *The development of*

- second language proficiency*. Cambridge: Cambridge University Press, pp. 93-109.
- Stubbs, M. (1995) 'Collocations and semantic profiles: On the cause of the trouble with quantitative studies', *Functions of language*, 2(1), pp. 23-55.
- Sun, Y.C., and Wang, L.Y. (2003) 'Concordancers in the EFL classroom: Cognitive approaches and collocation difficulty', *Computer Assisted Language Learning*, 16, pp. 83-94.
- Sunderman G., and Kroll J.F. (2006) 'First language activation during second language lexical processing: An investigation of lexical form, meaning, and grammatical class', *Studies in Second Language Acquisition*, 28, pp. 387-422.
- Swan, M. (1997) 'The influence of the mother tongue on second language vocabulary acquisition and use', in Schmitt, N and McCarthy, M. (eds.), *Vocabulary: Description, acquisition and pedagogy*. Cambridge: Cambridge University Press, pp. 156-180.
- Swain, M. (1985) 'Communicative competence Some roles of comprehensible input and comprehensible output in its development', in Gass, S. and Madden, C. (eds.) *Input and Second Language Acquisition*. Rowley, MA Newbury House.
- Swain, M. (1993) 'The output hypothesis: Just speaking and writing aren't enough', *Canadian Modern Language Review*, 50, pp. 158-164.
- Swain, M. (1995) 'Three functions of output in second language learning', in Cook, G. and Seidlhofer, B. (eds.), *Principle and practice in applied linguistics: Studies in honour of Henry G. Widdowson*. Oxford: Oxford University Press, pp. 125-144.
- Swain, M. (2000) 'The output hypothesis and beyond: mediating acquisition through collaborative dialogue', in Lantolf, C. (ed.) *Sociocultural Theory and Second Language Learning*. Oxford: Oxford University Press, pp. 97-114.
- Swain, M. (2005) 'The output hypothesis: theory and research', in Hinkel, E. (ed.) *Handbook of research in second language teaching and learning. Vol I*. New York: Routledge, pp. 471- 483.
- Swain, M. and S. Lapkin. (1995) 'Problems in output and the cognitive processes they generate: A step towards second language learning,' *Applied Linguistics*, 16, pp. 371-391.
- Szudarski, P. (2012) 'Effects of meaning- and form- focused instruction on the acquisition of verb- noun collocations in L2 English', *Journal of Second Language Teaching and Research*, 1(2), pp. 3-37.
- Takala, S. (1984) *Evaluation of Students' knowledge of Vocabulary in the Finnish Comprehensive School* (Reports of the Institute of Educational Research, No. 350). Finland: Jyväskylä.

- Talai, T. and Fotovatnia, Z. (2012) 'Data-Driven Learning: A Student-centred Technique for Language Learning', *Theory and Practice in Language Studies*, 2(7), pp. 1526-1531.
- Tanskanen, S. K. (2006) *Collaborating towards coherence*. Amsterdam/New York: Benjamins.
- Taylor, C. V. (1983) 'Vocabulary for education in English', *Word Language English*, 2(2), pp. 100- 104.
- Teichroew, M. (1982) 'Receptive versus productive vocabulary: A survey', *Interlanguage Studies Bulletin*, 6(2), pp. 5-33.
- Thurstun, J. and Candlin, C. N. (1998) 'Concordancing and the Teaching of the Vocabulary of Academic English', *English for Specific Purposes*, 17(3), pp. 267-280.
- Tsagari, D. and Floros, G. (eds.) (2013) *Translation in Language Teaching and Assessment*. Newcastle upon Tyne: Cambridge Scholars Publishing.
- Tsai, K.J. (2011) *A Multidimensional Investigation of a Data-Driven Approach to Learning Collocations*. Unpublished PhD thesis. University of Bristol.
- Tsui, A. B. M. (2004) 'What Teachers Have Always Wanted to Know - and How Corpora Can Help', in Siclair, J. (ed.) *How to Use Corpora in Language Teaching*. Amsterdam: John Benjamins, pp. 39-61.
- Van der Meer, G., (1998) 'Collocations as one particular type of conventional word combination. Their definition and character', *EURALEX '98 Proceedings*. University of Liège, Liège. pp. 313-322.
- Van Lier, L. (2000) 'From Input to Affordance: Social-Interactive Learning from an Ecological Perspective', in Lantolf, J. P. (ed.) *Sociocultural Theory and Second Language Learning*. Oxford: Oxford University Press, pp. 245.
- Vermeer, A. (2001) 'Breadth and depth of vocabulary in relation to L1/L2 acquisition and frequency of input', *Applied Psycholinguistics* 22(2), pp. 217-234.
- Waring, R. (1997) 'A comparison of the receptive and productive vocabulary sizes of some second language learners', *Immaculata* (Notre Dame Seishin University), 1, pp. 53-68.
- Waring, R., and Takaki, M. (2003) 'At what rate do learners learn and retain new vocabulary from reading a graded reader?', *Reading in a Foreign Language*, 15, pp.1-27. Available at: <http://nflrc.hawaii.edu/rfl/October2003/waring/waring.html> (Accessed: 21 August 2014).
- Webb, S. (2005) 'Receptive and productive vocabulary learning: The effects of reading and writing on word knowledge,' *Studies in Second Language Acquisition*. 27, pp. 33-52.
- Webb, S., and Kagimoto, E. (2009) 'The effects of vocabulary learning on collocation and meaning', *TESOL Quarterly*, 43, pp. 55-77.

- Webb, S., and Kagimoto, E. (2011) 'Learning collocations: Do the number of collocates, position of the node word, and synonymy affect learning?', *Applied Linguistics*, 32(3), pp. 259-276.
- Webb, S., Newton, J., and Chang, C-S. A. (2013) 'Incidental Learning of Collocation', *Language Learning*, 63 (1), pp. 91-120.
- West, M. (1953) *A general service list of English words*. London: Longman.
- Widdowson, H. G. (1978) *Teaching Language as Communication*. Oxford: Oxford University Press.
- Widdowson, H. G. (1980) 'Models and fictions', *Applied Linguistics*, 1(2), pp. 165-170.
- Widdowson, H. G. (1992) 'Language study and language learning', in van Essen, A. and Burkart, E. I. (eds.) *Homage to W.R. Lee: Essays in English as a Foreign Language*. Dordrecht: Foris, pp. 103-111.
- Widdowson, H. G. (1998) 'Context, community and authentic language', *TESOL Quarterly*, 32, pp. 705-716.
- Widdowson, H. G. (2000) 'On the Limitations of Linguistics Applied', *Applied Linguistics*, 21(1), pp. 3-25.
- Wolter, B., and Gyllstad, H. (2011) 'Collocational links in the L2 mental lexicon and the influence of L1 intralexical knowledge', *Applied Linguistics*, 32(4), pp. 430-449.
- Wolter, B., and Gyllstad, H. (2013) 'Frequency of input and L2 collocational processing: A comparison of congruent and incongruent collocations', *Studies in Second Language Acquisition*, 35, pp. 451-482.
- Wray, A. (1999) 'Formulaic language in learners and native speakers', *Language Teaching*, 32 (4), pp. 213-231.
- Wray, A. (2000) 'Formulaic sequences in second language teaching: Principle and practice', *Applied Linguistics*, 21(4), pp. 463-489.
- Wray, A. (2002) *Formulaic language and the lexicon*. Cambridge: Cambridge University Press.
- Wray, A. and Perkins, M.R. (2000) 'The functions of formulaic language: an integrated model', *Language and Communication*, 20(1), pp. 1-28.
- Yamashita, J., and Jiang, N. (2010) 'L1 influence on the acquisition of L2 collocations: Japanese ESL users and EFL learners acquiring English collocations', *TESOL Quarterly*, 44, pp. 647-668.
- Zimmerman, C. B. (1997) 'Do reading and interactive vocabulary instruction make a difference? An empirical study', *TESOL Quarterly*, 31, pp. 121-140.

Dictionaries

Oxford English Dictionary (online)

<http://www.oxforddictionaries.com/definition/english/collocation>

Longman Dictionary of Contemporary English

<http://www.ldoceonline.com/dictionary/collocationA>

Appendix A: Vocabulary level tests (K2 & K3)

This is a vocabulary test. You must choose the right word to go with each meaning. Write the number of that word next to its meaning. Here is an example.

- | | | |
|---|----------|----------------------------------|
| 1 | business | |
| 2 | clock | _____ part of a house |
| 3 | horse | _____ animal with four legs |
| 4 | pencil | _____ something used for writing |
| 5 | shoe | |
| 6 | wall | |

You answer it in the following way.

- | | | |
|---|----------|-------------------------------------|
| 1 | business | |
| 2 | clock | <u>6</u> part of a house |
| 3 | horse | <u>3</u> animal with four legs |
| 4 | pencil | <u>4</u> something used for writing |
| 5 | shoe | |
| 6 | wall | |

Some words are in the test to make it more difficult. You do not have to find a meaning for these words. In the example above, these words are business, clock, and shoe.

If you have no idea about the meaning of a word, do not guess. But if you think you might know the meaning, then you should try to find the answer.

Version 2 The 2,000 word level

1 copy
 2 event _____ end or highest point
 3 motor _____ this moves a car
 4 pity _____ thing made to be like
 5 profit another
 6 tip

1 accident
 2 debt _____ loud deep sound
 3 fortune _____ something you must pay
 4 pride _____ having a high opinion of
 5 roar yourself
 6 thread

1 coffee
 2 disease _____ money for work
 3 justice _____ a piece of clothing
 4 skirt _____ using the law in the right
 5 stage way
 6 wage

1 clerk
 2 frame _____ a drink
 3 noise _____ office worker
 4 respect _____ unwanted sound
 5 theater
 6 wine

1 dozen
 2 empire _____ chance
 3 gift _____ twelve
 4 opportunity _____ money paid to the
 5 relief government

1 admire
 2 complain _____ make wider or longer
 3 fix _____ bring in for the first time
 4 hire _____ have a high opinion of
 5 introduce someone
 6 stretch

1 arrange
 2 develop _____ grow
 3 lean _____ put in order
 4 owe _____ like more than something
 5 prefer else
 6 seize

1 blame
 2 elect _____ make
 3 jump _____ choose by voting
 4 manufacture _____ become like water
 5 melt
 6 threaten

1 ancient
 2 curious _____ not easy
 3 difficult _____ very old
 4 entire _____ related to God
 5 holy
 6 social

1 bitter
 2 independent _____ beautiful
 3 lovely _____ small
 4 merry _____ liked by many people
 5 popular

Version 2 The 3,000 word level

- 1 bull
- 2 champion _____ formal and serious manner
- 3 dignity _____ winner of a sporting event
- 4 hell _____ building where valuable objects are shown
- 5 museum
- 6 solution

- 1 blanket
- 2 contest _____ holiday
- 3 generation _____ good quality
- 4 merit _____ wool covering used on beds
- 5 plot
- 6 vacation

- 1 comment
- 2 gown _____ long formal dress
- 3 import _____ goods from a foreign country
- 4 nerve
- 5 pasture _____ part of the body which carries feeling
- 6 tradition

- 1 administration
- 2 angel _____ group of animals
- 3 frost _____ spirit who serves God
- 4 herd _____ managing business and affairs
- 5 fort
- 6 pond

- 1 atmosphere
- 2 counsel _____ advice
- 3 factor _____ a place covered with grass
- 4 hen _____ female chicken
- 5 lawn
- 6 muscle

- 1 abandon
- 2 dwell _____ live in a place
- 3 oblige _____ follow in order to catch
- 4 pursue _____ leave something permanently
- 5 quote
- 6 resolve

- 1 assemble
- 2 attach _____ look closely
- 3 peer _____ stop doing something
- 4 quit _____ cry out loudly in fear
- 5 scream
- 6 toss

- 1 drift
- 2 endure _____ suffer patiently
- 3 grasp _____ join wool threads together
- 4 knit _____ hold firmly with your hands
- 5 register
- 6 tumble

- 1 brilliant
- 2 distinct _____ thin
- 3 magic _____ steady
- 4 naked _____ without clothes
- 5 slender
- 6 stable

- 1 aware
- 2 blank _____ usual
- 3 desperate _____ best or most important
- 4 normal _____ knowing what is happening
- 5 striking
- 6 supreme

Appendix B: Target collocations, frequency and MI scores

collocations	BNC Raw Frequency	MI score
Open air	323	3.31
Key areas	190	3.44
Vast majority	857	8.30
Immediate future	137	4.08
Recent years	2791	5.05
Hard copy	175	4.50
Round trip	106	3.66
Domestic violence	105	5.54
Careful attention	91	4.47
Common sense	1003	5.27
Vast numbers	99	4.99
Broad agreement	57	3.83
Heavy emphasis	50	4.06
Ill health	326	5.54
Naked eye	149	7.14
Fine arts	215	5.69
Steady progress	67	5.61
Fresh start	154	4.05
Huge success	74	3.60
Careful planning	92	4.36
Early summer	230	3.35
Hard facts	65	3.22
Heavy losses	85	5.34
Low risk	133	3.48
Instant coffee	63	6.37
Poor condition	122	4.06
Heavy traffic	153	5.40
Long tradition	171	3.33
Safe return	56	3.01
Good faith	357	3.83

Appendix C: Clued-recall pilot test

The following items aim at assessing your knowledge of a number of adjective/noun combinations that usually go together (collocations). For each item, please fill in the blank with the word that completes the phrase and that begins with the letter provided. Please do not make random guesses. If you do not know the answer, please skip the item and leave it blank.

Here is an example:

Great deal of our flying is done by habit and this can lead to (s _____ problems) if anything goes wrong on the first flight in a strange machine.

You should answer in the following way:

Great deal of our flying is done by habit and this can lead to (serious problems) if anything goes wrong on the first flight in a strange machine.

Now, please start with the first below!

2. (R----- years) have witnessed changes in the overall structure of art education course.
3. If all the (f _____ water) was divided among the population, everybody would get around 40 million gallons each.
4. Extremely (s _____ water) may not contain sufficient calcium to maintain growth of the skeleton.
5. She was waiting in the outside lane of a (b _____ road) in St Helens to turn right when the accident happened.
6. There were few role models in the (w _____ society) which suggested that our creativity was so diverse and has such wonderful potential.
7. If you are experiencing difficulty in any of the (k _____ areas) of life listed below, you are not experiencing whole health.
8. It was concluded that the (n _____ effect) of all the plans outlined would be to affect employment by not more than 10%.
9. Simons's argument appeals to (c _____ sense), but we need to bear in mind an alternative.
10. It will be obvious that I am using the word liberal here in a very (b _____ sense).
11. The building is conveniently situated in the (h _____ street) of Stetchworth, just outside Newmarket, in Suffolk.
12. Drying clothes in the (o _____ air) if possible is preferable to hanging them indoors.

13. Conventions are here understood in a very (**n**_____ **sense**) in which they are solutions to co-ordination problems.
14. The pyramids acted as great sounding boards in the (**t**_____ **air**) of the desert.
15. She was tall and gracious, with a (**d**_____ **voice**) and a strikingly vivid personality.
16. Falling off a (**t**_____ **building**) is, after all, much more dangerous than falling off a low wall.
17. The castle was faintly lit by the (**s**_____ **light**) of the moon.
18. A faint glimmer of (**p**_____ **light**) was rising in the midnight-blue sky.
19. He has very (**f**_____ **hair**), resting timidly and lifelessly on his skull.
20. Under his grime, Oliver turned as pale as his own (**p**_____ **hair**).
21. Keith's (**i**_____ **future**) is secure while he has his parents to help look after him.
22. Some of these weaknesses might have been avoided if more (**c**_____ **attention**) had been given to commercial needs.
23. (**H**_____ **technology**) private care is available to the rich and inadequate public care to the poor.
24. Some of the Sicilian interiors glow with (**r**_____ **colour**) like the greatest of the Byzantine ones.
25. It appears that knowledge of the (**i**_____ **effects**) of tobacco has already led to a modification in smoking behaviour among older men.
26. The (**v**_____ **majority**) of people here are decent people, friendly, and they'll be civil enough to you.
27. If he was caught, Ewan would be in (**d**_____ **trouble**).
28. That official now holds a very prestigious (**s**_____ **post**) in the Education Department.
29. Some music has a (**r**_____ **variety**) of movement; other music depends on very little change.
30. The seriousness of the (**l**_____ **weight**) is often denied by the anorexic patient.
31. I hope we can get back together and make a (**f**_____ **start**).
32. His finger traced a (**n**_____ **band**) of blue shading on the map which was laid out across the table.

33. There would have been (**h**_____ **traffic**) at this hour and news of the roadblocks would have spread quickly throughout the country.
34. When the kettle boiled, she made two cups of (**i**_____ **coffee**).
35. Otley made me some hot milk and honey and helped himself to a cheese and pickle sandwich and a mug of (**s**_____ **coffee**).
36. Vass smiled a (**s**_____ **smile**), as warm as a winter's morning.
37. Put them in the (**h**_____ **seat**) and question them to find out.
38. In my own case no one can be expected to remember every (**f**_____ **detail**) of matters they have undertaken.
39. Military victory in the Civil War had been achieved by a (**h**_____ **emphasis**) on combined political and military control.
40. A shadow of apprehension crossed her face, immediately replaced by a welcoming and (**w**_____ **smile**).
41. His office, which works in (**f**_____ **arts**), craft, film and video and photography, may offer straight advice on projects.
42. Persuading others to accept the (**h**_____ **facts**) of life is not usually a very popular job.
43. A person may buy something in (**g**_____ **faith**), but may find out afterwards that the seller had no title to it.
44. He took out his knife and cut away some of the (**d**_____ **wood**) from the rose trees.
45. There is a (**l**_____ **tradition**) of migration, mostly to the United States, as people seek a better life and an escape from poverty.
46. Beer might be seen as a (**s**_____ **drink**) in some countries and as an alcoholic drink in others.
47. The Ukraine suffered particularly (**h**_____ **losses**) during the 1930s as there was a devastating famine during the early part of the decade.
48. Frankfurt and Regensburg had enjoyed a (**h**_____ **reputation**) for its technical skill.
49. The dog's owners have offered a reward for its (**s**_____ **return**).
50. The judge stressed that the girl would retain (**c**_____ **contact**) with her 64-year-old father.
51. One in six women are believed to be victims of (**d**_____ **violence**).
52. Eleanor picked up the (**f**_____ **flowers**) and began trimming their stems.

53. The prisoners are encouraged to confront the (t_____ **nature**) of their crimes and themselves.
54. Two squirrels died in transit, the rest were in (p_____ **condition**) on arrival.
55. Although the joint statement indicated (b_____ **agreement**), there were some differences over how to deal with the crisis.
56. The new album Change Everything is due for release in (e_____ **summer**).
57. The future of British coal lies in a high-tech industry employing (s_____ **numbers**) of people and producing coal at a competitive price.
58. We all collect books in (v_____ **numbers**), often leaving them unread for year.
59. When I went to that stress management course we were told to use (p_____ **resources**) like deep breathing.
60. Each business will need to have a solid safety net of (l_____ **risk**) performance areas.
61. We can't make any more hours in the day, but (c_____ **planning**) can allow us to use this time-saving piece of equipment as efficiently as possible.
62. It is said to be visible with the (n_____ **eye**), but I have never been able to confirm this, though with binoculars it is very easy indeed.
63. The symptoms include pounding head, (d_____ **mouth**), stiff face muscles, sweaty palms, tension in the neck and shoulders.
64. The full size terminal can store up to 37 pages in memory for review or printing to (h_____ **copy**).
65. That bullied student can't be the only young girl in the school with a full set of (f_____ **teeth**). Can she?
66. Tickets for this historic run are priced at 35.00 for the (r_____ **trip**), with the return trip to London on Friday night.
67. Forcing herself to walk slowly back to the apartment, she turned on the TV to catch the final (c_____ **stages**) of the race.
68. It was as if I'd had a (b_____ **dream**), but it was one I couldn't wake from.
69. Mayflies and several varieties of caddis disappear from acid streams, and the freshwater shrimp becomes scarce in (s_____ **water**).
70. He was on the point of taking his own life because he was depressed over his (i_____ **health**).
71. She felt sure the campaign would be a (h_____ **success**).

72. The warning tonight is that, with more (s_____ **weather**) forecast, it'll get worse before it gets better.
73. Difficulties at home have already caused a (s_____ **fall**) in Japanese investment abroad.
74. We are making a slow but (s_____ **progress**), of which the country should be proud.

Thank you very much for your time dedicated to taking the test.

Appendix D: MC distracters

Appendix E: Reading passages and comprehension questions (samples)

Session 1

I don't believe that today's wonders are similar in kind to the wonders of the Ancient World. They were all buildings such as the Pyramids in Egypt, or other architectural structures in the **open air**. Over the past 100 years, we have seen amazing achievement in **key areas** such as technology and science. These are surely modern wonders.

It is everywhere. More than billion people use it, and the **vast majority** of people are now using online services. Their numbers increase by 100 million every year. In 1994 there were only a few hundred web pages. Today there are billions. It has revolutionized the way we live and work. However, we are still in the early days. Soon there will be more and more interactivity between the user and the website, and we will be able to give instruction using speech.

In 1969, Neil Armstrong stepped out of his space capsule onto the surface of the moon and made his famous statement: 'That's one small step for a man, one giant leap for mankind'. Since then there have been space probes to Mars, Jupiter, Saturn, and even to the sun. One day, a space observatory will study how the first stars and galaxies began. So far, it seems that we are alone in the universe. There are no signs yet that there is intelligent life outside our own solar system, but who knows what is going to happen in the **immediate future!**

Surly nothing has done more for the comfort and happiness of the human race than the advances in health care! How many millions of people have benefited from the humble

MC distracters	MI score	MC distracters	MI scores
Basic sense	0.97	Good arts	-1.39
Special sense	0.46	Best arts	-0.32
Certain sense	0.72	Great arts	-1.92
Future years	0.31	full success	-0.33
Earlier years	0.95	true success	-1.25
Following years	0.31	large success	-2.62
Closer future	-0.02	important emphasis	-0.22
Very future	-4.18	High emphasis	-0.65
Final future	-1.79	Real emphasis	-1.30
Local violence	-0.76	Direct eye	-0.12
Public violence	-0.08	Big eye	-0.08
Social violence	-0.61	real eye	0.67
basic copy	-0.35	low health	-1.31
Soft copy	0.96	Serious health	-1.75
Bad copy	-0.06	Bad health	0.93
total trip	-0.23	Clean start	-0.64
Full trip	0.24	Right start	-4.09
Complete trip	0.82	Great start	-0.51
wonderful air	0.20	Extra numbers	0.92
Free air	0.92	Big numbers	-0.16
Outside air	-0.10	Major numbers	-2.68
Greater majority	-0.02	Consistent progress	0.40
Big majority	-1.06	General progress	-0.05
Enormous majority	-0.27	firm progress	-0.96
Major attention	-0.44	Great planning	-1.20
Complete attention	-0.03	Major planning	0.81
Good attention	-0.95	Good planning	-0.25
Strong areas	-0.44	Total agreement	0.69
General areas	-0.26	wider agreement	-0.35
Primary areas	0.65	Full agreement	0.29
Large traffic	-0.96	Later tradition	0.91
High traffic	0.78	Big tradition	-0.59
Great traffic	-0.51	General tradition	0.05
Nice coffee	0.88	Main facts	0.41
Current coffee	-0.85	Clear facts	-0.40
Quick coffee	0.44	Strong facts	0.26
Little risk	0.42	true return	-1.63
Small risk	0.54	Full return	0.29
Large risk	-2.04	Good return	-0.39
Big faith	-0.03	Old condition	-1.14
Real faith	0.10	New condition	-0.59
Full faith	-0.54	Low condition	-0.89
Past summer	-1.29	High losses	0.62
Final summer	-1.05	Great losses	-0.25
Recent summer	-0.59	Vast losses	-0.08

aspirin? How many lives has penicillin saved in **recent years**? Average life expectancy worldwide has risen dramatically over the past 100 years, from about 47 years in 1900 to

about 77 years today. Such information are available electronically online or in **hard copy** documents.

We are a world on the move. Airlines carry more than 1.5 billion people to their destinations every year. It is estimated that, at any one time these days, there are many people travelling in aeroplanes on **round trips** (to their destinations and back) as the total number of people who travelled abroad in the whole of the nineteenth century (but I have no idea how they worked this out!).

It is true that some of these modern wonders are now commercialized, and there is greed, drug abuse and even **domestic violence** inside the families that is related to them. However, it is a competition in which almost every country in the world takes part and pays **careful attention** to. Every four years, for a brief moment, we see the world come together in peace and friendship. We feel hope again for the future of mankind.

In 1724, Jonathan Swift wrote, 'Whoever makes two blades of grass or two ears of corn grow where only one grew before serves mankind better than the whole race of politicians'. In Europe our farmers have done this. In 1709, whole villages in France died of hunger. Now in Europe, we can't eat all the food we produce. If only our politicians would have **common sense** and could find a way to share it with those parts of the world where there is famine.

The last wonder of the modern world is simply that we are still here. We have had nuclear weapons for over 50 years that could destroy the world, but we haven't used them to do it. This is surely the greatest wonder of all.

Based on the passage, circle the correct answer for each item:

1- The passage mainly discusses . . .

- a) Ancient world wonders.
- a) Space and air travel.
- b) Internet and technology.
- c) Miracles of the modern world.

2- Which of the following points is true...?

- a) There are signs that we are not alone in the universe.
- b) Nowadays, internet users can interact with websites by speech.
- c) In Europe, the amount of food produced is more than the amount needed.
- d) Despite of all the problems in the world, there is still hope for peace.

3- The author believes that...

- a) There are similarities between world wonders today and in the past.
- b) Scientist will start to investigate planets such as Mars and Jupiter in the future.

- c) The most important wonder of all is being alive despite the existence of deadly weapons.
- d) Politicians are more important than farmers.

Session 4

Prince Charles is often portrayed as bad-tempered and spoilt. There are stories that every day seven eggs are boiled for his breakfast so that he can find one that is just cooked the way he likes it. And his bath towel is folded over a chair in a particular way for when he gets out of his royal bath.

He has an enormous private staff secretaries, deputy secretaries, press officers, two butlers, housekeepers, two chefs, two chauffeurs, ten gardeners, and an army of porters, handymen, cleaners and maids. The **vast majority** of them are expected to get everything right. When HRH (His Royal Highness) feels they have performed their duties well, they are praised a royal memo delivered to them in **hard copy**. However, if they have made mistakes, they are called into his study and told off.

Charles is eccentric and he admits it. He talks to trees and plants. He pays **careful attention** to wildlife and wants to save it, but he also enjoys **open air** activities like hunting, shooting and fishing. He dresses for dinner. Even if he's eating alone. He's a great socializer. Poets, explorers, writers, broadcasters, philosophers, and politician all eat at his table. Arriving at Highgrove, his family home, on a Saturday afternoon, guests are entertained in the height of luxury. They are then sent on their way before lunch on Sunday.

The prince also entertains extravagantly at Sandringham, one of the Queen's homes, at least twice a year. There are picnic lunches on the beach, the lavish dinners with organic food from Highgrove. Conversation is lively, but the heir to the throne has to be very

careful in what he says, because he knows too well that anything he says in private may be repeated in public.

Together Charles and his wife Camilla perform royal duties, both at home and abroad. He attends 500 public engagements a year. He visits hospitals, youth groups, charities, and business conferences. He hosts receptions to welcome visiting heads of state and VIP's. He goes on **round trips** to various countries extensively, as an ambassador of the United Kingdom, representing the two **key areas** of trade and industry.

Charles works hard to promote greater understanding between different religions, and is patron of the Oxford Centre for Islamic Studies at the University of Oxford. He is also President of the Prince's Charities, which are active in promoting education, business, the environment, and opportunities for young people. The group raises £110 million annually. After the floods in Pakistan in 2010, the prince started the Pakistan Recovery Fund, to raise money for health, education, reconstruction and fight against violence (especially family and **domestic violence** against women).

Since his second marriage, Prince Charles has everything he wants expect, as Diana used to call it, 'the top job'. Yet despite not being on the throne, he has worked hard to accomplish so much. The Prince of Wales has his own food company, Duchy Originals. It originally sold biscuits, but in **recent years**, it has been expanded to become one of Britain's best-known and most successful organic brands, with over 200 different products, including food, drinks, and hair and body care products.

Charles, well-intentioned, hard-working, conservative, old fashioned and a person with **common sense**, continues to do his duty as he sees it, but he is no longer alone. Maybe not in the **immediate future**, but one day he will be king, and his darling Camilla will be HRH the Princess Consort.

Based on the passage, circle the correct answer for each item:

4- The passage mainly describes . . .

- b) Prince Charles' personality.
- d) Prince Charles' life style.
- e) Prince Charles' interests and duties.
- f) All of the above.

5- Which of the following points is false...?

- e) Prince Charles is the president of Oxford University.
- f) Prince Charles has a food company.
- g) Highgrove is Prince Charles family home.
- h) Prince Charles is very strict with his private staff.

6- Which of the following points is true...?

- a) Duchy Originals basically sells biscuits.
- b) Camilla is prince Charles' first wife.

- c) Prince Charles is the heir to the throne.
- d) Saving the wildlife is the one and only interest of prince Charles.

Appendix F: Experimental groups' worksheets (samples)

Group 1/ Worksheet 1

Now, please translate the following English sentences into Arabic. Try to provide suitable translation(s) for each of the bolded word combinations.

- 1. It's up to your own **common sense** and the weight at which you feel comfortable.

.....

- 2. In **recent years**, tourism has made an increasing impact on farming.

.....

- 3. The advantage of this method is that it costs nothing, at least in the **immediate future**.

.....

- 4. **Domestic violence** is a major social problem in Chile.

.....

- 5. Three departments try to keep the use of **hard copy** to a minimum by the use of online services.

6. The rest of the day is spent on a **round trip** of Lake Zell, before returning to Kaprun and the hotel.

7. At an **open air**_market, my daughter bought a jacket that had been reduced in price.

8. Homes in the **vast majority** of Detroit suburbs cost \$10,000–100,000.

9. It was necessary to pay **careful attention** to planning and design to keep the original character of the building.

10. These are **key areas** for attention, which are stressed in the project.

Group 1/ Worksheet 4

Now, please translate the following Arabic sentences into English. Try to provide suitable translation(s) for each of the bolded word combinations.

١. تزايد إنتاج الأرز في السنوات الأخيرة الماضية حتى باتت الدولة الآن مكتفية ذاتيا.

٢. قد تبدو بعض هذه الأمور عند الإشارة إليها مسألة منطوق بديهي/سليم، ولكنها ليست دائما كذلك.

٣. السيد كريس باتن: ليس لدي أي خطط لزيارة بلاكبول في المستقبل القريب/العاجل.

٤. يعتقد أن امرأة من كل ستة نساء يقعن ضحايا للعنف الأسري، ولكنهن يعانين في صمت.

٥. للقيام بهذه المهمة فإن الأقسام تقوم بتخزين المعلومات إما إلكترونيا أو في نسخ ورقية أو مطبوعة.

٦. إنها رحلة ذهاب و إياب قدرها ١٩٠ ميلا أي ما قدره ٦٠٠ ميل في الأسبوع تقريبا.

6. The rest of the day is spent on a **round trip** of Lake Zell, before returning to Kaprun and the hotel.

.....

7. At an **open air** market, my daughter bought a jacket that had been reduced in price.

.....

8. Homes in the **vast majority** of Detroit suburbs cost \$10,000–100,000

.....

9. It was necessary to pay **careful attention** to planning and design to keep the original character of the building.

.....

10. These are **key areas** for attention, which are stressed in the project.

.....

Group 2 / Worksheet 4

Translate the following Arabic sentences into English. Pay special attention to the translation of the bolded word combinations. Use the attached data to help you search for the combinations, translate them and notice the differences between their translations in English and Arabic.

١. تزايد إنتاج الأرز في السنوات الأخيرة الماضية حتى باتت الدولة الآن مكتفية ذاتيا.

.....

٢. قد تبدو بعض هذه الأمور عند الإشارة إليها مسألة **منطق بديهي** / سليم، ولكنها ليست دائما كذلك.

.....

٣. السيد كريس باتن: ليس لدي أي خطط لزيارة بلاكبول في **المستقبل القريب** العاجل.

.....

٤. يعتقد أن امرأة من كل ستة نساء يقعن ضحايا **للعنف الأسري**، ولكنهن يعانين في صمت.

.....

٥. للقيام بهذه المهمة فإن الأقسام تقوم بتخزين المعلومات إما إلكترونيا أو في **نسخ ورقية** أو مطبوعة.

.....

٦. إنها رحلة **ذهاب و إياب** قدرها ١٩٠ ميلا أي ما قدره ٦٠٠ ميل في الأسبوع تقريبا.

.....

٧. ويوجد هناك ملعبا جديدا للنتس و مسبحا في **الهواء الطلق** ومضمارا للركض.

.....

٨. يتناول الغالبية العظمى من الناس الذين يعانون من مشاكل في الوزن طعاما عاديا غير صحي.

٩. يعتبر وجود عدد من الفنادق الصغيرة أحد متع جزر السيشيلز، حيث يمكنك التمتع بخدمة ودودة واهتمام كبير من المالكين.

١٠. يجب أن يتضمن ذلك استثمارا طويل الأمد في المجالات الرئيسية للتدريب والتعليم والمواصلات.

Group 3/ Worksheet 1

Circle the most appropriate adjective to go with each noun. Use the attached data to help you understand, decide your answer or check your decision.

1. It's up to your own **sense** and the weight at which you feel comfortable.
a. basic b. common c. special d. certain
2. In **years** tourism has made an increasing impact on farming.
a. recent b. future c. earlier d. following
3. The advantage of this method is that it costs nothing, at least in the **future**.
a. closer b. very c. immediate d. final
4. **violence** is a major social problem in Chile.
a. Domestic b. Local c. Public. d. Social
5. Three departments try to keep the use of **copy** to a minimum by the use of online services.
a. basic b. hard c. soft d. bad

6. The rest of the day is spent on a **trip** of Lake Zell, before returning to Kaprun and the hotel.
a. total b. full c. round d. complete
7. At **air** market, my daughter bought a jacket that had been reduced in price.
a. an open b. a wonderful c. a free d. an outside
8. Homes in the **majority** of Detroit suburbs cost \$10,000–100,000.
a. greater b. big c. vast d. enormous
9. It was necessary to pay **attention** to planning and design to keep the original character of the building.
a. major b. complete c. good d. careful
10. These are **areas** for attention which are stressed in the project.
a. strong b. general c. primary d. key

Group 3 / Worksheet 4

Fill in the blank with the most appropriate adjective to go with each noun. Use the attached data to help you understand, decide your answer or check your decision.

1. In**years** rice production had increased until the country was now self-sufficient.
2. Some of these matters may seem **sense** when pointed out, but this is not always the case.
3. Mr. Chris Patten: I have no plans to visit Blackpool in the **future**.
4. One in six women are believed to be victims of**violence**, yet many are suffering in silence.
5. To perform this task the departments store a variety of information either electronically or in **copy**.
6. It is a **trip** of 190 miles so that is almost 600 miles a week.
7. There is a new tennis court, running track and an **air** swimming pool.
8. The **majority** of people with weight issues have been eating a normal unhealthy diet.
9. One of the delights of the Seychelles is the number of small hotels, where you can enjoy friendly service and the **attention** of the owners.

10. That must include long-term investment in areas of training, education, and transport.

Appendix G: Bilingual corpus data (samples)

Group 2/Session 1

English text	النص العربي
hopes of people everywhere. For its first 45 years, the United Nations lived in the grip of globalization and governance. In the early years of the United Nations, the General	الناس وتحقيق آمالهم في كل مكان، وعليها أن تقوم بذلك. ولقد عاشت الأمم المتحدة طيلة سنواتها الخمس والأربعين الأولى في قبضة العولمة والحكم. خلال سنوات الأمم المتحدة الأولى، كان
that significant progress has been made in recent years. Advisory Service on	التقدم الهام الذي أحرز في هذا المجال في الأعوام القريبة الماضية. الدائرة الاستشارية المعنية
was proud to promote, took too many years and cost too many lives.	، الذي تعتز الأمم المتحدة بالدعوة له، قد استغرق سنوات طويلة وأودى بأرواح كثيرة للغاية.
illiterate. Of the 390 inhabitants under 20 years of age, three fourths live in the poorer	أميون. ومن بين السكان الذين تقل أعمارهم عن العشرين سنة وعددهم 390 نسمة يعيش ثلاثة أرباع في الأحياء الفقيرة،
the very poorest neighbourhoods a mere 52 years. Each marks an improvement over	أما في الأحياء الأشد فقرا في تلك المناطق فإنه لا يزيد على 52 سنة. ويزداد متوسط العمر المتوقع في كل
of weather-related natural disasters in recent years, including unexpected and	وفي السنوات الأخيرة الماضية تكبت هذه القرية بعدد متزايد من الكوارث الطبيعية التي لها صلة بالطقس، كان من بينها
foreign investment. Over the past 25 years, Asia has grown at an annual rate of 7	الاستثمار الأجنبي. وخلال الأعوام الخمسة والعشرين الماضية، بلغ معدل النمو في آسيا 7
fall in life expectancy has occurred in recent years, reflecting reduced public spending on	، طرأ في السنوات الأخيرة الماضية انخفاض حاد في معدلات العمر المتوقع، مما يعكس انخفاض الإنفاق العام على
modernization and education. Moreover, recent years have witnessed a renewed	التحديث والتعليم. وإضافة إلى ذلك بدأت تبرز خلال السنوات الأخيرة الماضية بشكل صارخ العودة إلى
20infected with HIV, half of them under 25 years old. It is a disease that attacks the	أصيب 5.6 ملايين نسمة بذلك الفيروس، وكان نصف هؤلاء دون سن الخامسة والعشرين. وهذا مرض يهاجم
Botswana today has a life expectancy of 41 years, when without AIDS it would have been	ويبلغ متوسط العمر المتوقع للطفل الذي يولد اليوم في بوتسوانا 41 عاما، بينما كان يمكن أن يكون هذا المتوسط 70 عاما لولا الإيدز
of the skilled and educated during their prime years, with what are bound to be tragic	من أصحاب المهارات والمتعلمين وهم في ربيع العمر، مما لا بد أن تترتب عليه آثار مفعجة
. fully, on time and without any conditions. In recent years, the oceans of the world have	كاملة، وفي مواعيدها، ودون أية شروط. لقد حظيت محيطات العالم بمزيد من الاهتمام في الجمعية العامة في الأعوام القريبة الماضية.
the rule of law has increased significantly in recent years and can be considered an	قد زاد عدد الدول التي تتطلب المساعدة في مجال تعزيز سيادة القانون وتقويتها بصورة كبيرة خلال السنوات القريبة الماضية مما يمكن أن يُعد

English text	النص العربي
all clean-up activities. The use of compressed air or regular vacuum cleaners is prohibited;	جميع أنشطة التنظيف. ويمنع استخدام المكنسات الكهربائية العادية أو العاملة بالهواء المضغوط؛
will receive, on an equal basis, free air time on State television to present their	وسوف يحصل جميع المرشحين على حصص زمنية مجانية متساوية من البث على الألفية التلفزيونية الحكومية لعرض
also receive work-therapy care in the open air in the countryside. concerning	كما يتلقون عناية علاجية عن طريق العمل في الهواء الطلق في الأرياف. أما بخصوص
100 per cent fresh air is used and the inside air is not recirculated in the Vienna	وحيث أن الهواء المستنشق في مركز فيينا الدولي هو هواء طلق بنسبة 100 في المائة وأنه لا تجري إعادة تدوير الهواء الداخلي.
is prohibited. There have been no indoor air quality tests performed to date to	وقد مُنع استعمال أي من المواد التي تحتوي على الأسبستوس. ولم تجر حتى الآن أي اختبارات لنوعية الهواء الداخلي
centres and to the freedom and safety of air and surface transportation. We therefore	التجمعات السكانية وحرية وسلامة النقل البري والجوي لذلك
the sea; the training taking place in the open air. The facilities were extremely	البحر، ويجري التدريب في الهواء الطلق. وكانت المرافق بدائية إلى أقصى حد
of rebates on international and domestic air transportation nor fulfilled certain	بالمستردات الخاصة بالنقل الجوي الدولي والمحلي، كما لم يفي ببعض
to his cell. He did not exercise in the open air and did not get any sunlight	لنزائته. ولم يكن يمارس هذه التمرينات في الهواء الطلق ولم يكن أي ضوء شمس يصل إليه.
were deprived of exercise or access to open air for long periods. When	كانوا محرومين من الحركة أو الخروج إلى الهواء الطلق مدداً طويلة. وعند
prohibited holding or taking part in an open-air public meeting for the purpose	حظر عقد أي اجتماع عام في الهواء الطلق أو المشاركة فيه بغرض
air base has become a centre for domestic air and land transport; the former School of	القاعدة الجوية السابقة في البروك، إلى جانب أشياء أخرى، مركزاً للنقل الجوي والبري المحلي؛ وتحولت مدرسة الأمريكتين السابقة
human society needs are water, food, clean air, shelter, energy, raw materials and living spa	والموارد الأساسية التي يحتاج إليها المجتمع البشري هي الماء والهواء النظيف والمأوى والطاقة والمواد الخام ومكان للعيش.
were used to cover the international air travel and living expenses of 24 speakers	في تغطية نفقات السفر الجوي الدولي وبدل المعيشة اليومي لـ 24 متحدثاً
other countries in the South Pacific. Hawaiian Air provides direct air service to Honolulu.	وبلدان أخرى في جنوب المحيط الهادئ. وتوفر شركة طيران هاواي خدمة نقل جوي مباشرة إلى هونولولو.

Group 2/ Session 4

English text	النص العربي
and State levels. With regard to the immediate future, we have identified the	وصعيد الولايات. وفيما يتعلق بالمستقبل العاجل، حددنا
will certainly have an impact on the future programme, given the fact that the	وسيكون لها بكل تأكيد أثر في البرنامج في المستقبل، وذلك بالنظر إلى أن
were expected to return home in the near future. In 1999, there were some positive	ويتوقع عودة الغالبية إلى وطنهم في المستقبل القريب. وفي عام 1999، ظهرت هناك بعض المؤشرات الإيجابية
very difficult to achieve. Looking to the Future Ladies and Gentlemen, Saving lives,	إيجاد الحلول عملية صعبة جداً. نظرة إلى المستقبل أيها السيدات والسادة، إن إنقاذ الأرواح
will place particular emphasis in the immediate future on the following	ستولي أهمية خاصة في المستقبل القريب جداً للمجالات التالية
and our cooperation efforts in the immediate future. As was repeatedly	وكما جرى التأكيد وجهودنا التعاونية في المستقبل القريب جداً. مرارا وتكرارا
to consider what we must achieve in the future and how we can work together to	للنظر فيما يجب أن ننجزه في المستقبل وفي الكيفية الواجب أن نعمل بها معا لتحقيق ذلك.
programme, but it faced an unpredictable future due to a lack of donor funding.	ولكنه واجه مستقبلا يتعذر التنبؤ به بسبب الافتقار إلى التمويل اللازم
challenges they have to confront in the immediate future than the new	التحديات التي يتعين عليهما التصدي لها في المستقبل القريب جداً أكثر من السلطات الديمقراطية الجديدة
behind us and look with confidence towards the future. Institutional capacity-building in a	وراءنا وأن نتطلع إلى المستقبل بثقة. بناء القدرة المؤسسية في
impact probabilities computed for specific future dates. The results are immediately	الممكنة في المستقبل وحساب احتمالات ارتباطه بالأرض لتواريخ معينة في المستقبل. وتُنشر النتائج فوراً
decision to assume that role in the near future. Japan is determined to do its utmost	قرارها المسؤول بالاضطلاع بذلك الدور في المستقبل القريب. إن اليابان عاقدة العزم على بذل قصارى
peace. We believe that our task for the immediate future is to ensure that this	السلام. ونحن نعتقد أن مهمتنا للمستقبل القريب جداً هي ضمان أن يظل هذا
Programme, it outlined the following future directions for its work to prevent and	لبرنامج المفوض السامي، يبين الاتجاهات المستقبلية التالية لعملها في مجال منع
challenges facing us and in charting a better future for all humankind through	للتحديات الأمنية والاقتصادية الأخذة في الظهور التي تواجهنا، وفي رسم مستقبل أفضل للبشرية من خلال

English text	النص العربي
scenario was apparent during an open air meeting organized by a	واتضح السيناريو الجديد المتأرجح خلال اجتماع عقده في الهواء الطلق
or were increasing slightly in Arctic air over the 1987-1997 time period. The	أو التي كانت تترادب بشكل طفيف في هواء المنطقة القطبية الشمالية على مدار فترة 1987-1997.
who carry out economic activities in the open air or in public places, since it	من يزاوون أنشطة اقتصادية في الهواء الطلق أو في أماكن عامة، حيث إن
so thin, once airborne, they can hang in the air for long periods of time before settling.	للرقة المتناهية لهذه الألياف، فإنه يمكنها، إذا انتشرت في الهواء، أن تبقى هناك فترات طويلة قبل أن ترسب.
, in addition to his post as commander of the air force. By 1966 also he had played an	، بالإضافة إلى منصبه كقائد للسلاح الجوي. وبحلول عام 1966
, fire protection, plumbing, heating, air conditioning, electrical systems, roofs,	أيضا كان قد لعب دورا
one in the future. Mr. ANDO noted that open- air meetings were subject to	للإنذار ضد الحريق وتبديل أنظمة التدفئة والتهوية وتكييف الهواء؛ وتحديث الشبكات واللوحات واللوحات الكهربائية؛ وتبديل الإضاءة والأسقف؛
to police approval (paragraph154). Why open- air meetings, in particular? What	في المستقبل. السيد أندو قال إن الاجتماعات التي تعقد في الهواء الطلق تخضع لموافقة
to release harmful, breathable fibres into the air depends on its existing degree of friability	لموافقة الشرطة (الفقرة 154). لماذا الاجتماعات التي تعقد في الهواء الطلق على وجه التحديد؟ وما هو
services were also established in the Army, Air Force and Navy; (c) Essays on	أن تطلق مادة تحتوي على الأسبستوس أليافا يمكن تنفسها في الهواء على مدى قابليته للتفتت
travel into and out of the country, by air , land or sea. This Division carried out	أنشئت مكاتب للشؤون القانونية وشؤون القانون الدولي الإنساني داخل دوائر الأركان في الجيش والقوات الجوية والبحرية؛ يتم إعداد دراسات
in 1985, problems had been noted in terms of air pollution and storage of dangerous	السفر إلى داخل البلد وخارجه، جوا وبرا وبحرا. ونفذت هذه الشعبة
more expensive because of the need to pay air travel and daily subsistence allowances.	بأنه لوحظت مشاكل من حيث تلوث الهواء و تخزين مواد ضارة
increased its monthly deliveries, mostly by air , from 12,000 to 16,000 tons.	فهو مكلف بقدر أكبر بكثير، نتيجة لضرورة دفع مصروفات السفر بالجو وبدلات الإقامة اليومية.
,marches, conferences, publicities, open- air concerts or children' courses;	رفع البرنامج الكميات التي يقدمها شهريا، عن طريق الجو أساسا، من 12 000 إلى 16 000 طن.
	مسيرات، أو عقد مؤتمرات، أو الدعاية، أو إقامة حفلات موسيقية في الهواء الطلق، أو وضع مقررات دراسية للأطفال

Appendix H: Monolingual corpus data (samples)

Group 3/Session 1

<p>hopes of people everywhere. For its first 45 years , the United Nations lived in the grip of globalization and governance. In the early years of the United Nations, the General that significant progress has been made in recent years . Advisory Service on was proud to promote, took too many years and cost too many lives. illiterate. Of the 390 inhabitants under 20 years of age, three fourths live in the poorer years . Each marks an improvement over of weather-related natural disasters in recent years , including unexpected and foreign investment. Over the past 25 years , Asia has grown at an annual rate of 7 fall in life expectancy has occurred in recent years , reflecting reduced public spending on modernization and education. Moreover, recent years have witnessed a renewed 20infected with HIV, half of them under 25 years old. It is a disease that attacks the Botswana today has a life expectancy of 41 years , when without AIDS it would have been of the skilled and educated during their prime years , with what are bound to be tragic fully, on time and without any conditions. In recent years , the oceans of the world have the rule of law has increased significantly in recent years and can be considered an</p>

transfer of these weapons, to clear mined areas in order to free land from its deadly and made specific progress in the following areas : (a) towards efficient and effective force proposals to improve these four key areas will follow in 2000 by the international community in three key areas : programmes to strengthen on a priority basis. Some of the key areas requiring immediate for change” across our various areas of responsibility. Many of the networks district is nearly 78 years, in the poorer areas 64 years - and in the very poorest . I have focused on strategic priority areas where, in my view, we can and must make . I wish to highlight a number of specific areas for particular attention by the Summit. in school are girls. Female enrolment in rural areas remains shockingly low. Short-changing member States should focus on three key areas: coordinated approach coordination in the following key areas : capacity-building, the community to meet and socialize, no safe areas for children to play. Slum dwellers live Internet and fax service to poor and rural areas . With help from civil society and the fall of Srebrenica and the other safe areas in Bosnia. The many reasons for those

challenges only if all of us feel a renewed sense of mission about our common concrete results and creating a greater sense of urgency about the need to address . We human beings must act according to common sense . For more than 40 years . In the meantime, I hope that the growing sense of urgency among a large number of greater involvement normally leads to greater sense of commitment. There is also a concern and first gave us the expression to the true sense of independence, sovereignty and civilizations as means to move forward. Common sense dictates that the of a younger generation with a clear sense of optimism, solidarity and values, of . Government of Rwanda maintains a strong sense of self-reliance when it comes to what . basic learning needs in the best and fullest sense of the term, an education that includes . It is not limited to nationality in a formal sense , that is, nationality acquired at birth or people affected by war, and restoring their sense of hope. This issue must become a should not be seen as such, it is simply common sense . The moderator not ideal, it was a product of common sense and good will. The against the Sudan. It is based simply on common sense and is aimed at providing

Group 3/ Session 4

<p>and State levels. With regard to the immediate will certainly have an impact on the were expected to return home in the near very difficult to achieve. Looking to the will place particular emphasis in the immediate and our cooperation efforts in the immediate to consider what we must achieve in the programme, but it faced an unpredictable challenges they have to confront in the immediate behind us and look with confidence towards the impact probabilities computed for specific decision to assume that role in the near peace. We believe that our task for the immediate Programme, it outlined the following challenges facing us and in charting a better</p>	<p>future , we have identified the future programme, given the fact that the future . In 1999, there were some positive future Ladies and Gentlemen, Saving lives, future on the following future . As was repeatedly future and how we can work together to future due to a lack of donor funding. future than the new future . Institutional capacity-building in a future dates. The results are immediately future . Japan is determined to do its utmost future is to ensure that this future directions for its work to prevent and future for all humankind through</p>
---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------	--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------

<p>scenario was apparent during an open or were increasing slightly in Arctic who carry out economic activities in the open so thin, once airborne, they can hang in the , in addition to his post as commander of the , fire protection, plumbing, heating, one in the future. Mr. ANDO noted that open to police approval (paragraph154). Why open to release harmful, breathable fibres into the services were also established in the Army, travel into and out of the country, by in 1985, problems had been noted in terms of more expensive because of the need to pay increased its monthly deliveries, mostly by ,marches, conferences, publicities, open-</p>	<p>air meeting organized by a air over the 1987-1997 time period. The air or in public places, since it air for long periods of time before settling. air force. By 1966 also he had played an air conditioning, electrical systems, roofs, air meetings were subject to air meetings, in particular? What air depends on its existing degree of friability Air Force and Navy; (c) Essays on air , land or sea. This Division carried out air pollution and storage of dangerous air travel and daily subsistence allowances. air , from 12,000 to 16,000 tons. air concerts or children' courses;</p>
-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------	---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------

<p>an apparently intractable civil war only a few a greater focus of attention in Africa in recent . services for us. During the past hundred .is lost to urban sprawl. Yet over the next 30 . As we look at developments over recent today as well as children born in future developing countries some begin as early as six has declined by 51 per cent in the past 36 regions and little change in the past 15-20 by civil society networks during the last few slowed the impressive progress of recent and social progress over the past few hosted large numbers of refugees for many over 400 factories have been closed in recent , we are glad to be able to state that in recent</p>	<p>years ago, has taken great strides thanks to years . This is reflected, for example, years , the natural environment has borne the years the demand for food in the developing years , we are encouraged that much years . For those children, the United Nations must years of age. Tobacco companies continue to years), about 10 million-15 million children years . Bangladesh, India and Pakistan have years , which had contributed to the years , particularly in the areas of years as well as their vision of Africa's future, years , longstanding generosity years and that the textile sector alone years our country has made significant</p>
-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------	----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------

Appendix I: Pre, post and delayed post-test (active recall)

Please translate the bolded Arabic word combinations into English combinations according to its context. Make sure that you provide the most appropriate translation(s).

١. ومن دواعي المنطق البديهي / السليم، أنه طالما كان هناك حرب، فإن هناك قتال.
.....
٢. ساد اتفاق واسع النطاق بأن البلدان النامية المستوردة للنفط في وضع شديد الصعوبة.
.....
٣. هناك أرقام لا يمكن رؤيتها بالعين المجردة إذ تتطلب معدات خاصة لكشفها.
.....
٤. صدر في أوائل صيف عام ١٩٩٩ في فرنسا كتاب بعنوان 'الشرق الأوسط بين الحرب والسلام'.
.....
٥. أصبح وقوع الكوارث الطبيعية في السنوات الأخيرة الماضية أكثر تكرارا.
.....
٦. حددت اللجنة عدة مجالات رئيسية للمناقشات.
.....
٧. ينبغي أن تكون هذه المسألة من أهم الأهداف في المستقبل القريب العاجل.
.....
٨. ويجري التدريب في الهواء الطلق.
.....
٩. إن الغالبية العظمى من أحجار الماس المنتجة في العالم تأتي من مصادر مشروعة.
.....
١٠. ولهذا من الضروري إيلاء اهتمام كبير بتفاصيل العمل الفعلية.
.....
١١. تم بذل الجهود لمعالجة مشكلة العنف الأسري المرتكب ضد المرأة.
.....
١٢. إن الوثائق عندما تنشر إلكترونيا تتاح معها أيضا نسخ ورقية.
.....
١٣. واسترعت الانتباه إلى الحاجة إلى تخطيط دقيق للعمليات.
.....
١٤. وتغطي الحكومة تكاليف رحلة ذهاب وإياب للطالب من وطنه إلى الجامعة.
.....
١٥. تقوم تقنية المعلومات بخدمة أعداد هائلة من العاملين والعملات في القرى النائية.
.....
١٦. نظرا لاعتلال صحته، تقرر إلغاء هذين اللقاءين.
.....
١٧. أحرزت كندا تقدما مطردا بالرغم من بطئه في تحقيق المساواة بين الجنسين.
.....

١٨. ومثلت الوحدة بداية جديدة في حياة الألمان.

١٩. ينبغي أن يكون هناك تركيز شديد على التنمية الاجتماعية والتعليم.

٢٠. ونشكر كل من أسهموا بشكل أو آخر في هذا النجاح الكبير والباهر.

٢١. ويستفيد من اتفاقية أبرمت مع وزارة الثقافة، عديد من المنظمات في مجالات الفنون الجميلة والسينما والتصوير الفوتوغرافي.

٢٢. تأثر معظم الناس الذين يعيشون في المدن بتلوث الهواء الناجم عن حركة المرور الكثيفة.

٢٣. فالإيرلنديون لهم تقاليد عريقة في العمل في خدمة الآخرين.

٢٤. فبإمكان المرء في المدن الريفية أن يجد زجاجة دواء إلى جانب القهوة سريعة التحضير والحليب المجفف.

٢٥. وهناك نقص في عدد غرف التدريس والمدارس، أما المدارس القائمة فهي في حالة سيئة.

٢٦. الكوارث الطبيعية تتسبب في خسائر جسيمة وفادحة في الأرواح والممتلكات.

٢٧. ومن بين القضايا المحددة التي سوف يتناولها الاجتماع حرية التنقل وسلامة العودة.

٢٨. على الرغم من هذه الحقائق الثابتة إلا أن أفريقيا تؤكد على ثقته في المجتمع الدولي.

٢٩. جميع الشركات قد نفذت المشاريع بمصداقية وحسن نية.

٣٠. وتجري لجميع المواطنين الكوبيين، سواء في ذلك الفئات المعرضة لمخاطر قليلة أو متزايدة، اختبارات الإصابة بالإيدز إذا رغبوا في ذلك.

Appendix J: Pre, post and delayed post-test (passive recall)

Please translate the bolded English word combinations into Arabic combinations according to its context. Make sure that you provide the most appropriate translation(s).

1. A book published in France in the **early summer** of 1999 was titled War and Peace in the Middle East.
.....
2. In **recent years**, natural disasters have become more regular.
.....
3. It is only **common sense** that when there is war, there will be fighting.
.....
4. There was **broad agreement** that oil-importing developing countries were in a difficult situation.
.....
5. The numbers are not visible to the **naked eye**; and need special equipment.
.....
6. They identified many **key areas** for discussions.
.....
7. That is why it is necessary to pay **careful attention** to the details of the actual work.
.....
8. Efforts are made to deal with the problem of **domestic violence** against women.
.....
9. When documents were published electronically, **hard copy** forms were also made available.
.....
10. This matter should be one of the most important goals in the immediate future.
.....
11. The training is taking place in the **open air**.
.....
12. The **vast majority** of diamonds produced in the world are from legal sources.
.....
13. Information technology serve those **vast numbers** of working men and women in distant villages.
.....
14. Because of his **ill health**, those meetings had to be cancelled.
.....
15. She drew attention to the need for **careful planning** of the work processes.
.....
16. The government provides for the cost of **round trips** of the successful student from his country to the university.

-
17. Most of the people lived in cities are affected by air pollution caused by **heavy traffic**.
.....
18. In rural towns, one can find small bottles of medicine among **instant coffee** and milk powder.
.....
19. There is a lack of classrooms and schools, and the existing ones are often in **poor condition**.
.....
20. Natural disasters cause **heavy losses** of human life and property.
.....
21. **Safe return** of women and children are among the issues to be addressed in the meeting.
.....
22. Canada has made slow but **steady progress** in understanding equality between men and women.
.....
23. Unification marked a **fresh start** in the life of the Germans.
.....
24. There should be **heavy emphasis** on social development and education programmes.
.....
25. We thank all those who led in one way or another to that **huge success**.
.....
26. A number of organizations in the fields of **fine arts**, cinema and photography, have signed agreements with the Ministry of Culture.
.....
27. Irish people have a **long tradition** of working abroad in the service of others.
.....
28. All Cuban citizens, both high-risk and **low-risk** groups, were given AIDS tests if they wished.
.....
29. Despite these **hard facts**, Africa confirmed its confidence in the international community.
.....
30. All companies had carried out the projects in **good faith**.
.....

Appendix K: Research information sheet

Information for the Research Participant

Dear participant,

Information about the researcher

I'm a PhD student at Newcastle University. I am majoring in Applied Linguistics and my research interest is on foreign language vocabulary acquisition.

Participants: undergraduate students majoring in English

Research duration: 12 weeks (approximately)

Procedures:

- Some preliminary data will be collected from you prior to the start of the classes.
- You will be given worksheets to work on during your usual class time.
- More essential data will be collected from you at the end of sessions.

Ethical considerations

It is VERY important to know that the collected information **will stay confidential** as the consent form clearly explains. Your participation will not affect any aspect of your study. The data collected from your participation will not be seen by any teacher of your institution, and will only be used for the purpose of this research project. They will be processed and analysed by the researcher. Names and personal information as well as your test results **will be entirely anonymous**.

For any further questions and concerns regarding my research, please do not hesitate to contact me!

Rezan Alharbi

Email: r.m.s.alharbi@ncl.ac.uk

Address: 2.34 KGVI

King George VI Building

School of ECLS, University of Newcastle

Queen Victoria Road

Newcastle upon Tyne

NE1 7RU

Appendix L: Consent form

NEWCASTLE UNIVERSITY

School of Education, Communication and Language Sciences

Participant Consent Form

Declaration of Consent

It is a university requirement that all respondents give their formal consent to take part in any research. For this reason could you please sign and date the declaration below.

Consent to the use of (worksheets and test papers)

I confirm that I have read the written statement provided by the researcher (entitled 'Information for the research participant') and had the opportunity to ask questions. I consent to participate in this research project. I understand that all the data will be kept confidential and I will be anonymous in the research report. I also know that the data gathered from this project will be used for the purposes stated in the Participant Information Form.

I understand that my participation is voluntary and that I am free to withdraw from the project at any time, without needing to give a reason.

I understand that whether I participate or not will have no effect on my grades/assessment.

Name of participant Date Signature Email

Researcher Date Signature

One copy to the participant and one to the researcher