Teaching English as a Foreign Language through a Data-driven Learning perspective –

using an annotated pedagogic corpus of English textbooks in a Greek high school class

by

Vasiliki Papaioannou

A thesis submitted for the degree of

Doctor of Philosophy in Applied Linguistics

at

Aristotle University of Thessaloniki

Department of Theoretical and Applied Linguistics

School of English

May 2018
I dedicate this thesis to my husband, Yannis,

who has always helped me and believed that I could do it.
ABSTRACT

This thesis examines the integration of a Data-driven learning (DDL) methodology in the learning context of a Greek Senior high school class. Its aim is to assess the usefulness of such an approach for the learning of English modal verbs. DDL, an inductive teaching approach which exploits data derived from corpora, has gained a fair degree of attention over the last 30 years, but still there is little research involving actual application in Secondary education EFL settings.

The present study addresses this gap by investigating the application of a blended DDL approach to the teaching of English modal verbs, specifically tailored to the needs of Senior high school Greek students. It is called blended because it first of all makes use of material from a variety of sources; it involves the use of a) a pedagogical corpus of course book material (ECCo) which we compiled specifically for the purposes of the present research endeavor, b) material from the Web, and c) a general online corpus, the Corpus of Contemporary American English (COCA). The proposed methodology is called blended for a further reason; it employs both printed concordances handouts and an online hand-on approach.

The proposed DDL methodology was applied to the experimental group, whereas the traditional course book-based approach was selected for the control group. The learning outcomes of the two approaches were evaluated with the use of pre, post and delayed post-tests. The data obtained were submitted to statistical analyses (SPSS 20) which aimed to measure learner’ performance. The overall results of this study indicated that (a) suitable DDL material which conform to the curriculum for the teaching of English in schools could be compiled by EFL teachers, (b) a DDL methodology can positively affect the learning of English grammar, (c) the improvement in the immediate and delayed post-tests is statistically significant, and (d) a DDL methodology suits best ‘low level’ students, which is those who had received the lowest grades in the pre-experimental test. The study adds to the EFL literature focusing on corpus-driven learning in upper secondary education.
ACKNOWLEDGEMENTS

I would like to thank, first and foremost, my primary supervisor Professor Marina Mattheoudakis for her continuous support, guidance and encouragement over the last few years. Thank you for believing in me and for being always there to offer your advice, help and inspiration throughout this thesis. After every meeting we had, I felt a little wiser.

I would also like to thank my committee members Associate Professor Eleni Agathopoulou for her insightful comments which improved the quality of my work, and Professor Georgios Mikros for his expertise on statistical matters, which assisted me in clarifying key issues in the statistical analysis of the test results.

Special thanks go to the following people who have contributed with comments and suggestions at several stages of this research: Professor T. Tsangalidis for the discussions we have had about modals, Assistant Professor A. Tantos for his advice on corpus compilation issues, and Dr. E. M. Joycey for his comments on the design of the tests.

Many thanks go to my students, who willingly participated in the pilot study, the training DDL sessions and the main study’s DDL lessons. Their comments and their questions helped me design better DDL lessons; their enthusiasm is what makes me want to be a better teacher.

This work would not have been completed without the support of my family and especially my mother, an amazing woman whose life has always inspired me to try harder and aim higher. Many thanks go to my sister for her encouragement and positive belief in my success, and my mother-in-law for her support and the endless babysitting hours during the writing stage of this thesis. And finally, to Yannis, who has been by my side throughout this PhD and has shown me how amazing Math really is.
TABLE OF CONTENTS

ABSTRACT..............................................................................................................................................i

ACKNOWLEDGEMENTS..........................................................................................................................ii

LIST OF FIGURES ....................................................................................................................................x

LIST OF TABLES ....................................................................................................................................xiii

LIST OF APPENDICES ............................................................................................................................xv

LIST OF ABBREVIATIONS .......................................................................................................................xvii

CHAPTER 1: INTRODUCTION ..................................................................................................................1

CHAPTER 2: CORPUS LINGUISTICS - THEORETICAL

BACKGROUND ...........................................................................................................................................6

2.1 Introduction ........................................................................................................................................6

2.2 A brief history of Corpus Linguistics ..............................................................................................6

2.2.1 Definition of ‘corpus’ and ‘corpus linguistics’ .............................................................................6

2.2.2 Uses of computerized corpora: past and present ......................................................................8

2.3 Corpus compilation – the case of ‘do-it-yourself’ (DIY) corpora .................................................10

2.3.1 Corpus compilation issues ..........................................................................................................10

2.4 Concordances and frequency counts .............................................................................................12

2.5 Collocation and colligation .............................................................................................................13

2.6 Indirect and direct pedagogical applications of corpora ..............................................................14

2.7 The corpus-based grammatical analysis .........................................................................................16

2.8 Conclusion .......................................................................................................................................17
CHAPTER 3: DATA-DRIVEN LEARNING ................................................................. 18

3.1 Introduction .................................................................................................... 18

3.2 Approaches to Language Learning with relation to Corpus Linguistics ....... 18
   3.2.1 The inductive approach and the noticing hypothesis ......................... 18
   3.2.2 Autonomous learning ........................................................................... 19
   3.2.3 Task-based learning ............................................................................ 20
   3.2.4 Communicative language teaching ..................................................... 20
   3.2.5 The lexical approach ........................................................................... 21
   3.2.6 Pattern (or Lexical) grammar ............................................................... 22
   3.2.7 Blended learning ............................................................................... 22

3.3 The Data-driven Learning ............................................................................ 23

3.4 Corpora for DDL .......................................................................................... 26
   3.4.1 The Web as a corpus ........................................................................... 27
   3.4.2 General corpora ................................................................................... 28
   3.4.3 Specialized corpora ............................................................................. 29
   3.4.4 Learner corpora ................................................................................ 30
   3.4.5 Pedagogic corpora ............................................................................. 32

3.5 Typology of DDL activities .......................................................................... 33

3.6 DDL potentials and limitations ................................................................... 35
   3.6.1 DDL potentials ................................................................................ 35
   3.6.2 Limitations of DDL .......................................................................... 36

3.7 Review of studies on DDL .......................................................................... 38
   3.7.1 Cases involving grammar teaching through DDL ............................ 39
   3.7.2 DDL case-studies in Secondary education ...................................... 40
5.3.3 Foreign language learning in Greece .............................................. 63
  5.3.3.1 The new curriculum for compulsory education ......................... 65

5.3.4 Foreign language learning in Upper Secondary Education in Greece.
............................................................................................................. 67
  5.3.4.1 The objectives set by the Syllabus for Upper-secondary education ......................................................... 68

5.4 The pilot study ..................................................................................... 69
  5.4.1 The pilot study participants ............................................................... 70
  5.4.2 The pilot study DDL training sessions ............................................. 70
  5.4.3 The contribution of the pilot study .................................................... 74

5.5 The main study .................................................................................... 74
  5.5.1 The main study participants ............................................................... 74
    5.5.1.1 The experimental and the control group ................................. 75

5.6 The instruments and materials ............................................................. 76
  5.6.1 The needs analysis questionnaire ...................................................... 76
  5.6.2 The ‘Attitudes towards Computer Assisted Language Learning’
    questionnaire ...................................................................................... 78
  5.6.3 The Quick placement test ................................................................. 79
  5.6.4 The pre, post and delayed - post experimental tests ......................... 80
  5.6.5 The ECCo Corpus .......................................................................... 82
    5.6.5.1 Corpus material choice ............................................................. 82
    5.6.5.2 The ECCo corpus compilation ................................................... 85
    5.6.5.3 Metadata information ............................................................... 88
    5.6.5.4 Corpus annotation .................................................................... 89
    5.6.5.5 ECCo concordancing ............................................................... 90
  5.6.6 The COCA ...................................................................................... 91
7.5.1 The test scores ................................................................. 132
7.5.2 Analysis of the score increase between pre and post-tests .......... 138
7.6 Analysis of the statistical significance ...................................... 140
  7.6.1 Within-group statistical analysis – paired samples tests .......... 141
  7.6.2 Between-group statistical analysis - ANOVA ....................... 142
7.7 Analysis of the test results for each Modal Verb ......................... 144
7.8 Conclusion ........................................................................... 149

CHAPTER 8: DISCUSSION OF THE RESULTS ..................................... 151
  8.1 Introduction .......................................................................... 151
  8.2 Discussion of the results ...................................................... 151
    8.2.1 Research question 1 ....................................................... 151
    8.2.2 Research question 2 ....................................................... 154
    8.2.3 Research question 3 ....................................................... 156
  8.3 Major findings of the study .................................................... 165
  8.4 Conclusion ........................................................................... 167

CHAPTER 9: CONCLUSION ......................................................... 168
  9.1 Introduction .......................................................................... 168
  9.2 Implications for pedagogy ..................................................... 168
    9.2.1 The case for richness and diversity of materials ................. 168
    9.2.2 The middle-way: the case for blended Data-driven language learning... ................................................................. 169
    9.2.3 The case for a combined DDL and Traditional learning methodology ................................................................. 170
    9.2.4 What does DDL have to offer to EFL teaching in Greek state schools? .................................................................... 170
9.3 Directions for future research .............................................. 171
9.4 Conclusion ................................................................................. 172
BIBLIOGRAPHY................................................................................. 174
Chapter 2

Figure 2.1 Types of pedagogical corpus applications (Römer 2010:19) ............. 14
Figure 2.1 The uses of corpora with relevance to language pedagogy (adapted from Johansson 2009: 40) .......................................................... 15

Chapter 3

Figure 3.1 The relationship between form and function in language, upon which the cycle of DDL is based, according to Johns (1994: 294) ............. 24
Figure 3.2 The blended DDL learning approach ............................................. 25
Figure 3.3 Concordance output used for investigating grammatical structures (Johns 1991: 7) ................................................................. 34
Figure 3.4 Collocate search for effort taken from COCA ............................ 34
Figure 3.5 Concordance output for effort in KWIC format taken from COCA ... 35

Chapter 4

Figure 4.1 Sub categorization of modality ................................................. 46
Figure 4.2 The meanings of modals, according to Quirk et al. (1995: 221) ...... 51

Chapter 5

Figure 5.1 The Greek Education System according to the EU’s Eurydice Report 2016/2017 ................................................................. 61
Figure 5.2 The Common European Framework of Reference descriptors ........ 66
Figure 5.3 Pilot Session 1 outline .............................................................. 71
Figure 5.4 Pilot Session 2 outline .............................................................. 72
Figure 5.5 Pilot Session 3 outline .............................................................. 73
Figure 5.6 Screenshot of ECCo extract saved in .txt format ......................... 86
Figure 5.7 Screenshot of ECCo corpus statistics ....................................... 87
Figure 5.8 Screenshot of ECCo header file ................................................. 89
Figure 5.9 Screenshot of ECCo annotation file ....................................... 90
Chapter 6

Figure 6.1  KWIC query of ‘innuendo’ in COCA ........................................ 95
Figure 6.2  The 1st training session handout ........................................ 96
Figure 6.3  Sample output of ‘innuendo’ in COCA .................................. 97
Figure 6.4  The 2nd training session handout ....................................... 98
Figure 6.5  Sample output of ‘swirl of’ in COCA .................................. 99
Figure 6.6  The 3rd training session handout ....................................... 100
Figure 6.7  The 4th training session handout ....................................... 101
Figure 6.8  Sample of the 5th training session handout ......................... 103
Figure 6.9  Comparing the words ‘deny/ refuse’ in the COCA interface .... 104
Figure 6.10 The training sessions’ evaluation questionnaire results .......... 106
Figure 6.11 Configurations of access to the data resources used in the present study’s DDL lessons (adapted by Lew 2009: 298) ...................... 108
Figure 6.12 The 1st DDL lesson handout .......................................... 109
Figure 6.13 The 2nd DDL lesson handout .......................................... 111
Figure 6.14 The 3rd DDL lesson handout .......................................... 114
Figure 6.15 The 4th DDL lesson handout .......................................... 118
Figure 6.16 The 5th DDL lesson handout .......................................... 121

Chapter 7

Figure 7.1  Distribution of original and normalized scores for the experimental group 1 pre-test results .............................................................. 131
Figure 7.2  Experimental group 1 (class B1) scores ................................. 133
Figure 7.3  Experimental group 2 (class B2) scores ................................. 134
Figure 7.4  Experimental group 1 (class B1) scores ................................. 135
Figure 7.5  Pre-test knowledge level for all groups ................................. 136
Figure 7.6  Immediate post-test knowledge level for all groups ....................... 137
Figure 7.7  Delayed post-test knowledge level for all groups .......................... 138
Figure 7.8  Sum of points difference between pre-test and post-test scores of students listed according to their pre-test scores ................................. 139
Figure 7.9  Sum of points difference between pre-test and post-test scores of students listed according to their post-test scores ................................. 139
Figure 7.10 Experimental group 1 (class B1) success rate in every modal verb . 145
Figure 7.11 Experimental group 1 (class B1) successful hits for every modal verb ........................................................................................................... 145
Figure 7.12 Experimental group 1 (class B2) success rate in every modal verb . 146
Figure 7.13 Experimental group 1 (class B2) successful hits for every modal verb ........................................................................................................... 147
Figure 7.14 Experimental group 1 (class B3) success rate in every modal verb . 148
Figure 7.15 Experimental group 1 (class B3) successful hits for every modal verb ........................................................................................................... 148
LIST OF TABLES

Chapter 4

Table 4.1  Modal auxiliary frequency lists in ECCo and COCA ..........................48
Table 4.2  Forms and meanings of the six modal verbs discussed in the present study ..........................................................54

Chapter 5

Table 5.1  The main study participants’ CEFR level ........................................80
Table 5.2  Frequency in the Corpus of Contemporary American English (in tokens) ...............................................................81
Table 5.3  The frequency of modal verbs in ECCo ...........................................88

Chapter 6

Table 6.1  The training sessions’ evaluation questionnaire results ......................105
Table 6.2  The grammar focus of the 5 DDL Units ........................................107

Chapter 7

Table 7.1  The data collection timeline ..........................................................126
Table 7.2  The senses of will, would, could, should, may, might in the pre, post and delayed post-experimental tests ..........................................................127
Table 7.3  Mean scores of Pre, Immediate Post and Delayed Post-tests in all three groups using simple counting .................................................128
Table 7.4  B1 class sample knowledge level score (after the normalization) .......131
Table 7.5  Descriptive statistics for experimental group 1 (B1) .........................132
Table 7.6  Descriptive statistics for experimental group 2 (B2) .........................133
Table 7.7  Descriptive statistics for Control group (B3) ...................................134
Table 7.8  Relative standard deviation (CV) values for the pre and immediate post-tests ........................................................................135
Table 7.9  Paired Samples Test for the experimental group 1 .......................... 141
Table 7.10 Paired Samples Test for the experimental group 2 .......................... 142
Table 7.11 Paired Samples Test for the control group ........................................ 142
Table 7.12 ANOVA for the pre-test ........................................................................ 143
Table 7.13 ANOVA for the immediate post-test .................................................. 143
Table 7.14 ANOVA for the delayed post-test ....................................................... 144
LIST OF APPENDICES

APPENDIX A: The Pedagogic corpus ECCo .......................... 201
   A.1 The header ........................................................................ 201
   A.2 ECCo sample .................................................................... 202
   A.3 The UCREL CLAWS7 tagset ............................................. 203
   A.4 Sample CLAWS7 tagger output ....................................... 209
   A.5 WordSmith Tools concordancer output ......................... 210
      A.5.1 Full sentences (keyword: past perfect verbs) ........... 210
      A.5.2 Blanking the keyword (past perfect tense verbs) ..... 210

APPENDIX B: The Questionnaires .............................................. 211
   B.1 The Questionnaire investigating the student’s attitudes towards
       Computer Assisted Language Learning .......................... 211
   B.2 The Questionnaire investigating the student’s attitudes towards
       Grammar Learning ..................................................... 215

APPENDIX C: The pilot study article ........................................ 216

APPENDIX D: The main study ................................................... 218
   D.1 The pre, post and delayed post-test ................................. 218
   D.2 Main study lesson plan 1: “Words and ideas can change the
       world!” ........................................................................ 226
   D.3 Main study lesson plan 2: “Shall we?” ............................. 227
   D.4 Main study lesson plan 3: “Will you join us?” .................. 228
   D.5 Main study lesson plan 4: “You can do it!” ..................... 233
   D.6 Main study lesson plan 5: “Shoot for the moon: you might get
       there” ........................................................................ 238
D.7 The traditional teaching methodology learning material ...... 242

APPENDIX E: Pre-, immediate post- and delayed post-test statistics ........ 251

E.1 Pre, immediate post and delayed post-test scores using simple counting................................................................. 251

E.2 Pre, immediate post and delayed post-test scores using normalized scores........................................................... 254

E.3 Distribution of original and normalized scores for all groups' pre, immediate –post and delayed –post-tests....................... 258
LIST OF ABBREVIATIONS

ANOVA - one-way analysis of variance
BNC – British National Corpus
CEC – Cambridge English Corpus
CEFR – Common European Framework of Reference
CING – Chemnitz Internet Grammar
CPE – Corpus of Professional English
COCA – Corpus of Contemporary American English
DDL – Data-Driven Learning
ECCo – English Coursebooks Corpus
EFL – English as a Foreign Language
EKPA – National Kapodistrian University of Athens
ESLC - European Survey on Language Competences
ESP – English for Specific Purposes
GSL – General Service List
IEP – Institute of Educational Policy
IFLC – Integrated Foreign Languages Curriculum
LDD – Learning Driven Data
MAux – Modal Auxiliary Verbs
NGSL – New General Service List
WaC – Web as Corpu
CHAPTER 1

INTRODUCTION

The advent of modern technology has provided the foreign language teachers with a learning methodology which makes use of the abundant data found in modern electronic corpora and has the dual advantage of making use of natural-occurring language as well as promoting discovery learning. This methodology, called Data-driven learning (DDL), has been defined as “one application of corpora that involves language learning based upon direct contact with language corpora” (Boulton 2008: 581) through “guided tasks or materials based on corpus evidence, such as concordance lines on handouts” (O’Keeffe et al. 2007: 21), in order to “see patterning in the target language and to form generalizations about language form and use” (Johns 1991a: 2).

One central aspect of DDL is that it changes the way we view language data; in DDL, language data is not only a tool, but a direct source of knowledge. In DDL students are presented with data in the form of concordances and are guided by questions that direct their attention to certain aspects of that data in order to help them understand the usage of a specific term and at the same time increase their "feel" for the common patterns of the language, i.e., the way things are commonly expressed in the language. This is advantageous for the language learner “since it invites him/ her to obtain, organize and study real-language data according to individual choice, giving him thus the realistic expectation of breaking new ground as a ‘researcher’, doing something which is a unique and individual contribution” (Leech 1997: 10).

Teaching English grammatical structures in secondary education is especially challenging since students seem to suffer from lack of motivation and appropriate materials. As noted above, DDL can reinforce motivation and active participation in the learning process through the construction of teaching materials that keep the learner involved, and moreover seem to keep up to date with the latest trends of communicative learning and authenticity in learning. Therefore, as Römer (in Aijmer ed. 2009: 90) suggests:

“It will be an important future task for the corpus researcher to create more DDL material that address particular language items, especially items which cause
constant problems for learners and that could be used directly in the EFL classroom”

The descriptive and pedagogic value of corpora has been recognized by researchers worldwide; nevertheless, the actual use of corpora in language learning settings has for a long time remained, as Bernardini (2004: 15) points out “somewhat behind such momentous breakthroughs”. As far as data-driven teaching methodology is concerned, research has been limited in terms of actual study cases. Johansson (2009: 41) echoes this when saying “what I miss are systematic studies testing the benefits of the (DDL) approach”, a belief also shared by other scholars such as Lessard-Clouston and Chang (2014) and Anthony, Chen and Flowerdew (2017). In other words, whereas the volume of research on theoretical approaches to the usefulness of corpora for EFL learning has been large, empirical studies on real-classroom use of electronic corpora are considerably fewer.

This is especially true in the case of empirical study cases in the setting of secondary education. Back in 2008, A. Boulton reported that only four out of the fifty studies considering DDL he had analyzed involved language teaching outside the university. In the 9 years which intervened between Boulton’s report and the present study, another handful of empirical studies have been added to the somewhat disappointing total of research investigating the application of DDL in secondary education settings. A report by Tribble (2012) stresses the fact that almost 80% of the respondents to a questionnaire addressed to practitioners who employed corpus techniques in classroom worked in tertiary education. The sparsity of secondary education-related DDL studies is also underlined in the latest meta-analysis of DDL empirical cases (Boulton and Cobb 2015).

Even fewer are the study cases focusing on direct uses of DDL with low and middle proficiency level learners (Poole 2016) as well as those focusing on the teaching of grammatical structures, a not so attractive part of EFL learning; the bulk of research seems to have focused on the teaching of vocabulary. Last, but certainly not least, research on the actual benefits of DDL methodology for teaching English to Greek high school students has been to our knowledge, to-date non-existent.
The rationale upon which the present thesis is based is to shed light to some of the less researched areas of data-driven language learning; more specifically, we chose to investigate the teaching of grammar to upper secondary state school learners, using a blended DDL approach which combines a variety of materials and tools, and to evaluate the results of our experimental study with the use of pre, post and delayed post-tests, with the aim to

a) investigate how a DDL methodology can be applied in a Greek senior high school setting, and

b) examine if the learners of English in Greek high schools can benefit from such an approach to language teaching

The present thesis draws on the legacy of the corpus-based approach to linguistics and language education in order to propose a teaching methodology which employs a specifically tailored DDL approach, with the following characteristics:

a. It is a blended approach (Ebeling 2009: 67) since it combines technology-based materials (the pedagogic corpus ECCo and the online giga corpus COCA, as well as the WordSmith Tools concordancer) and face-to-face sessions (in-class teaching).

b. It involves pure hands-on activities and teacher-tailored learning material in the form of handouts containing concordance output

c. It employs both house-made pedagogic (ECCo) and general online (COCA) corpora

d. It is essentially a discovery learning methodology; however, it is also teacher-guided, the teacher’s role being to initiate the learner into discovery learning

In our empirical study we created a DDL methodology which could be applied in the teaching of English as a foreign language in a senior high school class, with the design of DDL lessons which involves reading, interpreting and generalizing over concordances. The present thesis is structured as follows:

Chapter 2 overviews the history of corpus linguistics and discusses compilation and analysis issues, with a focus on the compilation of ‘do-it-yourself corpora’. It also
presents the notions of collocation and colligation, while focusing on how grammar is presented and analyzed in corpora.

Chapter 3 presents DDL in detail; we focus on its characteristics, its potentials and limitations, the different types of corpora involved in DDL and the types of DDL activities. It finishes with a review of the empirical studies on DDL with special emphasis on studies which focus on the teaching of grammar, and those which have been conducted in Secondary education.

Chapter 4 discuss the rationale behind our choice to teach the English modal verbs using a DDL methodology. We aim to present the challenges modal verbs pose for foreign language learners and discuss previous studies which have focused on this area of EFL learning and teaching.

Chapter 5 sets the research methodology that was followed in our empirical study, outlining the research hypotheses and the research questions that our study attempts to answer. It also gives a detailed description of the pilot study of our research; it presents the target group, the structure and the findings of the pilot study which preceded the main study of the present research. It proceeds with discussing the main study of our research. It begins by providing background information on the educational system of Greece and in particular the upper-secondary education that is of more importance for the purposes of our research, since our target group is upper-secondary education students. Furthermore, it gives insight into the strengths and weaknesses of the Greek educational system, focusing on EFL teaching in Greek Upper secondary education. It then proceeds with the presentation of the profile of the senior high school students target group, including their attitudes towards foreign language learning and the use of ICT for foreign language or computer assisted (henceforth called CALL) language learning. Lastly, the tools which were created specifically for use in my research are presented in detail. These tools include a needs analysis questionnaire, a questionnaire investigating the students’ attitudes towards ICT, the placement tests, the pre, post and delayed-post experimental tests and the creation of a pedagogic corpus.

Chapter 6 presents the application of the proposed DDL methodology in a Greek Senior High school class. It describes in detail the creation of DDL activities, based on the data derived from the pedagogic corpus, ECCo and the free online corpus, COCA, which
would focus on teaching the use of modal verbs in English and the design of five 45-
minutes DDL lessons based on those activities.

Chapter 7 discusses the results of the pre, immediate- post and delayed- posttests whose
design has been described in detail in chapter 5. The three tests were identical and were
administered to the three classes of senior high school students who participated in the
present case study and formed the two experimental and the one control group.

Chapter 8 discusses the findings of the present study. The results of the tests and
students’ performances in this study are discussed in the order of the research questions
presented in chapter 5. The chapter’s last part is a paragraph discussing the learners’
attitudes towards DDL, with evidence based on in-class observations and informal
interviews throughout the duration of the project.

Chapter 9 concludes by suggesting improvements or alterations which could be made
to the DDL methodology proposed in the present study, including the compilation of
the pedagogical corpus, the structure of the DDL lessons and the design of the activities.
It also discusses the theoretical and methodological implications of a DDL approach
and also aims to suggest areas in which our research could expand to in the future.
CHAPTER 2
CORPUS LINGUISTICS – THE THEORETICAL BACKGROUND

2.1 Introduction
There is a growing consensus that the merits of using corpora both for theoretical and applied studies of language are considerable. The usefulness of corpora can be traced in such fields as speech recognition and natural language processing (NLP) (Marcus et al. 1994: 273), historical linguistics, contrastive analysis and translation theory (Meyer 2002: 23) and, most importantly for the present study, language pedagogy through the development of language learning materials, such as dictionaries, grammars and course books (Torsello et al. 2008: 23). It is thus only natural that because of their importance and plethora of uses, corpora would result in the creation of a separate division in linguistics, called Corpus Linguistics.

The present chapter’s first section briefly narrates the history of Corpus linguistics, while the next section discusses corpus compilation and analysis issues, with a focus on the compilation of ‘do-it-yourself corpora’. It also presents the notions of collocation and colligation, while focusing on how grammar is presented and analyzed in corpora. The above issues are of utmost importance for this study, since our experimental case-study focuses on a corpus-driven approach to teaching English grammatical structures, which is based on ECCo (English Course books Corpus) a small do-it-yourself corpus.

2.2 A brief history of Corpus Linguistics

2.2.1 Definition of ‘corpus’ and ‘corpus linguistics’
In 1996, John Sinclair claimed that the linguist’s corpus is “a collection of pieces of language, selected and ordered according to explicit linguistic criteria in order to be used as a sample of the language”. Almost 20 years later, Kennedy (2014: 1) added to the definition the possible uses of a corpus by claiming that a “corpus is a body of written text or transcribed speech which can serve as a basis for linguistic analysis and description”.

Hyland (2006) offered an insight as to the nature of data contained in a corpus by characterizing it “a collection of naturally occurring texts used for linguistic study” (p. 58), while, in one of the most recent studies, Monaghan et al. (2014) claim that corpora are “the linguists’ name for ‘big data’” (as quoted in Christiansen et al. 2016: 144).

A corpus does not present new theories about language, but it can provide fresh insights on features of language use. More detailed is the description offered in Biber et al. (1998: 4) which presents the four main features of the ‘corpus approach’

1. It is empirical, analyzing the actual patterns of language use in natural texts
2. It utilizes a large and principled collection of natural texts as the basis for analysis
3. It makes extensive use of computers for analysis
4. It depends on both quantitative and qualitative analytical techniques

As corpora have since the 1960s been electronic, the definitions offered of them tend to include that characteristic too; Jones and Waller (2015: 5) define a corpus as “an electronically stored, searchable collection of texts”. More precise, we believe, is the definition provided in Charles (2017:107) that a “corpus is a collection of electronic texts designed according to set criteria and compiled for a specific purpose”.

Over the years, various definitions were formed regarding Corpus linguistics. Among others, Corpus linguistics has been described as “the study of language on the basis of text corpora” (Aijmer and Altenberg 2014:1) being concerned primarily “with the description and explanation of the nature, structure and use of language and languages and with particular matters such as language acquisition, variation and change” (Kennedy 1998).
The fact that corpus linguistics is essentially linguistics is unanimously agreed, however there is much debate concerning the status of corpus linguistics as a discipline or methodology; certain scholars, such as Stubbs (1993: 23-24) and Teubert (2005: 2), claim that corpus linguistics is a theoretical sub discipline of linguistics, however the majority of scientists (Curzan 2012: 11, McEnery et al. 2006: 5) agree it is being a methodology, a tool, or “not a branch of linguistics, but the route into linguistics” (Hoey 1998, as quoted in Sampson and McCarthy 2005: 5). A review of the various definitions of the term ‘Corpus Linguistics’ expressed over the years states the following:

“In terms of what corpus linguistics ‘is’, not only have various definitions been offered, but alternatives have been explicitly addressed and rejected. These include, as we shall see: corpus linguistics is a tool, a method, a methodology, a methodological approach, a discipline, a theory, a theoretical approach, a paradigm (theoretical or methodological), or a combination of these”.

Taylor (2008: 180)

As Kennedy (2014: 1) argues, “Corpus linguistics is not an end in itself, but is one source of evidence for improving descriptions of […] language”.

2.2.2 Uses of computerized corpora: past and present

Corpora always posed a challenge for the scholars, before even the term ‘corpus’ was adopted to refer to the systematic collection of texts. The manual concordance of the King James Version of the Bible (18th century) or the analysis of the Quran by the Arabic Grammarians (7th century) must have been enormous tasks for the scholars who undertook the compilation, equal to the creation of modern mega corpora by contemporary linguists.

First-generation corpora were almost exclusively written and contained about 1 million words each. One of the earliest modern computerized English language corpora is the work undertaken by Quirk in the late 1950s, whose aim was to collect data so as to conduct grammatical research. That was the London-Lund prosodically transcribed corpus of spoken British English, recorded from 1953 to 1987, which is the most characteristic case of spoken corpus among the so-called 1st generation corpora. That project resulted in the creation of a popular grammar of standard English, named A
Comprehensive Grammar of the English Language (Quirk et al. 1985). In the early 1960s, Busa organized the compilation of the complete works of Thomas Aquinas, in a project named *Index Thomisticus*, a reference work for students of philosophy and Latin. It was however the Brown Corpus in the 1960s, soon followed by the Lancaster-Oslo-Bergen (LOB) corpus, the one that made a real sensation, since it was the first machine-readable collection of texts.

The second generation of computerized corpora, which did not make use of the obsolete card punching computer technology and used instead the hard drive storage, was launched in the late 1970’s and evolved during the 1980s. The creation of 2nd generation corpora came as an answer to the demands for the creation of dictionaries and grammars. Major publishing houses were actively involved in some of the most influential projects in the field of corpus linguistics, such as the COBUILD project and the British National Corpus (BNC). They contain more than 100 million words each and are made up of both written and spoken language.

The 1990s were marked by the advent of the 3rd generation corpora. One such example is the Bank of English which in essence is the evolution of the 2nd generation COBUILD corpus with the addition of new texts. Other 3rd generation corpora have exceeded the 1-billion-word limit and have been also named ‘mega-corpora’. Notable cases of 3rd generation corpora are the British National Corpus (BNC), the Corpus of Contemporary American English (COCA) and the American National Corpus (ANC). They usually contain written and oral material, different genres and are often freely available and accessible online, either partly or in their entirety.

With the advent of the 21st century, researchers in the field have been pushing the boundaries and addressing the challenges of working with corpora, with the twofold aim to exploit the wisdom gained in corpus linguistics so that it can be of use in various domains and ameliorate the process of building and analyzing new corpora (Curzan 2012: 9). Despite the technical difficulties, the usefulness of collecting, analyzing and manipulating data led corpus compilation and application to new paths. Scientists also conducted research on minority languages (Aduriz et al. 2006: 11), lesser taught languages (Steinberger et al. 2006), to regional, sociocultural and subject-related

---

1 the EPEC written Basque corpus (Aduriz et al. 2006: 1) is one such case
varieties of more frequently taught languages (Davies and Fuchs 2015), as well as English as a Lingua Franca (Cogo and Dewey 2012).

Latest trends in the compilation of corpora include corpora of web-derived data (Ferraresi 2008). Web-derived data, abundant and freely available enable the compilation of giga corpora (Eide et al. 2016). Most recent research focuses on the use of data contained in social media interactions for the compilation of corpora; some of the latest publications are those of Huang et al. (2016) who make use of twitter data, and Passaro, Bondielli and Lenci (2016) who discuss the creation a corpus containing Facebook language.

2.3 Corpus compilation – the case of do-it-yourself (DIY) corpora

The creation of any corpus is always targeted; one must decide on the intended use of their corpus before the compilation procedure begins. This becomes obvious if we remember that corpus is, among other things, a ‘tool’; a tool is always intended to be used in one or more specific works. The main principles governing the compilation of any corpus are representativeness, balance and sample.

*Representativeness*, refers, according to Biber (1993: 243) “to the extent to which a sample includes the full range of variability in a population”. It is determined by the range of genres contained in the corpus (*balance*) and the selection of text chunks for each genre (*sampling*) (McEnery et al. 2006). Since representativeness is of main importance in the creation of large general corpora which are made up of a variety of genres and contain millions of words, it is not of major concern for the purposes of the present study and will only be briefly discussed. On the contrary, what interests us more is the factors and concepts guiding the creation of self-made or “Do-It-Yourself” (DIY) corpora (the term we use here is a loan from McEnery et al. 2006).

2.3.1 Corpus compilation issues

Modern technology, which made powerful computers and abundant electronic data available to the masses, has enabled individuals, teachers and linguists, to compile their own corpora which will aid their teaching or research efforts. Two issues which are of
considerable importance when compiling a corpus, is the size of the intended corpus and the annotation scheme, which may optionally be used. As regards the first, there is no optimum size; large multi-million-word corpora can offer sufficient examples, but so can a carefully constructed small corpus of a few thousand words (Leech 1991, McEnery et al. 2006).

If corpora are useful tools, annotation makes them powerful tools. The linguistic annotation of corpora is a handy mechanism whose purpose is to assign additional morphological, syntactic, semantic information to lexical items. In other words, annotation is the linguistic labeling of the primary data. By primary data we mean the digital language data which is the core of any corpus. The annotation’s role is to enhance the information that could be derived from a corpus by providing structural, lexical, grammatical and ethnographical information (Meyer 2012).

Annotation can take many forms, part-of-speech (POS) tagging, semantic parsing, and semantic annotation being the most prominent ones (McEnery et al. 2006). However, this categorization of annotation levels is not the only one in relevant literature; Aston and Burnard (1998) describe five forms of annotation, namely part-of-speech (or word class), lemmatization and morphological analysis, word-sense, syntactic role and pragmatic annotation; other, less common types of annotation, mentioned in Meyer (2002: 97), include discourse tagging and problem-oriented tagging.

The degree of accuracy in annotation depends on the level of analysis. Nowadays automatic annotators (taggers) are highly accurate when they conduct POS tagging, however semantic annotation, which has to face issues such as disambiguation, often requires the linguist’s intervention.

Various annotation schemes make automated analyses of language possible, a process which is definitely time saving. In pre-electronic corpora, the researcher would have to go through the texts on his own and make searches relying on his own investigating abilities. Nowadays corpora have either built-in annotation schemes (just like the tagged pedagogic corpus ECCo of the present study), or they are accompanied by powerful engines which support the grammatical or syntactic analysis of the data (for example the search interface of the online Corpus of Contemporary American English). Moreover, de Marneffe & Potts (2016) discuss different annotation schemes, such as
naturalistic, gold-standard (trained annotators), automatic (taggers, parsers) and custom annotation projects.

When compiling a corpus, one should not overlook the issue of including relevant metadata records as well. Metadata is the contextual information about the primary data, whose aim is to provide documentation for future users and to serve as a key for the retrieval of particular types of primary data. Both POS tagging and metadata information are provided in the compilation of the present study’s pedagogic corpus.

2.4 Concordances and frequency counts

Once the compilation of a corpus has taken place, linguists/teachers have to exploit it efficiently so as to receive an answer for the research question they had originally set. The computational tool that is used in this case is the concordancer, a type of “off-the-shelf computer program” (Wang 2017: 394) which presents a number of lines retrieved from the corpus in which the search word (or node) is the central word, together with seven or eight other words on its left and right. These lines, called concordances, are the output that we can investigate for language use and manipulate for the creation of word lists, translation memories and teaching materials, among others.

One should not overlook another important function of the concordancer, which is word frequency counts (or wordlists). They are invaluable for the rank ordering of all the words in any given corpus in order of frequency (O’Keeffe, McCarthy and Carter 2007). Within the dense ‘forest’ of corpus data what the concordancer does is “make the invisible visible” (Tribble 1990:11).

Concordancers can be available as software (e.g. WordSmith Tools, Xaira, IMS CorpusBench) or online tools (e.g. BNCweb, Sacodeyl, Just The Word). They can also be offered for free use (e.g. AntConc, MonoConc Easy, TextStat) or become available upon subscription (e.g. Sketch Engine, MonoConc Pro). Lastly there are also those built-in in modern online giga corpora (e.g. COCA).
2.5 Collocation and colligation

Words are not items to be learnt in isolation. They are found in meaningful contexts and according to Firth (1957) “you shall judge a word by the company it keeps”. When words “occur together more frequently than one would expect” (Willis 2003: 45), we say that they collocate. According to Firth (1957: 14), collocation is “actual words in habitual company”, while Kennedy (2014: 108) describes them as “recurrent combinations of English words”. All the aforementioned definitions agree with the statement by O’Keeffe, McCarthy and Carter (2007: 59) that collocations are ‘probabilistic events’, resulting from repeated combinations of chunks in the communication of native speakers.

The tendency for one word to co-occur with another also has grammatical implications (McEnery & Wilson 1998: 14). The collocation of a word or a chunk with a particular grammatical structure is called colligation. The introduction of the notion of a grammatical collocation (colligation) is attributed to Firth, while a more detailed description is offered by U. Römer.

“What collocation is on a lexical level of analysis, colligation is on a syntactic level. The term does not refer to the repeated combination of concrete word forms but to the way in which word classes co-occur or keep habitual company in an utterance”

Römer (2005: 13)

We agree with the claim by Römer (2005) that it is useful to teach a word in its syntactic and semantic contexts. This statement is in agreement with the one by Sinclair (1997) that a closer inspection of the verbal environment of a word or phrase could be very important. This is a credible way to investigate which terms are normally selected together by the competent speaker of English and thus assist our students to focus on language that is actually produced by native speakers.

Previous corpus-based findings regarding the syntactic behavior of modal verbs, discussed in Römer (2004), has been taken into account for the design of DDL lessons in the present study (see chapter 6); one such case is the tendency of auxiliary modal verbs such as can, may, shall, would to enter into colligation with bare infinitive.
2.6 Indirect and direct pedagogical applications of corpora

Methods which make use of corpora have been quite influential in language pedagogy, especially during the last three decades, even though it is true that one of the earliest instances of corpus-based approaches to language teaching, The General Service List (West 1953), took place much earlier.

The applications of corpora to language pedagogy can be distinguished into indirect and direct applications (Figure 2.1). Other terms which may refer to those categories are corpus-based and corpus-driven. What differentiates them is their intended use (Figure 2.2), or, more specifically, whether they will be used for research purposes or direct classroom use.

Figure 2.1: Types of pedagogical corpus applications (Römer 2010:19)

The indirect applications take advantage of the wealth of knowledge available in corpora and exploit them for the benefit of language teachers and learners without really letting either of them get too close to the ‘unknown’ territory of corpora. In this case, linguists, computational linguists, lexicographers and materials writers employ corpora to inform reference materials, such as dictionaries, reference grammars and ELT textbooks, or to revise and inform the teaching syllabus (Flowerdew 2012).
Modern grammars and lexicographical projects make extensive use of concordance output. One well known example of a grammar which was partly based on corpus evidence is the Comprehensive Grammar of the English Language (Quirk et al. 1985).

Lexicography was the first discipline to fully utilize corpus data. The first cases of writers of dictionaries using data derived from some kind of corpora can be traced back in the 18th century. Nowadays, corpora and lexicography are closely linked since dictionary publishers are actively involved in the compilation of mega or giga corpora. Major modern dictionaries (the Collins Cobuild Dictionary, the Cambridge Dictionary of International English, the Longman dictionary of Contemporary English, the Longman Essential Activator and the Oxford advanced Learners Dictionary) were corpus-based (Aston and Burnard 1998) and this trend continues well into the 21st century with major publishing houses being involved in the creation of mega or giga corpora, such the Oxford English Corpus, the Cambridge International Corpus and the Longman Corpus Network.

The direct applications of language corpora, also called Data-driven Learning (DDL)\(^2\), focus on the learning process and investigate how corpus-derived material can be of use in the language classroom, in the form of language learning material.

---

\(^2\) See Chapter 3 for a detailed discussion on DDL
Corpus linguistics has been an invaluable tool in the analysis of grammatical structures in English. This is strongly stressed in Xiao and McEnery (2013: 1) who claim that “research in grammar has probably been influenced most profoundly by the corpus-based approach”. Corpora offered us a better understanding of grammar, by providing accurate information on the frequency of grammatical structures, and an abundance of examples giving us insight as to how language is used. By providing large samples of naturalistic data, such resources complement native speaker intuitions and controlled psycholinguistic methods, thereby putting linguistic hypotheses on a sturdier empirical foundation (de Marneffe and Potts 2017). Investigation of natural occurring language for testing linguistic hypotheses on the grammatical function of words is also suggested by Halliday (1992).

“Fundamental work is needed on the probabilistic modelling of systems in a paradigmatic grammar of this kind. But in my view this effort is more likely to be successful if we first find out more of the facts; and that can only be done by interrogating the corpus”

Halliday (1992: 76)

Furthermore, the abundance of examples of real language in corpora, eliminates the need for the invention of artificial example sentences (Francis 1993). Howatt and Widdowson (2004: 359) suggest that “grammar is subservient to lexis”, but we’d rather say that they stand as equals, they are inextricably linked and they complement each other.

The surge of novel work using corpus methods to study the grammar is underlined by the six conferences on ‘Grammar and Corpora’ organized so far3. Furthermore, corpus data has informed numerous studies on grammatical structures in English. To name but a few of such cases, Ellis et al. (2014) studied second language verb-argument constructions, Nesselhauf and Römer (2012) investigated the case of progressive with future reference, Biber et al. (1998) examined the grammatical associations of nearly synonymous words, while Aston and Burnard (1998), Conrad (2010, 2009) and

---

Kennedy (2014) discussed more general issues concerning the corpus-based methods in grammar description and analysis.

One of the grammatical areas research has focused on is modal verbs, a challenging part of English grammar. The corpus statistics offer solid evidence of the ubiquity of modal items in everyday speech and writing in O’Keeffe et al. (2007), while Nattinger (1992: 36) argued that the syntax of modals can be treated as a lexical phrase, in particular a ‘syntactic string’. The corpus-driven investigation of modal verbs found its way into the language teaching classroom in the work of Römer (2004), who employed a corpus-based analysis of the English modal verbs to provide evidence of the artificiality of the presentation of modals in the EFL textbooks, advocating thus in favour of the contribution of corpus data for the improvement of teaching materials.

Conrad (2000) predicted that corpus-studies will revolutionize the teaching of grammar in terms of a shift of focus from monolithic grammatical descriptions and structural accuracy, to combined grammar and vocabulary learning, attention to use and register. Nevertheless, Timmis (2015) argues that even though corpus research has led to fuller grammatical descriptions, the aforementioned revolution has yet to become a reality.

2.8 Conclusion

The present chapter discussed the ways corpora can be employed in linguistics research, providing thus the basis for a more focused discussion that will take place in the next section (chapter 3), regarding the specific applications of corpora in applied linguistics, in an effort to move beyond description to taking action (Conrad 2016). The use of concordances extracted from corpora for pedagogical purposes, called Data-driven Learning, a methodology which has been claimed to provide discovery teaching and involve learners in “discovery learning” will be discussed in detail. This discussion will provide the theoretical framework underpinning the rationale of the present study.

---

4 For a description of English Modal Verbs see chapter 4
CHAPTER 3

DATA-DRIVEN LEARNING

3.1 Introduction

An ever-growing interest in incorporating Corpus Linguistics into language teaching has been documented over the last three decades (O’Keefe et al. 2007). The direct application of corpora in the language classroom, called Data-driven learning (DDL), which is the central focus of our study, is discussed in detail in the present chapter.

The next section is a short overview of the language learning theories underpinning DDL. Afterwards, DDL is presented in detail; we focus on its characteristics, its potentials and limitations, the different types of corpora involved in DDL and the types of DDL activities and last a review of the empirical studies on DDL is given, with special emphasis on those studies which focus on the teaching of grammar, and those which have been conducted in secondary education.

3.2 Approaches to Language Learning with relation to Corpus Linguistics

Since corpus linguistics is more of a tool or a means than a linguistic theory, it is based on a greater or lesser degree onto the theoretical foundations provided by learning theories that preceded it. In the paragraphs which follow, we briefly discuss the main learning theories underpinning Data-driven learning (or corpus-driven learning), and how they are related to and complement each other.

3.2.1 The inductive approach and the noticing hypothesis
According to the Noticing Hypothesis (Schmidt 1990) language input does not become intake unless it is consciously registered. There are two teaching techniques which can facilitate noticing, one of them being input enrichment, or repeated exposure of learners to the target structure over a period of time, and input enhancement, or emphasizing the target structure (Sharwood Smith 1993). In other words, the means to engage students in successful noticing is the availability of appropriate authentic texts, which “need […] to capture student attention and be comprehensible” (Krashen 1989: 19).

The importance of exposure to and investigation of authentic language is also one of the main principles underpinning the inductive approach. Here, one has to rely on examples of use; the only way to do this is by observing and generalizing. This involves hypothesis formation and then the testing of those hypotheses.

Some educators propose that an inductive process in grammar learning is essential (Seliger 1975) while others advocate for a deductive approach (Shaffer 1989) or are in favor of a combined methodology (Corder 1973, Chujo and Oghigian 2008). Characteristic of this approach is the creation of the Chemnitz Internet Grammar (CING), a learner-centered pedagogical grammar which has two components: a rule-based grammar and a corpus where the learners are able to conduct their own investigations (Hahn 2000).

3.2.2 Autonomous learning

The concepts of autonomy and independence have been much discussed in language learning literature. One manifestation of autonomy is students’ initiation of questions (Loewen 2005), which leads them to seek information about linguistic items which may be problematic for them. This procedure, which draws emphasis on the individuality of the learner (Mishan 2005), can be purely autonomous, or it may involve teachers drawing learners’ attention to certain words, phrases or texts, by providing guidance and support.

The notion of autonomy is central to corpus-driven language learning; as Aston (2001: 23) points out, one of the aims of corpus-driven, or data-driven learning is to help learners achieve greater independence through the development of metalinguistic awareness, which can enhance the learner’s ability to detect frequently occurring
patterns in data and interpret them. Free access to materials and tools, alongside with appropriate training and support on how to use those resources are ways in which a corpus-driven methodology can promote autonomy and learning (Mishan 2005).

The materials should also foster independent learning by raising the consciousness of the learners and making them more aware of the learning process (Nunan 1988). In Rutherford and Sharwood Smith’s (1995: 274) definition, consciousness-raising is claimed to be “the deliberated attempt to draw the learner’s attention specifically to the formal properties of the target language”. This is best done by engaging students into working on tasks focusing on target language.

3.2.3 Task-based learning

Task-based learning differs from the traditional Presentation-Practice-Production paradigm in that students begin by carrying out a specific task and they proceed to compare their findings. In this way their attention is directed towards specific functions of the target language. Willis (1996) defined task as “a goal-oriented activity in which learners use language to achieve a real outcome” and described the cycle of task-based learning, in which the learning content is first introduced by the teacher, followed by a discovery learning task; the last stage of the cycle involves students sharing their findings and focusing on form in order to generalize newfound knowledge.

One of the advantages of the task-based approach is that real communication takes place, since the students are encouraged to work in pairs or groups and share their findings (Bowen 2010). The element of communication along with the fact that the tasks are based on authentic materials are key concepts behind the paradigm of communicative language teaching and learning.

3.2.4 Communicative language teaching

The communicative language teaching (or communicative approach) is a long-established paradigm in language learning. The communicative approach makes use of real life situations. When this approach to language learning is adopted, classroom materials are often authentic, in other words produced for the native speaker audience.
The communicative approach has been claimed to be highly motivational since it offers the learners a chance to get a glimpse of the way language is used in the ‘real world’ outside the classroom (Mitchell 2002). This is the case in corpus-based approach as well, since natural occurring language data mirrors real life situations.

If our aim is not simply to provide a formal description of language but to describe the use of language as a communicative tool, corpora are the means to do so (Meyer 2012). As communicators, we do not proceed by selecting syntactic structures and words to fit into these slots; the natural process is to convey meanings by selecting central lexical items which themselves tend to keep company with certain other lexical items (Francis 1993). This is the essence of the lexical approach.

3.2.5 The lexical approach

Michael Lewis (1993) is the one who first used the term ‘lexical approach’ and who has claimed that “lexis is the basis of language”. He also suggests that traditional language teaching has underestimated lexis by putting grammar first in importance. The “key principle” in lexical approach is that “language consists of grammaticalized lexis, not lexicalized grammar” (Lewis 1993:34); therefore, lexis should be the basis of any well-structured syllabus. In much simpler terms, it expresses the idea that language acquisition is based on the ability to comprehend and produce lexical phrases, or language ‘chunks’; these chunks enable learners to master lexical structures or patterns in the target language, traditionally called grammar structures.

Nattinger (1980: 341) pointed out that language production is the successful comprehension and merging of lexical units. For a learner to successfully engage in meaningful communication, he should be able, among other things, to guess the meaning from context, notice language patterns and make use of reference tools such as dictionaries and language corpora.

Another notion, the Lexical Priming (Hoey 2005), which has been developed over the last decade, connects Corpus Linguistics (the lexical aspect) with Psycholinguistics (the priming aspect). The lexical approach gave rise to pattern grammar, an approach to language learning presented in the next paragraph.
3.2.6 Pattern (or Lexical) grammar

All words can be described in terms of their pattern. This is what the linguists working on the COBUILD dictionary project in the 80’s advocated when they created Pattern Grammar, a model for describing the interaction between lexical items and the grammatical patterns they form a part of. Pattern grammar (also named Lexical grammar in Francis 1993) has been described as an approach based on corpus investigation, which encourages the observation of items that occur together with “a high relative frequency” (Hunston et al 2014: 101). Francis and Hunston (2000) argue in favor of using a corpus to reveal the typical patterning of lexical items, since as they state “intuition is not always a reliable guide” (p. 3).

3.2.7 Blended learning

Over the last years, digital and online resources have been used as complementary to the traditional teaching/learning paradigm, giving thus rise to a novel approach in teaching called blended learning. Blended learning, also called hybrid learning and mixed mode learning (Spring et al 2016: 88), has been named the mixing and matching of face-to-face and online techniques in the learning procedure (Picciano 2013). Blended learning, in one of its forms, is related to educational technology and focuses on those contexts where teacher and learner are co-located (Graham et al 2013).

Its advantages over traditional teaching approaches are numerous; blended learning fosters learner autonomy since, at least some part of learning involves direct access to online material which they have to investigate in order to answer questions or do activities. It also promotes computer literacy since it requires that learners spend a considerable amount of time exploiting online resources. This has led to an encouraging surge of research, the majority of which focuses on tertiary education, but with a few number of studies existing which have investigated its potential in school education settings (Halverson et al. 2017).
3.3 The Data-driven Learning

Corpus linguistics has provided a multifaceted methodology to language learning and teaching, which may be called “corpus based” or “corpus driven” depending on whether the teacher makes use of corpora as sources of knowledge to inform his decisions about the teaching procedure, or he/she uses corpora directly as learning material in class.

The term DDL was coined by Johns (1991a, 1991b, 1994, 2002) who made use of concordances in the English for Specific Purposes (ESP) classroom. According to the definition provided by Johns, DDL is an approach which relies on the ability to notice patterns in target language and form generalizations about language form and use (Johns 1991a: 5). Johns (2002: 24) saw DDL as a process which “confront(s) the learner as directly as possible with the data”, and whose aim is “to make the learner a linguistic researcher” in a classroom where “every student is Sherlock Holmes”.

Various other definitions have been given to DDL during those 25 years. One of the most recent ones is that by Boulton (2016: 268) who claims that DDL “involves the use of dedicated concordances to explore large language corpora”. Much broader is the definition by Gilquin and Granger (2010: 359) that DDL is “using the tools and techniques of corpus linguistics for pedagogical purposes”.

As implied by Johns’ aforementioned definition, the ‘center-pin’ (Mishan 2005) of DDL is research activity. It is essentially an inductive approach to language learning since the learner uses the instances of language use contained in a corpus as the starting point in order to make inferences about language use and reveal the meaning of words or form grammatical rules (Boulton 2011). In the DDL approach, the focus lies on the learner; the learner is the agent who investigates language, finds clues about the meaning and relations between linguistic items and comes up with findings regarding meaning and rules. In this environment, the teacher is the facilitator of the learning process, coaching the learners on the best ways to form queries and interpret results (Anthony 2016).

For example, in order to teach grammar inductively, we offer relevant texts to the students and guide them through relevant tasks to discover the grammatical structures in those texts. This approach to language learning is illustrated in Johns’ cycle of DDL (Figure 3.1).
This process of discovery encourages the exploitation of finite (in the case of hand-made corpora) or infinite (in the case of the Web) resources. In other words, the backbone of the DDL methodology is the learner-conducted investigation of authentic materials in class, even though it can often be teacher-guided. The focus is on the learner; he/she is the one who asks the questions and searches the available resources for answers and then shares the findings with the rest of the class, with the instructor keeping the role of facilitator by providing input and encouraging pair and group work for better results. DDL has also been claimed to promote “grammatical consciousness-raising” (Rutherford 1987), since the learners, by looking at a number of examples from the concordance lines, familiarize with lexical items and get useful insights into the formation and usage of grammatical structures (Lee 2011).

A DDL methodology may be direct or indirect. The indirect use of corpora mainly influences syllabus design and teaching materials, and it is therefore more closely related to researchers. In the direct approach learners and teachers are actively involved in DDL by studying corpus-derived texts in order to discover language patterns and word meanings (Bernardini 2002).

A further categorization of DDL involves the way concordance and frequency list output is used in class. When learners explore corpora and form hypotheses without the direct assistance of a teacher, DDL is called inductive. It has also been called ‘pure’ or ‘hands-on’ approach. In the case of instructor guided learning, the DDL is deductive or ‘paper-based’ (Johansson 2009: 42). In deductive DDL the teachers usually conduct the investigation of corpora in search for appropriate texts, selects the concordances,
possibly edit them as well, taking into consideration the proficiency level of their students and then offer them for use in class, printed on handouts. In other words, different approaches to DDL depend upon whether we opt for unmediated or mediated use of corpora (Thomas 2016) in the language classroom.

Certain scholars such as Johansson (2009) suggested a combined deductive and inductive approach, appropriately tailored to the needs of the students. In a similar wavelength is the middle-ground methodology suggested in the present study; we call for a blended DDL approach which employs both inductive and deductive DDL elements; on the one hand, we employ paper-based and teacher edited DDL activities and, on the other hand, direct student access to corpora, teacher-supervised formation of searches and student-led interpretations of the results. This blended DDL approach, similar to the ‘web-facilitated blended learning approach’ (Graham et al 2013), is illustrated below:

Figure 3.2: The blended DDL learning approach
There have also been those voices who advocate for a combined traditional EFL learning and DDL-based EFL methodology. Such is the claim by McCarthy and Carter (1995) who argued that learners may need to supplement an inductive approach in addition to a deductive approach offered by the Presentation- Practice- Production paradigm, and propose their three Is: illustration, interaction and induction.

A positive sign of the interest that DDL has stirred in the language learning community is the biannual *Teaching and Language Corpora conference* (TaLC). Furthermore, several Master dissertations and Doctoral theses have dealt with different aspects of the pedagogical applications of DDL, focusing mainly on the learning of vocabulary (Balunda 2010, Shaw 2011, Otto 2017), grammar (Rapti 2010, Yunus 2015) syllabus design (Redrupova 2009), writing skills (Sepehri 2015) and students’ evaluation of DDL (Sripicharn 2002). It is also worth mentioning the increasing availability of pedagogically orientated concordancers, such as www.lextutor.ca, https://www.sketchengine.co.uk/skell/, www.just-the-word.com and http://flax.nzdl.org/greenstone3/flax.

3.4 Corpora for DDL

The heart of DDL methodology is the corpus. A corpus is a tool, a means to achieve the learning goal; therefore, its very existence is defined by the language use which is intended by the syllabus of any given course. Any corpus material should be consciously selected with a specific purpose in mind. Meyer (2002: 28) states that corpora can be valuable resources in order to improve foreign language instruction and should therefore “be compiled in a manner that ensures that will be most useful for their potential users”. Moreover, Chujo and Oghigian (2010: 201) claim that “the first and arguably most essential consideration for classroom DDL exercises is choosing the appropriate corpus with regard to size”, while, Braun (2010: 92) asks for “pedagogically viable corpora” that would be ideal for use in the foreign language classroom. Last, Van Els et al (1984) point out that any corpus-driven approach should take into account the goals of the learner and “be representative of the language use which is defined in the objectives (of the course)” (p. 204).

Bennett (2010: 13) identifies eight types of corpora – generalized, specialized, learner, pedagogic, historical, parallel, comparable and monitor, with the first four being most
useful for employing the corpus approach directly in the classroom. These are the ones to be discussed in the following paragraphs, with the addition of the Web as a corpus which may not be prominent in (especially older) bibliography, but it is an emerging type of corpus that deserves our attention and was employed in the methodology of our study. Last, since our study primarily focused on the use of a pedagogic corpus, greater emphasis is to be given to this type of corpora.

3.4.1 The Web as a corpus

Many learners around the world are already searching the web in ways not dissimilar to DDL (Boulton 2016: 288). Whether the Web is a corpus or not is hotly debated. Meyer (2002: 28) referred to the Web as “the ultimate corpus”, Kilgarriff & Grefenstette (2003: 334) justified their claim that “the web is a corpus” based on their prior definition of the corpus being “simply a collection of texts”. Lew (2009: 290) claimed that the Web is in essence a “corpus-like resource”, whereas Sinclair, one of the most prominent scientists against this notion, argued that the Web cannot be a corpus, because its dimensions are unknown and constantly changing, and also because it hasn’t been designed by linguists (2004: 4).

Whatever the case might be, a great many studies have pointed out the merits of using the Web as a corpus, even if, technically, it might not be one. Despite its ‘noise’ which might be created, among others, by the presence of tokenization errors – that is the errors which occur when replacing data with symbols-, misspelling and non-words (Schäfer and Bildhauer 2013), the sheer size of the web means that web data often give results that are close to traditional corpora (e.g. Rohdenburg, 2007), and even to native-speaker judgements (Keller and Lapata, 2003). Despite Sinclair’s criticisms as these have been mentioned above, he also recognizes that “the web itself… [is a] huge source of language that is available in the classroom or the study at home” (Sinclair 2004: 297).

If the Web is the world’s favourite ‘unofficial’ corpus, then “Google, as a quick n’ dirty corpus tool” (Robb 2003) is the most widely used concordancer. People use it for language-learning reasons, for example to search for the meaning or grammatical
behavior of a word, or to come up with examples of use of a specific word or chunk. Evidence that Google is already being used in this way for language teaching and learning is provided in Clerehan et al. (2003) and Conroy (2010). Googling is so popular among language learners that it seems it can foster language awareness and critical thinking about language (Milton 2006), despite the messiness of web data and the limitations of search engines.

Googling the Internet is thus not entirely dissimilar to DDL. It is a practice which may be actively encouraged by the teachers because it is attractive, it offers a familiar, intuitive, and easy way to begin simple DDL which brings immediate benefits (Shei 2008). Googling may encourage the perception of DDL as ordinary practice (Boulton 2016: 288) and thus provide a handy first step on the road to more ‘hard-core’ DDL (Conroy, 2010).

The merits of using the Web as a corpus are evident, however there are scholars (Wu et al. 2016) who look at the web-as-a-corpus approach with a critical eye due to the fact that such an uncontrolled collection of texts may contain grammatical errors, inappropriate content or be at least daunting for the average, especially the low-level, learner. In order to compensate for these shortcomings, Wu et al. (2016) propose the construction of a collocation learning system, FlaxCLS, from a Wikipedia corpus, claiming that Wikipedia contains far less grammatical errors than the Web.

3.4.2 General corpora

General (or generalized (Bennett 2010), large scale (Flowerdew 2012), large reference (Tribble 2001), large mixed (Aston 2001)) corpora are the broadest type of corpus. Their two distinguishing features are first, that they are often very large and secondly, that they contain general language, usually from many sources and various registers.

Large general corpora which are said to have exceeded 1 billion words of text are nowadays the British National Corpus (BNC), the Corpus of Contemporary American English (COCA), the Cambridge International Corpus and the Oxford English Corpus. These large corpora have won over the years the appreciation of linguists and teachers because they offer large amounts of genuine, unedited material and they can be
exploited as reference tools to check our intuitions about language (Bernardini 2001:221).

Since the late 1980s general corpora have been used as the basis for the creation of grammars and dictionaries. The Collins Cobuild grammar (late 1990s) is such an example. Other corpus-influenced materials include textbooks; one such example is the Touchstone series (Mc Carthy et al 2005) which includes material based on the Cambridge International Corpus of North American English.

Due to their easy online accessibility, large general corpora, such as the BNC and COCA have often been employed in the classroom (Charles 2017). The studies which made use of large corpora report the merits of using them, these being their user friendliness, accessibility, abundance of natural occurring language, variety of registers, existence of built-in concordance tools and detailed annotation schemes (Tribble 2000, Shaw 2011, Boulton & Landure 2016, Charles 2016, Noor et al 2017, to name but a few). However, Chujo et al (2013) argue that it is difficult “to capitalize on large corpora [such as COCA and BNC], because [they] would expose students to completely authentic, native speaker language […] that would be far beyond the learners’ levels” (p.68). This is the reason why a great number of scholars, eager to employ corpora in the classroom, turn to specialized and pedagogic corpora, or opt for a blended approach incorporating more than one type of corpora, similar to the case of our study.

3.4.3 Specialized corpora

A specialized (also called exemplar (Tribble 2001), or dedicated (Noguchi, Orr and Tono 2006) corpus is characterized by its small, in most cases, size and also by its content; it consists of texts of a certain type and aims to be as representative of this type of language as possible (Bennett 2010). A specialized corpus offers advantages from a DDL perspective due to its focused content which addresses specific target groups. Aston’s (1988: 8) definition says that a specialized corpus “above all, is targeted; it focuses on a particular group of learners, in particular settings”.

While the majority of studies have focused on the use of small specialized corpora, usually compiled by the “researcher-cum-practitioner” (Flowerdew 2012), there are
distinguished cases of much praised large specialized corpora. This is the case of the Michigan Corpus of Academic Spoken English (MICASE), or the CHILDES corpus of computerized transcripts of utterances by children, which comprises 130 subcorpora.

Despite the few aforementioned exceptions to the rule, specialized corpora tend to be small in size and occupy the biggest part of the literature on the applications of corpora on pedagogy. Aston (2002) claims that such corpora seem to offer practical advantages over large general corpora; they are easy to compile, even by ordinary teachers and also, using the web, it is relatively easy to gather texts on a particular subject. Moreover, they are often simpler to analyze, and if they contain texts of a specific domain, they are more likely to reveal recurrent patterns.

A great number of empirical studies have reported the use of specialized corpora; Tribble (2001) discusses the merits of using a small specialized corpus for teaching writing skills to adult learners, Braun (2007) created a small experimental corpus (ELISA) consisting of 25 video-based narrative interviews of native speakers of Australian English to teach vocabulary and cultural information about Australia, while Reguzzoni (2008) made use of a specialized self-made corpus of Maritime English to teach an ESP class. In a more recent case study, Kitao and Kitao (2013) employed a corpus of DVD subtitles developed from the first three seasons of the US situation comedy Modern Family, in order to teach speech acts.

In contrast to the aforementioned studies which involve mostly small specialized corpora to be used with specific scientific target groups, Noguchi et al. (2006) discuss the development of a major international Corpus of Professional English (CPE). CPE may technically be a specialized corpus, but it stands out because of its size, since the ultimate goal, as reported, is to reach 100 million words and secondly because it addresses the needs of a much broader target group, in the sciences and other disciplines.

3.4.4 Learner corpora

A learner corpus is a kind of specialized corpus that is made up from the oral or written production of language learners. A fairly recent description of learner corpora by Flowerdew (2012) points out the following characteristics: Learner corpora may be oral
(e.g. the Vienna-Oxford International Corpus of English (VOICE)) or written (e.g. the Trinity Lancaster corpus), they can also contain language samples by native speakers, or EFL learners and finally they may address L1 or EFL learners.

As regards their use, Granger (2009: 20) classifies learner corpora into two categories; those intended for delayed pedagogic use (DPU) by a similar population of learners to those who have produced the data and those for immediate pedagogic use (IPU), produced and used by the same group of learners. DPU corpora tend to be larger and more generalized. They are usually compiled by linguists, with the aim to either study learner language or to produce reference materials for a specific learner group. IPU corpora are directly related to the actual teaching procedure, since they are created by teachers and the learners are both the producers of the corpus content and the beneficiaries of its in-class use.

Research on learner corpora has influenced three domains in language pedagogy; curriculum design, the production of reference tools and classroom EFL teaching (Meunier 2002). The learner-corpus use for language learning purposes has been reported (Cotos 2014) to slowly but surely make its way toward the classroom. Cases of DIY learner corpora have also been reported; one such case is YoLeCorE learner corpus, which has been created with the aim to record and analyse young learners L2 input and output (Zapounidis 2017). The scientific community’s belief concerning potential and significance of learner corpora is reflected in the birth of the term Learning Driven Data (LDD) (Seidlhofer 2002). LDD describes a language learning methodology which uses a “secure ‘home base’ through focusing on familiar, non-threatening texts” (Seidlhofer 2002: 230) of learner corpora, in order to promote the language learner’s autonomy.

Learner corpus research is one of the most dynamic domains in applied corpus linguistics, with numerous such corpora compiled in several countries, a comprehensive list of which is provided in the informed webpage of the Institute for Language and Communication of the Université catholique de Louvain5. Here we should also mention the significant number of scholars reporting the empirical evidence of learner corpus

---

use in classroom (Nesselhauf 2004, Cotos 2014, Hardy et al. 2015, de Souza Guedes et al. 2016 are only a few of them).

3.4.5 Pedagogic corpora

The practice of compiling corpora for various purposes may be centuries old, but the first use of such resources with the aim to teach and learn a foreign language can only be traced back to the middle of 20th century. One of the earliest attempts is the 2.5 million words corpus whose compilation resulted in the publication known as The General Service List (GSL; M. West 1953). GSL was a list of around 2,000 high-frequency words of important vocabulary that could be of service to learners of English as a foreign language. This may well be one of the first pedagogic corpora.

As suggested by their name, pedagogic corpora are intended for teaching and learning. Textbooks and other available print resources can be scanned to allow the compilation of pedagogic corpora which are used to examine language patterns and teach vocabulary, grammatical structures or writing and culture. Pedagogic corpora stand out from the other aforementioned types of corpora because they are characterized by two unique features. Firstly, they are created with a particular target group in mind. In order to create a pedagogic corpus, one has to set clear educational goals; what to teach, how to teach it, as well as the special characteristics of the learner target group. Secondly, since these corpora are intended for classroom use, they are most often created by the teacher, who may also be a linguist.

Since their first steps in the 1950s, pedagogic corpora have been extensively used for foreign language teaching and learning purposes. Most of them involve teaching English for specific purposes, with a smaller number targeting general English.

Chujo et al (2006) discuss the design criteria of a specially developed Japanese-English newspaper corpus which was created to teach vocabulary to low-level students of English. Reguzzoni (2008) reports the compilation of an in-house pedagogic corpus of maritime English. Chujo et al (2013) created a simple corpus from standard school book texts to teach low level (remedial) grammar to university students. There have also been study cases involving other target languages, such as the one reported in Tyne (2012:114) who compiled a pedagogic corpus of newspaper articles in Spanish.
The majority of case studies on DDL with pedagogic corpora report the compilation of small specialized corpora. The size issue can be easily explained by the limitations set upon such endeavors, namely lack of time and resources for the ordinary teacher. Most empirical studies of exploiting specialized corpora were conducted in tertiary education settings and the educational goal was usually the learning of relevant terminology. An exception to the rule of small pedagogic corpora which focus on terminology is the New General Service List (NGSL). NGSL, which was created in 2013, enhanced the former NGL (West 1953) and now includes 2800 high frequency words that can be of ‘general service’ in the foreign language classroom. It was derived from a 244-million-word subsection of the 2 billion-word Cambridge English Corpus (CEC).

One should not oversee the advent of such pedagogic corpora as ELISA (Sabine Braun 2005) which adopt a consistent pedagogical conceptualization of the entire corpus process from compilation to annotation, enrichment and search. Its aim is through contextualisation to lead learners to effective communicative interaction in similar instances of discourse. ELISA’s concept was adopted by the European Minerva project SACODEYL (2005-08) (Pérez-Paredes 2010). In one of the most recent studies involving the use of pedagogic corpora, Wu et al (2016) describe the construction of the online collocation platform FlaxCLS which draws material from Wikipedia articles exclusively for pedagogic purposes.

3.5 Typology of DDL activities

DDL activities are created with the concordancer output, which may come in three different forms; frequency lists, keyword-in-context (KWIC) lines and samples of texts. The aim is to observe recurrences and induce generalizations, through “student-centered exploratory learning activities” (Luo 2016: 1255) which stimulate discovery and motivate learners.

In paper-based DDL, concordance output is selected, edited and printed by the teacher and then offered to students for analysis. One of the earliest cases discussing paper-based DDL application in class was the one by Johns (1991). Johns created a handout of concordances which included example sentences with different senses of the modal ‘should’. After they had examined the concordances, the students were to determine
what labels could be given to the categories, and possibly what the relationship between them might be.

**Figure 3.3:** Concordance output used for investigating grammatical structures (Johns 1991: 7)

In paper-based DDL, once the teacher has searched the corpus, s/he can edit the concordancer output so as to create activities suitable for the needs of the learner target group. Those tasks mainly ask learners to investigate concordances for recurrent patterns and generalize (Tian 2005); those tasks may be gap-filling, matching, or multiple-choice tasks.

In hands-on DDL, students conduct the corpus queries themselves, and the output of those searches has a quite different layout. One such approach is described in Noor et al. (2017) and involves hands-on activity in which the learner searches an online corpus for the different adjective collocations of the word ‘effort’ (Figures 3.4 and 3.5).

**Figure 3.4:** Collocate search for *effort* taken from COCA
3.6 DDL potentials and limitations

Much hope has been placed on DDL over the years; many studies reported the findings of using this “bottom-up” (Flowedew 2008: 14) approach in actual classes. The findings of those studies bring into light both the potentials and the limitations of this approach to language learning.

3.6.1 DDL potentials

DDL brought a major change in language pedagogy; rather than being spoon-fed the knowledge, students learn by discovering. They are not passive recipients and the teacher acts as a research facilitator rather than a transmitter of knowledge (Jiezhi and Jianhua 2009). Using computers as a powerful hypothesis-testing device helps “learners remember the knowledge which they have formulated themselves rather than formulations which have been imposed on them” (Lamy and Mortensen 2017: 1). Moreover, working with corpus data can support intuition and memory for the better understanding of problematic linguistic items, such as idioms (Gavioli 2001: 131). Therefore, each learner becomes a researcher and assumes responsibility for their own learning.

It has also been claimed that concordancing activities are fun and amusing (Lamy and Mortensen 2017); this can make learning more attractive, especially for low level students. Furthermore, working on online DDL platforms can be ideal for mixed level
classes because it allows learners to work at their own level and in their own ways (Willis 1998).

One of the greatest attractions of DDL is its strength to promote autonomous learning (Aston 2001, O’Keeffe, McCarthy and Carter 2007). Using concordancers can stimulate users to induce their own generalizations (Gavioli 2001); it also helps them become more aware of language and pay more attention to forms rather than simply use them. Also, there is the likelihood that this awareness will increase proficiency (Fox 1998).

Analysis activities based on concordance lines can be used as consciousness-raising activities, providing learners with opportunities to discover language before expecting them to produce it. Furthermore, DDL material is an excellent source of ‘authentic data’ (Barlow 1996, cited in Partington 2001); those examples of authentic language use can help learners note the gap between their use of a word and the way(s) native speakers use it (Willis 1998).

Not only learners but also teachers are likely to gain from the wealth of knowledge contained in corpus data. Working with corpora can effectively increase their competence and awareness, enrich their understanding of the language and enhance their ability to critically evaluate other sources of information (Aston 2001, Partington 2001).

Regarding DDL’s advantages, especially as far as grammar learning is concerned, there have been scholars who pointed out that such a pedagogy may help students to approach grammar communicatively (Tognini-Bonelli 2001) and also provide learners with “relevant, authentic, and interesting examples as opposed to made-up traditional ‘grammar examples’” (Lamy and Mortenson, 2012: 11) a belief also shared by Tsui (2004) who claims that DDL fills the gap of limited exposure to target language.

3.6.2 Limitations of DDL

Gilquin and Granger (2010: 366) identified four areas which could prove problematic for the effective application of a DDL methodology; “logistics, the teacher’s point of view, the student’s point of view and the content of DDL”.

36
Any DDL-inspired methodology is essentially computer based. At least one computer with internet access is needed to compile and search a corpus, as well as to create the relevant activities. The cost of DDL rises in those cases where subscription to concordancing software is selected by the teacher. Even more demanding is the case of a pure hands-on DDL approach; here the students demand individual, or in pairs, access to computers and a well-equipped pc lab is required. This can be overwhelming for the school budget and timetable as well, since often one pc lab is shared among all teachers.

The bulk of empirical case-studies on DDL have focused on tertiary education, and even then, not on mandatory courses, rather than on few elective courses, especially when direct corpus access by students is concerned (Granath 2009). The reason behind this may be the teachers’ fears that DDL may be an approach suitable only for advanced, highly motivated, computer literate students. Even the pioneer of DDL himself, Tim Johns, reported that fellow teachers often objected that DDL “may be all very well for students as intelligent, sophisticated, and well-motivated as [the ones] at Birmingham University, [but] it would not work with students as unintelligent, unsophisticated and poorly-motivated as theirs” (Johns 1991: 12).

Teachers have been said to favor deductive teaching methods over inductive ones, because the latter are thought to be time consuming (Sun and Wang 2003). Another point made by Mauranen (2004) is that ordinary teachers are not familiar with the nature and potential uses of corpora and she reports that it is “surprisingly hard to convey the logic of what corpus data is like, how a corpus works and what kinds of questions can be asked” (Mauranen 2004: 209). Also, introducing corpora in the EFL classroom may lead the teaching process into unknown paths, since any discovery-based learning calls for the unpredictable. Flowerdew (2009) mentions several teachers’ concerns that a corpus presents language out of its original context and stresses the importance of pedagogic mediation to contextualize the data for the students. The aforementioned limitations of DDL call for appropriate teacher training, as well as for actions to convince teachers to use corpora (Romer 2006).

Regarding learners’ perceptions about DDL, it is a fact that the majority of students are not familiar with concordancers and they may thus feel more comfortable studying a language the traditional way, rather than by engaging in a more active role (Chujo et al
2012). This calls for appropriate student training on DDL, first with a softer version of DDL activities and gradually with more autonomous tasks.

These limitations may account for the fact that, even after 30 years of DDL research and in-class application, the regular use of corpora in the language learning classroom is still a rare occurrence (Granath 2009)

3.7 Review of studies on DDL

We name empirical (or systematic) the study which has been conducted based on observations through an experiment, survey, or other scientific means. Winston and Blais (1996) also mention that in an empirical study there is “no mention of control or manipulation” (p.601).

A review of studies in the field of DDL by Boulton (2007) presented 39 empirical case studies. More recent is the meta-analysis of studies of corpus-based learning by Cobb and Boulton (2015), which identified “132 papers which seek to evaluate some aspect of corpus use in foreign or second language (L2) learning and teaching” (p. 483). Mizumoto and Chujo (2015) added to the aforementioned report the case-studies conducted in the Japanese classroom, proving thus that a great more research has taken place in this domain.

There are two categories of empirical research; qualitative and quantitative. The former is primarily exploratory and uses unstructured or semi-structured techniques such as individual interviews, group discussions and observations. Its main aim is to provide insight into ideas, opinions or motivations. Examples of recent qualitative case studies on DDL can be found, among others, in Kotamjani et al (2017) and Yoon and Jo (2014).

The quantitative research (see Chujo et al 2016, Wong and Lee 2016 and Paker and Özcan 2017 for recent quantitative case studies) differs regarding both its goals and methods; it generates data, transforms it into statistics and gives numerical evidence and answers to the research questions. Its tools are tests, structured questionnaires and surveys, from which numerical data can be drawn (Wyse 2011).

Research on DDL has been based on both types of case-studies, however quantitative empirical research is sadly “still limited and narrowly focused” (Vyatkina 2016: 207). On the same wavelength, Tomlison and Masuhara (2010) comment on the lack of
Experiments in which a control class makes use of commercial material and an experimental group follows the same syllabus with materials designed to meet the (language) learning needs of the specific group of learners and stress the usefulness of a comparison in the improvement of those two groups through appropriate pre and post-tests.

3.7.1 Cases involving grammar teaching through DDL

Corpus linguistics has provided an invaluable means to inform our knowledge of grammar and to critically assess the existing grammar books (Folse 2016). Focusing on DDL, Johns (1991: 3) stated that “one of the effects [of DDL] is a re-evaluation of the place of grammar, not only in grammatical analysis, but also in language-learning and language-teaching”. In a relatively recent meta-analysis, Römer (2010) provided a list of studies which made use of corpus-derived data to address problematic areas in EFL grammar learning; among others, focusing on reflexives (Barlow 1996), linking adverbials (Conrad 2004), modal verbs (Römer 2004), if-clauses (Römer 2005), progressives (Römer 2006) and passive voice (Conrad 2016).

Significant is also the direct or indirect in-class exploitation of corpora for grammar learning. Estlingvannestal and Lindquist (2007) report that the application of a dual paper-based and hands-on DDL approach to 1st year university students for the learning of English grammar has been effective as it boosted the student’s motivation and helped them assume responsibility for their own learning. Chujo et al (2013) illustrate how low level (remedial) grammar to university students can be taught with an ad hoc corpus. Tertiary education students were also the target group of the experimental study by Tian (2005) whose DDL-based methodology focused on the teaching of subjunctive verbs and reduction of adverb clauses to modifying adverbial phrases. The role of paper-based DDL for grammar learning is discussed in Smart (2014), whereas Lin and Lee (2015) present the teachers’ evaluation of a DDL methodology for the teaching of grammatical structures. The growing interest in the use of corpora in the classroom is also pointed out in the study by Alruwaili (2016) which focuses on the teaching of general verbs the DDL way. In one of the most recent studies, Akinci and Yildiz (2017) report the effectiveness of data-driven learning (DDL) in teaching verb + noun (V+N) collocations to advanced Turkish learners of English.
A DDL methodology has also been adopted to teaching grammatical structures of languages other than English, with positive results. In one of those cases, Vyatkina (2016) reports the effectiveness of paper-based DDL for teaching verb-preposition collocations in German, especially as far as new collocations are concerned.

3.7.2 DDL case-studies in secondary education

The majority of studies on corpus-driven language learning are carried out in tertiary education environments, with adult and most often quite advanced students. That was the case of the first DDL empirical studies by Johns in the early 1990s, and the researchers who followed seemed reluctant to abandon this trend.

We believe there are two reasons to explain the researchers’ attachment to experiments carried out in university settings. The first is owing to the fact that the researchers who have reported relevant empirical DDL study cases are employed in tertiary education and thus have access to college-level learners. The second reason could be that learners in tertiary education are (most often) adults and their knowledge level of the foreign language is advanced. This makes it easier to deal with such unfamiliar notions as corpora and concordancers. Davies (2004: 260) justifies his choice to work with advanced students by saying that such an approach shows “a more complete picture of the different ways in which students can use large corpora to study and analyze the grammar of the second language”. However, there are also researchers such as Tyne (2012) who argue that it would be feasible to deal with younger students who learn languages at school as part of the curriculum.

Kaltenböck and Mehlmauer Larcher (2005: 80-81) call for a cautious approach when using corpora in a foreign language classroom, “if we want our students to fully benefit from a DDL approach overcoming any obstacles set by their age, computer competence or level of autonomy”. On the other hand, Sinclair (1997: 30), in the very early stages of DDL, made the optimistic prediction that “young learners will gain access to the corpus and will become self-taught DDL students”.

The absence of DDL empirical studies in secondary education could be justified by the lack of relevant teacher training and the lack of expertise in corpus consultation by learners (Braun 2007). In a similar wavelength is the remark by Widdowson (2003) that
corpora merely contain textual records of discourse situations, and the (re)contextualisation of these records, which is crucial if anything is to be learned from them, can be difficult and requires pedagogical mediation; this also calls for appropriate teacher and learner training on interpreting and using corpus data.

Given the curricular requirements in the school context, the aforementioned ‘re-contextualization challenge’ could be addressed by corpora which would also be complementary to school curricula, to facilitate both the contextualisation process and the practical problems of integration (Braun 2007). As suggested in Braun (2005), one way of achieving this is through the ‘enrichment’ of corpora with support materials such as corpus-based learning activities and exercises to support the construction of knowledge and provide learners with opportunities to practice and test their knowledge. This is of great importance if we want to expose our students to real data and to offer them a learning environment in which they will engage into tasks that will teach them to learn critically, form their own hypotheses, ask questions and search for the answers in the wealth of data provided in corpora.

Among the, admittedly, very few case studies reporting the application of a DDL methodology in Secondary education settings is the one conducted by Braun (2007) who used a corpus of interviews by native speakers to teach vocabulary in Secondary education. The usefulness of this methodology was evaluated by comparing the results of computer-based tests between a corpus group and a control group. The findings of Braun’s study showed that the corpus group scored significantly better than the control group.

Vocabulary was also the learning target of the experimental study by Yilmaz and Soruc (2015) who involved forty EFL learners, aged between 14 and 16, in a private English language education center. In this study, pre and post-test results revealed that the experimental group, which was taught investigating concordances, performed better than the control group. Another finding was that this difference in performance was statistically significant.

In another study, Lee (2011) exploited a corpus of Harry Potter books in order to teach prepositions to a class of Taiwanese students. In this study case, the process of creating a self-compiled corpus is discussed, as well as the advantages and disadvantages of classroom concordancing, without any quantitative or qualitative evidence. Similar
seems to be the case in the study conducted by Fischer-Starcke (2008) which reports the use of corpus linguistic techniques in Secondary school foreign language literature teaching. Fischer-Starcke’s study discusses the benefits and challenges of exploiting wordlists, keywords and concordance lines for classroom use, without offering any quantitative evidence regarding the effectiveness of the approach. The possibility of future application of a DDL methodology in Secondary education is also discussed in Heyvaert and Laffut (2008). They discuss the findings of a questionnaire investigating Secondary education Flemish teachers’ attitudes towards corpora and corpus-assisted learning. Their research reports the Flemish teachers’ call for targeted workshops which will allow them to familiarize themselves with corpus techniques.

In a later empirical research, Tekin (2015: 79) followed the DDL methodology to teach vocabulary to teenagers using the BNC with promising results; his quantitative evaluation of the experiment showed that in the post-test the experimental group performed significantly better than the control group. Another study by Castillo (2016) reports the implementation of a DDL paper-based methodology for the teaching of general English to low-level 10th grade students in Ecuador. Castillo used quantitative methods, pre and post-tests and questionnaires, in order to explore high school students’ responses and perceptions to corpus-based vocabulary learning. The findings of this study reveal that the use of corpus-based techniques improved the level of the students’ vocabulary.

In one of the most recent studies, Lee (2017) discusses a DDL approach to teaching writing and grammar in the Korean secondary school ELT classroom. More specifically, this study presents a certain number of concordance-based activities which were designed to remedy students’ most frequent error types; the evaluation of this approach is done through in-depth interviews and questionnaires, designed to investigate the students and teachers’ attitudes towards DDL. As reported by Lee (2017), the students and teachers’ responses towards the DDL methodology were “relatively positive” (p.1). The challenges of introducing public school teachers to the applications of concordancing were also presented in a much larger scale project, called project MORE, which was conducted by Davis and Russell-Pinson (2004) and involved 29 middle schools and 1,100 students. In project MORE, a corpus of 600 oral narratives of native and non-native speakers was compiled and ESL-relevant classroom activities
were produced. The teachers’ responses to questionnaires investigating their attitudes towards corpus-based materials reveal the need for more in-service training.

The only study case reporting the implementation of a DDL approach in Greek secondary education settings is the one by Giagkou and Antoniou-Kritikou (2008) and Antoniou et al. (2010) which discusses a small-scale experiment of data-driven learning of Greek as mother tongue. The administration of questionnaires investigating the students’ attitudes towards the novel methodology showed promising results; the students were willing to use DDL not only in class, but also for homework, even though learning how to use the corpus and interpreting the results proves to be a demanding and time-consuming task for non-native and non-expert users.

3.8 Conclusion

The present chapter introduced Data-driven Learning, a unique approach to language learning which makes exclusive use of language corpora. We discussed the language learning theories comprising the theoretical framework of DDL, described the various types of corpora employed in DDL, presented its dynamics, potentials and limitations. The present chapter also referred to DDL case studies published over the years, pointing out the lack of adequate studies (Hunston 2002, Johansson 2009), especially on secondary education settings. The creation of secondary education-appropriate DDL material and its application in a Greek Senior High school class will be the focus of the next sessions of our thesis.
“Where should I go?” – Alice

“That depends on where you want to end up”

– The Cheshire Cat.

_L. Carroll, Alice’s Adventures in Wonderland_

CHAPTER 4

MODAL AUXILIARY VERBS

4.1 Introduction

Our decision to focus the DDL lessons on the learning of English Modal Auxiliary verbs is discussed in the present chapter. Firstly, an introduction to the notion of modality and relevant information on the formation, syntactic behavior of modal verbs is offered, alongside with a discussion on the senses of the six modals we chose to include in the present study. In the last section of this chapter, the rationale behind our choice to teach the modal verbs using a DDL methodology is discussed. We aim to present the challenges modal verbs pose for foreign language learners and discuss previous studies which have focused on this area of EFL learning and teaching.

4.2 Modality

Modality is a system, or a ‘semantic category’ (Downing and Locke 2006: 380), by which the speaker introduces various degrees of certainty or possibility into his/ her message (Von Fintel 2006, Palmer 1990) and shows how far they are “committed to the truth of a statement” (Willis 2003: 109). Palmer (1990) states that modality is a semantic term which refers to the meaning of modal verbs and defines it as the grammaticalisation of speaker’s (subjective) attitudes and opinions. In another definition of modality, Huddleston and Pullum (2002: 173) point out that modality is “centrally concerned with the speaker’s individual judgement of the likelihood that something is true, necessary, permitted, etc.”. Lyons (1997) notes that the basic notions
of modality are those of possibility and necessity. Palmer (1990) presents three scales of modality:

a. wish and intention,

b. necessity and obligation, and

c. certainty and possibility.

According to Hoye (1997) the aforementioned notions characterising modality “derive from the fact that human beings often categorize their attitudes and experience in terms of the way things might or must be, or might have been or must have been other than they actually are or were” (p. 40).

Modality can be epistemic and non-epistemic or root modality whose subcategories are deontic, or dynamic modality (Nuyts 2001, Depraetere and Reed 2008, Palmer 1990). Epistemic derives from the Greek for ‘knowledge’ and concerns what is probable, possible or (logical) necessary based on the speaker’s knowledge, and deontic derives from the Greek for ‘duty’ or ‘binding’ or ‘that which is needed’ and involves the issues of obligation, necessity, prohibition or permission, granted on relevant written or moral laws. The third category, dynamic modality concerns ability and volition; it determines whether certain entities involved in the situation are capable to do what is implied in the clause.

However, boundaries between those categories are fuzzy; different communicative uses may create ambiguity and different interpretations of meaning. A common ground to differentiate between the three categories is provided by Nuyts (2001), who states that epistemic, deontic and dynamic modality “differ in how they relate to the participants in the state of affairs” (p133). According to Nuyts (2001) epistemic modality is speaker-oriented, dynamic modality is agent-oriented and deontic modality can be located between the aforementioned two; the moral value is directed at the first-argument participant, but the speaker is the source of this judgment. A similar categorisation is offered by Yule (1998), who discusses epistemic and root modality. According to Yule (1998: 90), epistemic modality means “deductions from speaker/ writer” and root modality means “requirements from speaker/ writer”.
4.3 Categorization of modals

Modality can be realized, or “marked” as termed in Palmer (2001: 7), in various ways, for example by verbs, adverbs, nouns, adjectives and conditionals. Verbs can be divided into *lexical* and *auxiliary*, or ‘helping’ (Quirk et al 1985: 120) verbs. A verb phrase (VP) contains one lexical “and (optionally) up to four auxiliaries” (Van Gelderen 2010: 105). Auxiliaries depend on the lexical verb and add grammatical information in the sentence. This small class of verbs is furtherly divided into *primary auxiliary* like ‘be’ and *modal auxiliaries* or *modals* (Master 1996) like ‘can’. Modal auxiliaries have been termed those verbs which combine with another verb to indicate mood or tense and, therefore, constitute “the principal means” (Depraetere and Reed 2008), the “primary source” (Von Hintel 2006) or “main component” (Nordberg 2010) of expressing modality in English. In the following sentences⁶, the verbs in italics, are primary auxiliaries and the verbs in bold are modal auxiliary verbs.

- They *are* being pulled along with space, as space expands.
- The president *has* warned against taking anything for granted.

---

⁶ All the example sentences provided in this chapter have been adapted from concordances taken from the Corpus of Contemporary American English (COCA).
• That must have been a good luck charm.
• “This might be a permanent career move,” he says.
• We can’t be forgiven if we fail to learn their lesson.

While there are numerous modal verbs in English, only nine are the central (Quirk et al 1985, Romer 2004) or core (Leech et al 2009) modals; those are ‘can’, ‘could’, ‘may’, ‘might’, ‘shall’, ‘should’, ‘will’, ‘would’ and ‘must’. Apart from these, there are also to be found marginal (or peripheral) modals (dare, need, ought to etc.), modal idioms (had better, would rather, be to etc.), semi-auxiliaries (or semi-modals) (have to, be about to, be able to etc.), and catenatives (appear to, happen to, seem to etc.) (Quirk et al 1985: 137), which are however not as frequent as the central modal verbs. In fact, Leech et al (2009) list central modal verbs as being over five times more frequent than semi-modals. It is for this reason that our DDL lessons focused on central or core modals only. The central modal verbs are further categorized into primary modals (‘can’, ‘may’, ‘must’, ‘will’, ‘shall’) and their past tense forms which are called secondary modals and include ‘could’, ‘might’, ‘would’ and ‘should’ (Leach 2004, Palmer 1990).

4.4 Identifying modals: formation, syntactic behaviour and meaning

What makes modals a challenging case in the learning of English grammar is their formation and syntactic behavior which differentiates them from those of the lexical verbs, as well as their semantic indeterminacy, which is the semantic diversity of the meanings carried by modal verbs. Commenting on this, Depraetere and Reed (2008) point out that “it is often difficult to pin down the ‘meaning’ communicated by the modal unequivocally” (p.282).

The present study targets both primary and secondary modals since an investigation of the frequency of those lexical items (see table 4.1) showed that both categories are equally important in English language. Moreover, we believe that attention should be paid to both overused and underused modals. However only six out of the nine core modals were taught in the present study’s DDL lessons, owing to the fact that time allotted for the running of those lessons was limited. The six modals we chose were the following: ‘will’, ‘would’, ‘could’, ‘should’, ‘may’ and ‘might’. The reason behind that
choice is our in-class observations which showed that students seemed to face more difficulties when learning those particular modal verbs.

Table 4.1: Modal auxiliary frequency lists in ECCo and COCA

<table>
<thead>
<tr>
<th></th>
<th>ECCo</th>
<th>COCA Jun2014</th>
<th>COCA Jul2017</th>
</tr>
</thead>
<tbody>
<tr>
<td>Would</td>
<td>271</td>
<td>1,053,721</td>
<td>1,209,440</td>
</tr>
<tr>
<td>Will</td>
<td>237</td>
<td>901,805</td>
<td>979,101</td>
</tr>
<tr>
<td>Could</td>
<td>187</td>
<td>707,998</td>
<td>821,097</td>
</tr>
<tr>
<td>Should</td>
<td>98</td>
<td>354,248</td>
<td>405,415</td>
</tr>
<tr>
<td>Might</td>
<td>69</td>
<td>240,578</td>
<td>278,300</td>
</tr>
<tr>
<td>May</td>
<td>67</td>
<td>363,181</td>
<td>462,805</td>
</tr>
</tbody>
</table>

After the selection of the modals was made, it became imperative to decide what to teach about these lexical items. Our aim was not only to teach the relevant grammar rules; we wanted to offer our students the chance to practice with grammar which reflects authentic language as spoken by native speakers (Conrad 2016). This grammar would include both the grammatical properties of modal auxiliaries, as well as their senses, which will be discussed in detail in the next section.

4.4.1 Structural characteristics

The modal auxiliary verbs possess unique characteristics which differentiate them from all other verbs in English. The grammatical characteristics which distinguish modal verbs are also called the *NICE properties*, NICE being an acronym standing for *negative contraction*, formation of question with *inversion*, *code* and *emphatic affirmation* (Huddleston 1976: 333). More specifically,

a. Negation is formed simply by adding ‘not’ after the verb  
   e.g. *He says he cannot catch any fish.*

b. Questions are formed by inversion of the verb and the subject  
   e.g. *Can he be brave for 20 seconds?*
c. Modals can replace a main verb to avoid repetition
e.g. I can meet you there, if you can.
d. Modals can be used for emphasis
e.g. Not only can he do a rhino, he does half human, half rhino.

There are also other structural characteristics of modal auxiliary verbs, mentioned, among others, in Palmer (2001), Nuyts (2001), Leech and Svartvik (2013), which are summarized as follows:

a. Modals do not co-occur,
b. They have no –s forms for their 3rd person singular,
c. They are normally followed by the bare infinitive, with the exception of ‘ought to’,
d. They have no imperatives, and
e. ‘Must’ has no morphologically past tense form.

4.4.2 Tense and time reference

Modality may be marked by time. Most English textbooks refers to modals as either being present (i.e. ‘can’, ‘shall’, ‘will’, ‘may’) or past tense (i.e. ‘could’, ‘should’, ‘would’, ‘might’), the latter being the past representations of the former. However, Leech (2009) points out that “the present – past tense distinction, if it is considered to exist for modals, is highly irregular and problematic” (p. 80), while Palmer (1990) claims that although morphologically ‘will’, ‘can’, ‘shall’ and ‘may’ have the past tense forms ‘would’, ‘can’, ‘shall’ and ‘might’, semantically they do not behave in a uniform manner. ‘Could’ and ‘would’ have a past sense only at certain times, ‘might’ rarely and ‘should’ never. Palmer (1990) supports the view that only dynamic modality has past tense reference; in epistemic modality, only proposition can be marked for past time, and in the case of deontic, neither modality nor proposition refer to the past. Although ‘might’ and ‘could’ are formally past forms, they are used to express present states of affairs (Downing and Locke 2006). This is what Quirk et al (1985) refer to as “abnormal time reference” (p.128); the past forms of modals can be used to refer to present and future time, and also modals which are not traditionally considered as having a past form, such as ‘must’ and ‘need’, may refer to the future.
Regarding future, there is a close link between modality and future time. Epistemic modality may refer either to present or future, deontic modality always refers to the future with dynamic modality usually having a future reference. Although scholars traditionally agree that English has no future tense, ‘will’ is said to express futurity (Palmer 1990); Lewis ((1986), as cited in Imre 2017) claims that the sense ‘futurity’ refers not to the tense, rather than the ‘logical inevitability’ of the event described.

4.5 Meanings expressed by the central modals

Every modal verb conveys meaning. Polysemy is a problematic issue in the comprehension and production of utterances using modals, a fact mentioned by several scholars such as Nuyts (2001), Ruppenhofer and Rehbein (2016). Also their meaning has both a logical and a practical (or pragmatic) element (Leech 2004: 66). Kennedy (2002: 74) seems to agree with that when he states that the difficulty in mastering modals lies in their polysemy and the similarity of their meanings, which tend to overlap. There is a scaled approach to defining the senses of modals. For example, sentence (a) shows strong possibility (conviction), while sentence (b) expresses possibility and sentence (c) weaker possibility. However, this scaled approach is questionable since the speaker’s conviction may be intensified by adverbs (Hoye 1997), such as ‘well’ (sentence d), which heighten the possibility.

(a) Canada will be producing a billion dollars’ worth of arctic diamonds a year.

(b) Aerobic exercise keeps up blood flow to the brain, and may also benefit brain function through other means.

(c) The answer to the first might be maybe, the answer to the second is most definitely no.

(d) Such claims may well be justified.

Nevertheless, no attempt to teach the modal verbs could be complete without making learners aware of their meanings (or ‘senses’) and uses. Those are closely related to modality and they may adhere to one of the following three categories of modality; epistemic (expressing likelihood or certainty), deontic (which refers to permission or
duty), and dynamic (which expresses someone’s own ability, or willingness). Quirk et al. (1985) divide modality into two types, namely *intrinsic* when there is some kind of human control over the events, and *extrinsic* which involves human judgement of the situation described rather than human control over these events. Intrinsic modality involves such notions as permission, obligation and volition whereas extrinsic modality carries the meaning of possibility, necessity and prediction (see Figure 4.2). Of course, such metalanguage need not appear in the lesson itself, but the uses of modal verbs in everyday language is something students should be aware of. Different descriptions of the meanings expressed by modal verbs have been offered in the past.

It is a fact that the meanings we teach our students depend on the proficiency level of the class. The more advanced the class is the more detailed the nuances in meaning we may teach. Since the target group of the present study is B1+ students, we chose to deal only with the main senses of six modal verbs. The chief meanings of the six modals we chose to focus on in our study were taken from the theoretical framework provided in the work of Klages & Römer (2002), Willis 2003, Leech (2004), Downing and Locke (2006), Depraetere and Reed (2008) and Palmer (1990), and are the ones discussed in the following section and summarised in Table 4.2.

Figure 4.2: The meanings of modals, according to Quirk et al. (1995: 221)
4.5.1 Will

Will is used to convey the meaning of:

- ‘prediction’ or ‘predictability’ as mentioned in Leech (2004). It refers to the future expressing a judgement or prediction of high probability rates (Klages & Römer (2002), or as Palmer (2014) points out “a reasonable conclusion… from previous knowledge” (p.57)
  e.g. On the other hand, this doesn’t mean democracy will come.

- ‘epistemic necessity’ or ‘certainty’. In this case, Downing and Locke (2006: 382) state that ‘will’ expresses “a confident assumption by the speaker as observer, based on experience, known facts or what is usually the case”. On the same wavelength is Palmer’s (1990) claim that will “refers to what is reasonable to expect” (p.57). There is strong epistemic commitment (Downing and Locke 2006).
  e.g. Italy will build Turin/ Venice and Milan/ Naples links.

- ‘willingness’ (‘volition’ in Depraetere & Reed (2008), Romer (2004) and Willis (2003)).
  e.g. Maybe there will be somebody on the jury that will vote to spare his life.

4.5.2 Would

Would carries the semantic notions of

- ‘predictability’ when the orientation frame is past time (Downing and Locke 2006)
  e.g. I asked Roberts what would happen when the troops left.

- Dynamic would in narrative refers to ‘volition’ (Romer 2004 and Depraetere & Reed 2008), or ‘repeated states’ or ‘happenings’ (Downing and Locke 2006)
  e.g. People would constantly be asking the character why he is smoking.

4.5.3 Could

Could is used to express

- ‘past ability’,
e.g. The trees were leafless by then, so a person could have a view of the sidewalk from across the street.

- ‘epistemic possibility’ or ‘general possibility’, or ‘tentative possibility’ (Downing and Locke 2006)
  e.g. It appeared obvious that the adventure could not possibly succeed.

- ‘root possibility’ or ‘hypothetical meaning’
  e.g. If you could have one power taught at Hogwards, what would that be?

- ‘request’ and ‘asking for permission’; what characterises it is the ‘formality’ or ‘state meaning’ (Leech 2004)
  e.g. 5th CALLER: Pardon me? Could you ask that again?

4.5.4 Should

Should conveys the meaning of

- Root necessity in the form of ‘advice’, ‘suggestion’ (or ‘instruction’ in Willis 2003) or non-binding ‘obligation’ (Downing and Locke 2006). The speaker merely states an advisable course of action instead of uttering an obligation.
  e.g. He believes faculty and students should be lifelong learners.

- Epistemic necessity in the form of ‘deduction’ or ‘predictability’ with a medium degree of conviction (Downing and Locke 2006). Here, one states what is reasonable to expect, even though implicitly the use of should admits non-fulfillment.
  e.g. Trains should be chugging through Arizona again today.

4.5.5 May

May expresses the modal meanings of

- Permission
  e.g. Dan may go first.

- Root possibility or ‘probability’
Producers also say they *may* add another out-of-town stop of ‘Annie 2’.

- Epistemic possibility or ‘truth’ (Palmer 1990), or extrinsic possibility denoting weaker conviction (Downing and Locke 2006)
  e.g. This *may* be a policy disaster for the country.

- ‘wish’, which Leech (2004) further divides into ‘benediction’ (a) and ‘malediction’ (b)
  (a) Good luck, and *may* you always be happy.
  (b) *May* you choke from the strings of your own parachutes!

4.5.6 Might
Concerning the senses of *might*, they are very close to those of the modal auxiliary *may*. Palmer claims that *might* conveys the meaning of ‘weak possibility’, or as Downing and Locke (2006) point out, possibility in a more ‘neutral’ form.

- e.g. Radiation[…] *might* also make a section of a city uninhabitable for years.

Table 4.2: Forms and meanings of the six modal verbs discussed in the present study

<table>
<thead>
<tr>
<th>Nonnegative form</th>
<th>Negative form</th>
<th>Meanings</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Uncontracted</td>
<td>Contracted</td>
</tr>
<tr>
<td>Could</td>
<td>Could not</td>
<td>Couldn’t</td>
</tr>
<tr>
<td>May</td>
<td>May not</td>
<td>Mayn’t</td>
</tr>
<tr>
<td>Might</td>
<td>Might not</td>
<td>Mighn’t</td>
</tr>
<tr>
<td>Should</td>
<td>Should not</td>
<td>Shouldn’t</td>
</tr>
<tr>
<td>Will / ’ll</td>
<td>Will not</td>
<td>Won’t</td>
</tr>
</tbody>
</table>
Weak volition/willingness, strong volition/insistence

<table>
<thead>
<tr>
<th>Would / `d</th>
<th>Would not, `d not</th>
<th>Wouldn’t</th>
<th>Volition</th>
</tr>
</thead>
</table>

4.6 The learning of modal verbs by non-native speakers

Our decision to focus our DDL methodology on the teaching / learning of modal auxiliary verbs, was motivated by the fact that MAux count among the most challenging areas of English grammar learning, as even native speakers find it difficult to successfully interpret their possible meanings and use them (Bald 1991, as cited in Klages and Römer 2004). Even more so, modal verbs are a challenging concept for EFL learning. In fact, as past studies have reported, the modal auxiliaries pose a huge challenge for EFL teachers and learners by being “a grammatical problem area” (Römer 2004: 185), “particularly troublesome for learners of English” (Palmer 2003:1).

In a study case investigating the learning of modals by German learners, Klages and Römer (2002) state that “the difficulties EFL learners encounter with respect to modal verbs in English are at least partially due to the learners’ limited access to the cultural values encoded in the descriptive labels used in the traditional paraphrases” (p.201). This, as discussed in the aforementioned study, is due to the sometimes misleading and inaccurate descriptions of lexico-grammatical phenomena in the traditional EFL course books. In order to compensate for this shortcoming, Römer (2004) proposes a corpus-driven approach to language learning and teaching, since the language found in mega corpora, such as the BNC that she used in her research, is the type of English which is used and understood by native speakers nowadays, and that is what the EFL students should learn. Similar is the view expressed by Hinkel (1995) that non-native speakers’ difficulty in grasping the meanings of modal verbs is because they use them in different contexts from those of native speakers. What is suggested in this study is that, since modal verbs reflect cultural values and are therefore culture and context dependent, their learning should be treated as thus.

In a recent study discussing the difficulties in teaching the meanings of modal verbs, Imre (2017) characterizes modals as an “eternal challenge” (p.94) and proposes an
alternative to the “theory-first, practice-later” methodology. This approach could offer the students the opportunity to investigate language through well-chosen samples of authentic language and lead them thus to formulate their own rules. He also suggests that the students practice the newfound knowledge by checking modal verb occurrences in TV series.

Wegener (2006) has investigated the learning of the structure and contextual uses of modal verbs adult learners of English for Specific Purposes (ESP). She addresses the problematics of making German ESP students effectively use language patterns expressing empathy and politeness, notions conveyed through modal verbs. The findings of this study call for ESP materials in the form of realistic texts from course books and authentic texts which should ideally “reflect the different job responsibilities and potential workplace situations where the learners have to speak English” (p.133).

In another study involving ESP learners, Tenuta, Oliveira and Orfanó (2012) discuss that Brazilian learners use a less varied expression of modality and should therefore be exposed to a great repertoire of vocabulary and structures through appropriate material which would raise the students’ awareness of how to use modal verbs effectively.

Qian (2017) investigates the use of modal verbs in the writing of Chinese EFL learners and claims that Chinese learners tend to overuse certain MAux, such as ‘need’, ‘should’ and ‘must’, while they underused others, for example ‘would’ and ‘could’. Qian claims that this was due to the learners’ ignorance of the pragmatic meaning of modal verbs. Similar are the findings in other studies examining the acquisition and use of modal verbs by Chinese learners (Xiao 2017, Yang 2018); Xiao (2017) states that Chinese learners significantly overuse MAux both in speaking and writing, while Yang (2018) claims that the students tend to overuse or underuse modals in academic writing.

Surujiu, I. and A. Scarabnaia (2012) examine the difficulties that Moldovan students of English face when studying the modals, and they claim that the difficulty in learning modals is due to the lack of similar grammatical forms in the native language. In order to account for this shortcoming they propose interactive, cooperative activities.

The learning of modal verbs is also problematic for Greek learners of English, as shown in relevant studies. Kostaki-Psoma (2015) stresses the complexity of modality in language use, a fact also mentioned in Manika (2009) who claims that modal verbs are “one of the most difficult aspects of learning English” (p.308). Kostaki- Psoma (2015)
investigated the learning of ‘can’ and ‘could’ by Greek learners of English. She states that the multiple meanings and functions of modals may confuse the learners. She discusses how the Phonetic intricacies of Greek language may cause mispronunciation by Greek students. She also claims that ‘can’ and ‘could’ have no one-to-one equivalents in Greek and may therefore confuse Greek learners of English. That is why she suggests the adoption of a more communicative approach, in which there is ample exposure of learners to language, presentation of forms in context and contextualized meaningful practice.

Manika (2009) also investigates the learning and use of modal verbs by Greek EFL in State Schools. What her research revealed is that Greek learners face difficulties in making correct associations between modal verbs and their corresponding notions in authentic communication settings. The reason she provides for this mismatch is their tendency to translate English modals with their Greek equivalents. What she suggests in order to overcome this difficulty is to describe, and thus teach, modality not as distinct lexicogrammatical elements, rather than as interrelated parts of specific speech acts. The pragmatics of modal verbs use should not therefore be neglected and the correct ‘form-meaning connections’ (Manika 2009: 309) should be made if the aim to facilitate the learners into learning the appropriate sense and use for each modal is to be achieved.

4.7 Conclusion

For the above-mentioned reasons, it becomes evident that the difficulty to learn the Modal verbs derives from the fact that one’s knowledge of them and one’s ability to use them appropriately is not limited to their grammatical function and senses but to their pragmatic element as well. Their teaching and learning is initially lexical, and learners should be pointed as to see “how the meanings are exploited with subtle differences in discourse” (Willis 2003: 110) if they want to communicate successfully overcoming any cultural differences between their mother tongue and the target language. Concordances have been claimed to be useful for expanding knowledge about language, for example grammatical structures, that the learner is already familiar with (Hunston 2002, Braun 2010). The latter is true in the educational reality in Greece, since it is a fact that all students in Greece are taught the modal verbs in lower secondary
education; therefore, our approach ideally offered new insight to linguistic knowledge that was acquired in the past.

We believe that our own intuitions or the examples given in the available textbooks can only account for a limited range of the different instances of modal auxiliary verbs as they are used by native English speakers in different settings. DDL becomes thus the ideal methodology for the learning of modal auxiliary verbs by providing an abundance of natural occurring examples of everyday language. DDL material mirrors actual language use and is most efficient in assisting learners to engage in cross-cultural communicative situations.
“Science progresses by means of ‘essential tensions' expressed at the boundaries of tradition and revolution.”

(Kuhn 1962, in Minarik 2008: 38)

CHAPTER 5

THE METHODOLOGY

5.1 Introduction

In this chapter, section 5.2 presents the research hypotheses and the research questions that our study attempts to answer. Section 5.3 provides background information on the educational system of Greece and in particular the upper-secondary education that is of more importance for the purposes of our research, since our target group is senior high school students. Section 5.4 gives a detailed description of the pilot study of our research; it presents the target group, the structure and the findings of the pilot study which preceded the main study of the present research. The next sections discuss the main study of our research. We present the target group that is the Senior High school students who participated in this case study and we provide information on the profile of those students, as well as their attitudes towards foreign language learning and the use of ICT for (foreign language) or Computer Assisted Language Learning (henceforth called CALL. Moreover, the tools which were created specifically for use in our research are presented in detail. These include a needs analysis questionnaire, a questionnaire investigating the students’ attitudes towards ICT, the placement tests, the pre, immediate post and delayed post experimental tests and the creation of a pedagogic corpus of English Course books which are approved by the Greek Ministry of Education, Research and Religious Affairs for use in Greek Upper Secondary education. The last section discusses the use of a free online electronic concordance program, the Corpus of Contemporary American English (henceforth COCA).
5.2 Research methodology and research hypotheses and questions

Our methodology is in accordance with the three ‘maxims’ which should govern the effective use of DDL:

- don’t scare the students;
- use the corpus when the dictionary does not tell you enough,
- disguise the corpus as a dictionary

(adapted from Kilgariff et al. 2015: 61)

Upon the evident lack of empirical studies on the benefits of DDL for EFL teaching in Greek secondary schools, the objectives of our research were the following:

c) To investigate if learners of English in Greek high schools can extract benefits from a DDL approach;
d) To examine how learners of English in Greek high schools can benefit from such an approach to language teaching

More specifically, our study attempts to give an answer to the following research questions:

1. Can a DDL teaching methodology be applied – either directly or indirectly - in a Greek high school classroom setting?
2. How can methodology and materials be appropriately adapted so as to have a data-driven approach to EFL teaching?
3. Is DDL more effective than traditional grammar teaching in learning the English modal verbs?

More specifically:

a) What are the actual benefits of DDL teaching methodology for learners of English in Greek high schools?
b) In the case of the present study, was the DDL methodology more beneficial than the traditional grammar teaching methodology?
c) Is any improvement in the competence of the learners who have been taught the Modal Verbs following a DDL methodology statistically significant?
d) Are the gains of a DDL methodology retained eight months later?
5.3 The Greek educational system

The educational system in Greece, as stated in the official site of the Greek Ministry of Foreign Affairs, is divided into three levels, namely primary, secondary (including lower secondary education, and upper secondary education), and tertiary. One could also add a fourth level, pre-school education for children aged 4 to 6. Education is compulsory in primary and lower secondary levels, that is for students between the ages of 4 and 15. In primary and lower secondary education the Greek Educational System dictates a common core curriculum provision (as described in the EU’s Eurydice report “The Structure of the European Education Systems 2016/17”: 6).

Figure 5.1: The Greek Education System according to the EU’s Eurydice Report 2016/2017

5.3.1 The upper-secondary education in Greece

When students graduate from lower secondary education, they have the choice to either leave school and look for a job or attend courses in upper secondary education. While upper secondary education is not compulsory in Greece, most students attend one type of ‘Lykeio’ offered, such as the General High school, or the Vocational High School or high schools specializing in Music or Arts. As stated in the EU’s 2016 Eurydice report, programmes at this level are typically designed to complete secondary education in preparation for tertiary education or provide skills relevant to employment, or both (EACEA/ Eurydice 2016: 10). Pupils enter this level typically at the age of 15.
Upper secondary schools offer a combination of General Education courses and Advanced Placement courses. According to the description provided in the Fullbright Institute site, students who wish to pursue studies in Higher Education by sitting the Panhellenic exams opt for specific Placement Courses which fall into the following three categories or ‘orientations’ 1) Humanities, 2) Science and 3) Economics and Computer Science. This is considered to be a very demanding procedure and students spend considerable time and effort to ensure a place in tertiary education.

5.3.2 Strengths and weaknesses of the Greek educational system

Statistical data provided in the ‘Education and training monitor report for Greece - 2015’ issued by the Directorate General of Education and Culture of the European Commission reveal that Greece rates below average in most key indicators, such as for example underachievement in basic skills, employment rates of recent graduates and adult participation in learning. Moreover, Greece is one of the countries with the lowest levels of digital skills in the EU, as it sadly ranks 25th out of 28 Member States in terms of internet usage, and in 2012 as many as 65% of Greeks had few or no digital skills, compared to the EU average of 47%.

Last but not least, according to the 2009 PISA database, as stated in the OECD indicators site, individual schools in Greece were among the least autonomous with respect to resource allocation, assessment and curriculum, such as choosing textbooks and determining course content. Indeed, the European Commission’s ‘Education and training monitor report for Greece - 2015’ states that “Greece has a highly centralized school education system…” and that “Greece is the only country where the relevant education authorities take decisions nearly on all matters relating to teaching, including the choice of teaching methods”.

The initiative “the digital school” by the Greek Ministry of Education, Research and Religious Affairs was created to account for the aforementioned two deficiencies of the Greek educational system. It introduces the use of internet technologies and use of IT in classroom. However, Greece sadly rates very low in the use of ICT in class. This is stated in the report by Dendrinos, Zouganeli and Karavas (2013) and it is partly
explained by the lack of appropriate infrastructure in Greek schools and by the lack of training of Greek teachers as regards the use of ICT in class.

5.3.3 Foreign language learning in Greece

A fairly recent Eurobarometer report, requested by the European Commission in 2012, states that 88% of Europeans think that knowing languages other than their mother tongue is very useful. The fact that an even higher percentage (98%) declares that mastering foreign languages is useful for the future of their children underlines the belief that foreign languages are considered essential in order to work, move and communicate in the multicultural European Union. In Greece, like the majority of European countries, “English is by far the most taught foreign language, with the percentage of students learning English in general upper secondary education exceeding 90%” (Eurydice Report 2012: 3).

Greece rates above average⁷ among the EU state members in a recent Eurostat survey concerning the proportion of pupils learning foreign languages in primary education. The teachers are all university graduates who have completed four-year undergraduate studies with a focus in linguistics, British and American literature and culture and translation. Moreover, many of them have pursued postgraduate studies either in Greece or abroad. They are therefore well motivated and highly qualified.

As mentioned in the previous paragraph, one of the priorities set by the average Greek family is the foreign language education of its younger members. This is not the case only in Greece but in other European Countries as well, as revealed by the data contained in the European Survey on Language Competences (ESLC). In particular, “students [in the 15 of the participating countries or regions covered by the ESLC] consider it useful to learn English for their future education, work or for getting a good job” (Eurydice Report 2012: 5).

In the case of Greece, students’ motivation to learn foreign languages may be attributed to a number of reasons; one reason might be the not-so-widely used native Greek language, another being Greece’s financial reliance on tourism and the growing number of Greeks working in the tourism industry. A third reason could also be the fact that

---

⁷ Occupying the 11th place among the 28 EU member states
Greeks have always been extrovert; since antiquity they’ve traveled abroad, first as merchants and warriors, later as colonists, immigrants, students and travelers. Therefore, foreign language learning is considered to be a sine qua non for all Greek students, regardless of gender, family background, educational preferences and future goals.

As discussed above, the Greek educational system is highly centralized and decision making regarding foreign language teaching in all grades of education, including upper secondary lies exclusively in the Ministry of Education, Research and Religious Affairs. All guidelines are to be found in the National Curriculum available to all state school teachers. Curriculum, according to Rogers (1989: 26 cited in Richards 2001) is a much broader concept, as in essence is “all those activities in which children engage under the auspices of the school”, which includes “not only what pupils learn, but how they learn it, how teachers help them learn, using what supporting materials, styles and methods of assessment, and in what kind of facilities”. The National Curriculum for the teaching of English in senior high school contains a detailed description of those skills that are deemed necessary for upper secondary education students.

The ESLC (European Survey on Language Competences), a EU-funded report on foreign language learning in Greece, which was conducted by the University of Athens (E.K.P.A.) and the National Institute of Educational Policy (I.E.P.), gathered useful data on the foreign language proficiency of secondary school students in Greece, its ultimate goal being to contribute towards the improvement of foreign language learning in European schools.

According to the ESLC report, a new curriculum for foreign language learning in Greek schools, the Integrated Foreign Languages Curriculum (IFLC), which was developed in 2011, adopts a generic approach to language learning as dictated by the popular Communicative Approach. According to IFLC specifications, attention in foreign language learning shifts from linguistic to communicative competence and conforms to the 6-level scale comprising the European standard for language proficiency specified by the Council of Europe. However, IFLC only provided vague specifications for primary education. Despite the main tendency in EU to use the European Framework of Reference (CEFR) as a tool for defining student attainment levels, the Greek
Educational System has not yet made any provision of regulations or recommendations for the use of CEFR to define minimum attainment levels (Eurydice Report 2012: 6).

Moreover, as the ESLC report explicitly mentions, the IFLC’s ‘communicative’ orientation to language learning sadly fails to integrate ICT into foreign language teaching due to lack of equipment, software and teachers’ reluctance owing mainly to lack of suitable teacher training. In other words, in Greece there have been few opportunities to successfully incorporate ICT in foreign language learning. This, as mentioned in the ESCL report, can also be regarded as “(a failure) of the school to take advantage of students’ existing digital literacy and extend it into the development of strategies for learning”. Nevertheless, recent studies have shown educational technology and course and syllabus design to be considered of foremost importance by in-service EFL teachers in Greece. Evidence from a 2011 study by N. Sifakis shows that “more than 64% (of the in-service teachers that participated in the survey) identify internet as their top or top-by-one priority, whereas 55% believe to be in need of further training in the implementation of appropriate educational computer software” (p. 395). In the same study, the second highest priority among those questioned is the course and syllabus design, as 75% of the teachers “wanted to find out more about creating their own teaching materials” (Sifakis 2011:395)

Our research provides the methodology and material that could be used for a more effective use of ICT in EFL classroom, giving the teachers insights on how to ‘modernize’ teaching and allow students to make good use of their ICT skills.

5.3.3.1 The new curriculum for compulsory education

The Greek school curriculum includes English as a compulsory foreign language in primary school (from the first grade onwards), through the end of junior secondary school. As stated in the previous paragraph, a new curriculum for foreign language teaching in schools, the Integrated Foreign Languages Curriculum, was developed in 2011 within the framework of the new National Curriculum and was published in the Official newspaper of the Greek Government in 2016. It provides a common framework for the teaching of all offered languages in primary and lower secondary schools, including the objectives, the didactic materials and assessment.
One of its major characteristics is that it describes the desired linguistic competences, functional, grammatical, lexical and pragmatic, per level rather than per grade. In particular, the new curriculum for foreign language teaching aims to assess and certify the students’ language competences on the six-level language proficiency scale of the Common European Framework of Reference for Languages (CEFR), an international standard for describing language ability. The structural elements necessary to reach each CEFR level are listed in detail and are in agreement with the indicators of communication competency.

Figure 5.2: The Common European Framework of Reference descriptors

<table>
<thead>
<tr>
<th>Level</th>
<th>Listening</th>
<th>Reading</th>
<th>Speaking</th>
<th>Writing</th>
</tr>
</thead>
<tbody>
<tr>
<td>C2</td>
<td>Has no difficulty in understanding any kind of spoken language, delivered at fast native speed</td>
<td>Can understand a wide range of long and complex texts</td>
<td>Has a good command of idiomatic expressions</td>
<td>Can write clear, smoothly flowing, complex texts in a logical structure</td>
</tr>
<tr>
<td>C1</td>
<td>Can understand enough to follow complex topics, though he/she may need to confirm details</td>
<td>Can understand in detail lengthy, complex texts, provided he/she can read difficult sections</td>
<td>Can express him/herself fluently and spontaneously</td>
<td>Can express him/herself with clarity and precision</td>
</tr>
<tr>
<td>B2</td>
<td>Idiomatic usage influences the ability to understand</td>
<td>Has a broad active reading vocabulary, but may experience some difficulty with low-frequency idioms</td>
<td>Can interact with a degree of fluency and spontaneity that makes regular interaction</td>
<td>Can express new and views effectively in writing</td>
</tr>
<tr>
<td>B1</td>
<td>Can understand the main points of clear standard speech on familiar matters regularly encountered in work, school, leisure</td>
<td>Can read straightforward factual texts on subjects related to his/her field and interest</td>
<td>Can exploit a wide range of simple language to deal with most situations</td>
<td>Can write personal letters and notes asking for or conveying simple information of immediate relevance,</td>
</tr>
<tr>
<td>A2</td>
<td>Can understand enough provided speech is clearly and slowly articulated</td>
<td>Can understand short, simple texts containing the highest frequency vocabulary</td>
<td>Can communicate in simple and routine tasks requiring a simple and direct exchange of information</td>
<td>Can write short, simple formulaic notes relating to matters in areas of immediate need</td>
</tr>
<tr>
<td>A1</td>
<td>Can follow speech which is very slow and carefully articulated, with long pauses for him/her to assimilate meaning</td>
<td>Can understand very short, simple texts a single phrase at a time, picking up familiar names, words and basic phrases</td>
<td>Can interact in a simple way but communication is totally dependent on repetition at a slower rate of speech</td>
<td>Can ask for or pass on personal details in written form</td>
</tr>
</tbody>
</table>

For example, CEFR indicators dictate that the B2 level learner should “…converse naturally, fluently and effectively” and “…sustain relationships with native speakers without unintentionally amusing or irritating them or requiring them to behave other than they would with a native speaker”.

---

8 The CEFR descriptors are presented in detail in https://www.coe.int/en/web/common-european-framework-reference-languages/table-1-cefr-3.3-common-reference-levels-global-scale
The aforementioned language competence descriptors have served as the basis for calibration of the contents of foreign language course books published by the Ministry of Education for use in Greek primary and lower secondary schools.

The main principles underpinning the new curriculum for foreign languages are those of effective communication and understanding of the world. It is based on the belief that the role of foreign language education at school is to provide the learners with practical knowledge in order to successfully communicate in modern multicultural societies, but it may also enhance cultural awareness. In other words, it may provide the students with quality education and not only skills.

The new integrated framework for the teaching of foreign languages also sets the basic principles for its effective application. More specifically, it provides guidelines concerning the optimum distribution of students in classes, according to their level of knowledge rather than their age or alphabetical listing. Moreover, it makes a special reference to the choice of appropriate teaching materials with particular reference to the active involvement of the teacher in the creation of appropriate teaching and supporting material which will take into consideration the individual characteristics of the students. Lastly, it stresses the importance of employing ICT tools in foreign language classroom, for effective education that meets the needs of modern learners. This is to be achieved with the necessary technological equipment at schools as well as the appropriate teacher training.

The latest addition to the guidelines concerning the teaching of foreign languages in compulsory education, is the new curriculum for the teaching of English in the 1st and 2nd grade of primary school, introduced in 2016, which focuses on foreign language teaching to very young learners.

5.3.4 Foreign language learning in Upper-secondary education in Greece

English has been until the school year 2017-18 a compulsory-optional subject in all three years of general high school, with two 45-minute lessons per week. The classes

---

9 Students have to choose one from the languages offered; English, German or French. In 3rd grade they may also choose a second foreign language.
are mixed ability, since the students are placed in classes with alphabetical listing, and they have 18-27 students.

An ongoing reform of the education provided in senior high schools is about to change the status quo of foreign language teaching in upper secondary education. From the school year 2018-19 onwards, foreign language classes will be optional in the second and third grade of senior high school. This may have a negative impact on the mastery of foreign languages among Greek students in upper-secondary education, which today is quite impressive; in 2015, 94% of upper secondary school students were taught at least one foreign language (Eurostat [2] 2017).

The situation will be probably worse in the case of the, already inadequate, second foreign language teaching in Greek senior high schools. Enabling citizens to communicate in two languages other than their mother tongue, also known as the ‘Barcelona objective’, is an ambitious goal, agreed in 2002 by the EU member states. However, data provided by the Eurostat service shows that in 2014 only 2.4% of upper secondary education students in Greece were learning two or more languages. In the respective ranking list of all 28 EU countries, Greece occupied the last place.

5.3.4.1 The objectives set by the curriculum for upper-secondary education

Contrary to compulsory education, the curriculum concerning the teaching of foreign languages in general upper secondary education has remained unchanged since it was published in the official paper of the Greek Government in 1999.

As stated in it, the development of the curriculum for the teaching of foreign languages in upper secondary education has taken into consideration not only the communicative needs of the learners, but also the characteristics of teenagers’ personality and the modern learning theories.

Rather than focusing on vocabulary and grammar competence, the curriculum lists various instances of communication that a learner should be able to cope with. It describes the communicative needs of upper secondary education students and sets general and side objectives. In it there is a separate paragraph describing the communicative goal of expressing beliefs, ideas, wishes and emotions; it includes language functions, language exponents and examples. In the case of modal verb use,
which is of central importance for the purposes of the present study, it dictates that students become familiar with making predictions, expressing certainty, necessity, obligation, ability or lack of it, expressing judgment, making requests and suggestions, providing advice, giving orders and asking for and giving permission or forbidding someone to do something.

The curriculum also sets the guidelines for the design and application of effective learning activities. It states that any activities should expose the student to realistic communication instances aiming to exchange information, provide solution to problems, and reach common decisions and actions. To achieve those goals, the learning activities should match the learner's level of knowledge, give him/her the opportunity for self-study, and self-assessment and also provide the chance to work in pairs or groups.

Since the present experimental study takes place in a state school, it is imperative that both control and experimental groups be taught targeting the aforementioned communicative goals; also, any suggested learning methodology should take into consideration the suggestions of the curriculum regarding the creation of effective learning activities.

5.4 The pilot study

The main part of this empirical research study was conducted in the 2nd Senior High school (2ο Geniko Lykeio) of Orestiada, a small town of around 20,000 residents at the north-eastern borders of Greece. The researcher is a full-time EFL teacher at this school and has classes with all 196 students. Due to this, the students were very well acquainted with the practitioner and they felt thus free to make comments on the learning process and ask for assistance when they needed it.

Prior to the main experimental study, a pilot study of a smaller range was conducted. This pilot study gave the author of this study the chance to try different DDL activities with a small group of students and examine their attitudes towards this methodology. It was carried out in three 45-minute lessons and provided the basis for the design of the three training sessions on DDL which preceded the actual DDL-lessons of modal verbs.
5.4.1 The pilot study participants

During February – May 2013 three DDL-designed learning sessions were administered to 22 2nd grade students, who were 16-17 years old. The rationale behind the choice of the students belonging to that class was that we intended to run the main DDL lesson with 2nd grade students as well; therefore, we decided it was appropriate for the pilot sessions to be conducted with students of the same level. All students belonged to the same class, and the group included in total 13 girls and 9 boys. This class was chosen, firstly because, after the administration of the placement test, they proved to be level B1+ and secondly, because their classes timetable was suitable for the purposes of our research. More specifically, the school has a computer laboratory with 15 computers in total, which are used by the class teachers; therefore, we had to ensure that at the time of our DDL sessions no other teacher would need to make use of them.

5.4.2 The pilot study DDL training sessions

Session 1

Pilot Session 1 involved training on handling concordances taken from ECCo with the use of Wordsmith Tools 6 and the Corpus of Contemporary American English (COCA).

Initially there was discussion in class on how we deal with an unknown word or a case when we cannot decide on which item to choose between two near synonymous words (e.g. ‘deny/refuse’). The second phase of Session 1 involved lab work. All students worked on a pc with internet access, and they all worked in pairs. In a previous lesson during which the students had read a NY Times article on bullying, the majority of them said that the most ‘difficult’ phrase to identify in the article is the phrase ‘swirl of innuendo’. Therefore, the word ‘innuendo’ was written on the board and they were asked to find the meaning of the word based on the procedure which involved googling, examining monolingual and bilingual dictionaries and investigating an online Corpus interface. This procedure is outlined in figure 5.3; students worked autonomously in the
first 3 activities, however they were guided by the teacher in the stage 4 of the pilot session 1, since they had no previous knowledge of corpora. Therefore, stage four was preceded by a brief description of corpora and their uses.

Figure 5.3: Pilot Session 1 outline

1. ‘Google’ the word and guess its meaning by the ‘hits’ they had
2. Search an online English-English dictionary (eg word-reference.com)
3. Search an online English-Greek dictionary
4. Enter the website of COCA. Type the keyword ‘innuendo’ and go through the results so as to realize that the output lines are examples of the usage of the word ‘innuendo’

Session 2

This training session focused exclusively on the use of COCA platform. The aim was for the students to familiarize themselves with raw concordance output. This session built on the knowledge gained in the previous pilot session. The phrase ‘swirl of innuendo’ was written on the board. By now, the students knew the meaning of ‘innuendo’ but not of ‘swirl’. Instead of asking them to consult a dictionary, or google the word, the students tried a pure DDL approach with a direct consultation of an online corpus. More specifically, they were asked to enter the COCA website and fill in a handout with information they found by searching the corpus. The handout outlined the steps needed so as to successfully identify the word ‘swirl’ (Figure 5.4). This procedure was slightly more demanding that the one in pilot session 1, as the corpus search was more advance; the students had to choose sub corpora and conduct a KWIC search. The teacher guided them and answered every question they had.
In the 3rd session use was made of the language data contained in the handmade pedagogic corpus ECCo. This session involved use of printed handouts which were distributed to each student. It contained 15 edited concordances lines, full sentences only, which included instances of Past Simple and Past Progressive. Each student worked on his/ her own handout, however they were encouraged to interact, compare their answers and share their findings with the class. The aim of session 3 was for the students to look at the sentences, search for repetitive patterns and form generalizations about the formation and usage of those two tenses.
Figure 5.5: Pilot Session 3 outline

**Let's discover language!**

- Have a look at the following sentences. All of them are sentences taken from coursebooks.

a) **Do you have any unknown words?**
b) **Do you see any repetitive patterns?**
c) **Can you guess what today's lesson is going to be about?**
d) **Can you form generalizations and make your own rules out of the examples given?**

1. I realized I'd left them on the kitchen table as soon as I closed the front door. I'm so angry with myself. It's such a stupid thing to do!
2. Backing silently towards the bedroom door, she kept her eyes on the window. She expected a horrible alien face to appear there at any second. Suddenly she felt something behind her..
3. He was accepted and finally graduated with honours.
4. Well, Jennifer, one of my classmates, was wrongly accused of stealing another girl's mobile phone.
5. I closed my eyes and opened them again. Several Japanese tourists were admiring the cow. Some were taking photos, others were laughing!
6. Then something caught her eye - one of the dummies was moving! The dummies were coming to life and they started to walk slowly towards her.
7. While they were drinking their coffee, she showed them a leaflet.
8. During the night of 19 September, 1961, an American couple called Betty and Barney Gill were driving home to Portsmouth, New Hampshire from a holiday.
9. Rose panicked and tried to run away, but they were getting nearer. She screamed in horror.
10. E: Why did you hit my bike? Couldn't you see where you were going? F: Well, no, I couldn't because the sun was in my eyes.
11. Father: The theatre was full! I think the audience really enjoyed it - they were laughing a lot! Teen Boy: I'm just glad you liked it!
12. Emily's heart was racing and her hands were shaking.
13. She shouted at the headteacher and said that he was being unfair. She said some other things she shouldn't have said, too.
14. Liz: The acting was OK, the script wasn't bad. I can't say that it was boring, but I wasn't thrilled about it either.
15. James: Wasn't there anything you liked about it? Rick: There were lots of shots of spaceships and battles in space.
5.4.3 The contribution of the pilot study

The pilot study helped us to refine and finalize the content and layout of the DDL lessons. Both printed handouts with edited and unedited concordances were provided, as well as handouts with step-by-step instructions for a hands-on DDL procedure. Students’ comments while they worked on the DDL activities, as well as informal interviews that took place after the end of those DDL sessions provided valuable insight on how we could create a DDL lesson that would be easy to administer, comprehend and handle, with clear instructions and ‘user-friendly’ computer-based activities. In particular, the vast majority of the participants (75%) said that they would prefer to do more DDL activities, while 80% stated that corpus work would help them avoid certain errors in the future and almost all participants decided that they prefer the hands-on instruction more than the paper-based one. Lastly, when they were asked to comment on the DDL approach as compared to the traditional one they were so familiar with, they said it was ‘novel, ‘interesting’, ‘useful’ but they would prefer to be given more feedback by the teacher.

5.5 The main study

5.5.1 The main study participants

After the outcomes of the pilot study were assessed, the DDL activities were slightly modified and refined so as to be more attractive for the ‘DDL-naïve’ target group. The pilot study target group was considered inappropriate to be also the target group for the main part of our study for two reasons. Firstly, during the academic year 2014-2015, those 22 students were 3rd grade, or ‘final year’, students, who tend not to attend classes regularly; it is a sad fact that seniors in Greek high schools often skip classes in order to study for the end-of-year Panhellenic examinations. Secondly, their number was considered insufficient.

Therefore, we chose instead to involve in our study all 2nd grade students, who were also level B1+; this was found after taking into consideration the results of the Quick placement test they had taken at the beginning of the school year. For this reason, they could be taught using the B1+ level pedagogic corpus we had created. The participants
were all Greeks, their native language being Greek and four of them were bilinguals, with Bulgarian being their second language. In total 64 students, 35 girls and 29 boys participated in the study; they were all 2nd grade students who belonged to three classes. 24 students came from class B1, 20 students from class B2 and 20 students participated from class B3. Three students from class B2 and four from B3 respectively did not have sufficient participation in all the DDL sessions, as they were often absent from classes. For those students who were absent when pre and post experimental tests were taken, the opportunity was given to sit for them in a later lesson. All classes consisted of mixed English language proficiency students, as shown by the results of the Oxford Quick placement test (U.C.L.E.S. 2001) they sat prior to the DDL sessions.

5.5.1.1 The Experimental and the Control Group

As the present study is an experimental one, it was deemed necessary to involve two groups of students; one was the experimental group and the other the control. Classes B1 and B2 formed the experimental group and class B3 the control. The reason behind this had to do mainly with the school timetable. Class B3 could inconveniently spend only a few hours in a computer lab since when this class had EFL lessons, the school computer lab was usually occupied by other classes having ICT lessons.

The experimental group was exposed to a DDL methodology for the learning of modal verbs, using handouts of printed concordances as well as an electronic concordance program. The handouts contained printed concordances based on material taken from the pedagogic corpus ECCo (English Course-books Corpus) designed and created for the purposes of our research. The electronic concordance program used for the hands-on DDL methodology was the one provided by the Corpus of American English (COCA). The control group was taught using the assigned school course-book which, during the academic year 2014-15, was ‘Take off’ CEFR level B1+ by Hillside Press. The choice of the book was made based on the following rationale; it was included in the list issued by the Ministry of Education, it was level B1+ and therefore we could have comparable results with those of the experimental groups, and last, but certainly not least, it was one of the cheapest course books in the list. Here, we have to note that although the Ministry commissions and prints course books for all subjects taught in
both primary and secondary school, course books for foreign languages in general senior high school are not provided; the teacher is encouraged to choose one from the Ministry-assigned list of books by well-known publishers and the students pay for them themselves. Therefore, the cost of the chosen course book was a powerful incentive behind our choice.

5.6 The instruments and materials

The implementation of the proposed DDL methodology could not have taken place without the use of carefully designed materials which are presented in detail in the next sections. The materials created specifically for use in the present study include a) a needs analysis questionnaire, b) a questionnaire investigating their attitudes towards Computer Assisted Language Learning, c) pre and post experimental tests, and d) a pedagogic corpus. Other ready-made materials which were used in our research involved the Oxford Quick Placement test and a free online corpus. All DDL-based lessons were designed with a focus on English modal verbs. Those will be discussed in detail in chapter 6.

5.6.1 The needs analysis questionnaire

Before the compilation of the corpus a questionnaire was distributed to all students, in September 2014. The aim of the questionnaire was to investigate students’ preferences and opinions as far as grammar and vocabulary learning is concerned and thus to provide the necessary evidence so as to check our initial hypotheses that a) grammar is especially challenging for Greek learners of English and b) students, albeit subconsciously, opt for strategies which are close to the DDL methodology when facing a lexical or grammatical structure they are not familiar with.

Both quantitative and qualitative data were provided by the questionnaire, which included 10 closed and 1 open-ended questions. The 10 closed questions rated the respondents’ ‘level of agreement’ on a 5-point Likert scale from 1 (strongly disagree) to 5 (strongly agree). The Likert-type rating scales are commonly used to measure
“student views on various educational interventions” (Jamieson 2004: 2017), one of their assets being their ability to “extract more accurate feedback of users” (Quan 2016: 279).

The closed-type questions investigated students’ attitudes towards grammar and vocabulary learning. Among other questions, they were asked whether they enjoy learning grammatical structures, what is the best way for them to learn grammatical structures and vocabulary in a foreign language and what is their opinion regarding the importance of learning grammatical structures. The open-ended question investigated how they deal with an unknown lexical term (see Appendix B.2).

The data gathered by the responses the students offered in this questionnaire was analyzed using the Microsoft Office Excel program. The results confirmed our initial hypothesis; the students regard grammar learning as especially challenging. They admit the importance of knowing grammatical rules, but they do not like learning grammar. As shown in the questionnaire findings, a significant 76% ‘strongly agree’ or ‘agree’ with the statement “I must know grammar in order to speak/write in English” and an equal 76% ‘strongly disagree’ or ‘disagree’ with the statement “I like to study English grammar”.

Students’ responses to other questions highlighted the main principles underlying the DDL approach. One of these is the preference for a learning methodology that makes use of texts, since 58% of the students ‘strongly agree’ or ‘agree’ with the statement “The best way for me to learn is to read texts”. Another is the respondents’ preference for a learning methodology which does not rely solely on rules. More specifically, 48% of the students state they ‘strongly disagree’ or ‘disagree’ with the statement “I learn best by studying rules”, while 58% admit that they “must know grammar rules in order to speak/write in English”.

Moreover, the students’ responses underlined the importance of learning word chunks instead of individual words and the connection between grammar and vocabulary. More specifically, they stressed the importance of word chunks when offering their opinion on the statement “I must know many word chunks in order to speak/write in English”; here 71% of the students said that they either ‘strongly agree’ or ‘agree’, whereas an equally significant percentage of 70% said that they either ‘strongly agree’ or ‘agree’ with the statement “I must know many individual words in order to speak/write in
Furthermore 48% of the students state that they either ‘strongly disagree’ or ‘disagree’ with the fact that grammar and vocabulary are two completely separate parts of English.

Last, the students’ answers in the questionnaire’s one and only open-ended question “When I come across an unknown word I …..” provided another interesting finding; the majority of the students filled in the open-ended statement with “look in a dictionary” (64%), with the second and third most popular choices being “search the internet” (38%) and “guess by the co-text” (29%).

5.6.2 The ‘Attitudes towards Computer Assisted Language Learning’ questionnaire

Before conducting the DDL training sessions as well as the actual DDL-based lessons on modal verbs, we created a questionnaire\textsuperscript{10} investigating the students’ attitudes towards using a personal computer and web-based resources to assist their learning process and asked both the experimental and control group to respond to its questions. The aim of this questionnaire was to investigate if and how students use a computer in their daily life for communication, entertainment, and most important for learning purposes; the answers would provide support and justification for the present study’s choice to use computers and web resources to teach a foreign language.

The questionnaire contained 6 multiple-choice and 4 yes/no questions and investigated the following: (a) in what ways and how often the students use a pc for (language) learning purposes and (b) whether they can see any advantages or difficulties in using a pc for (language) learning purposes.

After the data was collected, its analysis, using the Microsoft Office Excel program, revealed some interesting facts. Firstly, all students are computer literate; that can be attributed to the fact that IT lessons are compulsory in all Greek high schools. Secondly, only a very small percentage (3%) of the students use their pc for learning purposes; the majority opt for social media (48%) and online gaming (30%). Most of the

\textsuperscript{10} See Appendix B.1
respondents said that they use a PC for learning purposes once (21%) or twice (37%) a week, with the time spent on such activities being one hour per week (58%).

One interesting finding of this questionnaire was that despite the students’ low use of computers for learning purposes, in their answers they acknowledge its usefulness. 90% of the students state that Computer Assisted Language Learning could be advantageous; its main advantages being convenience (47%) and the ability to provide immediate answers (31%). The vast majority of the students claimed that it could be even more advantageous than traditional language learning methods and that it could raise their interest in learning a foreign language.

Those findings provided us with useful insights as to whether any computer-based language learning methods would attract the students’ interest. What can anyone conclude by seeing the answers in this questionnaire, is that computers are indispensable parts of the students’ everyday life, and also that, even though students are not familiar with e-learning, they have a positive attitude towards it, they acknowledge its merits, and they would not be afraid to give it a try.

5.6.3 The Quick placement test

One year before the actual DDL learning sessions took place, all 64 students who participated in this empirical study took the paper and pen version of the Quick Placement Test (UCLES 2001). The administration of this test takes place during the first week of lessons every September and all first graders sit for it. The reason is that it provides an accurate and reliable testimony of the learners’ level of proficiency so as to choose the appropriate course books.

All students who enter Senior High school in Greece have had 7 or 9 years of English classes if they have attended primary schools following the revised curriculum (EPEAEP schools). Moreover, as mentioned in chapter 4, the majority of them have had English classes either in a private evening school – a form of language education that thrives in Greece - or with a private tutor, as they are under great pressure to obtain an official language certificate.
The Quick placement test was administered during one 50 minute lesson; all students were given a copy of the test as well as an answer sheet. They were asked to do only the first part; the listening part was not administered. The importance of not leaving any questions unanswered was also stressed. The Quick Placement test consists of 40 questions testing reading skills and core competence, and the time allotted for its completion is 30 minutes.

One useful feature of this placement test is that it reports scores aligned to the CEFR levels. This is of vital importance for our study, since the creation of a pedagogic corpus that would be used for the DDL lessons should contain texts that corresponded to a particular CEFR level. This was a *sine qua non* for our study because the experimental and the control groups’ teaching material should be the same CEFR level. The teaching methodology would differ, but not the content of teaching.

As shown in the table that follows, the majority of students were at B1 level with the next largest number being at level B2; consequently, it was decided that the most appropriate CEFR level for our target group was B1+.

<table>
<thead>
<tr>
<th>CEFR Level</th>
<th>Students</th>
<th>Students percent</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>A1</td>
<td>6</td>
</tr>
<tr>
<td>2</td>
<td>A2</td>
<td>3</td>
</tr>
<tr>
<td>3</td>
<td>B1</td>
<td>38</td>
</tr>
<tr>
<td>4</td>
<td>B2</td>
<td>14</td>
</tr>
<tr>
<td>5</td>
<td>C1</td>
<td>3</td>
</tr>
</tbody>
</table>

5.6.4 The pre, post and delayed-post experimental tests

After the training sessions had taken place, a pre-experimental test\textsuperscript{11} on modal verbs was administered to all three classes, that is, is both the experimental and the control

\textsuperscript{11} See Appendix D.1
groups. Pre- and post-tests are used to “determine [training course’s participants] knowledge level of the course content” [I-TECH 2010: 1].

The test was created specifically for the purposes of this research and included 34 multiple choice questions out of which 21 covered all main functions of the modal verbs will, would, may, might, could, should. Multiple choice tests questions have been claimed to have the advantages of versatility, reliability and validity (Brame 2013) and can therefore be an effective way to assess existing knowledge and learning outcomes.

Each of 34 question items in our test is followed by a list of four suggested alternatives; one of them is the solution to the question and the other three are the incorrect alternatives, also known as distractors. The sentences of the multiple-choice questions were made with material taken from the pedagogical corpus ECCo and the online Corpus of Contemporary American English (COCA).

The six modal auxiliary verbs ‘will’, ‘would’, ‘may’, ‘might’, ‘could’ and should include both high (e.g. ‘will’) and low (e.g. ‘might’) frequency modal verbs (see table 5.2) which previous teaching experience showed that are very likely to confuse students. The test also included distractors focusing on other aspects of the English language, such as tenses, articles, prepositions and connectives.

| Frequency in the Corpus of Contemporary American English (COCA) (in tokens) |
|-----------------|-----------------|-----------------|
|                  | Spoken section  | Newspaper section |
| 1. May           | 63.946          | 54.627           |
| 2. Might         | 45.338          | 39.336           |
| 3. Should        | 88.819          | 59.606           |
| 4. Could         | 137.737         | 113.880          |
| 5. Will          | 213.890         | 224.317          |
| 6. Would         | 266.727         | 188.137          |

The pre-experimental test was administered during the class hour prior to the DDL-based lessons. In the beginning the students were given the appropriate instructions. More specifically, they were told they were going to take a test in English that has 34
multiple choice questions. They were also told that it would take them about 20 minutes to answer all of the questions and that they should not worry about how many questions they think will get wrong.

In the lesson following the DDL-based learning sessions, a post-experimental test on modal verbs was administered, identical to the pre-experimental test. Students’ scores on both these tests as well as on the delayed-post experimental test they sat for 8 months after the post-experimental test, provided us with the data necessary to carry out the statistical analysis which led to conclusions regarding the usefulness of the proposed DDL methodology.

5.6.5 The ECCo Corpus

The merits of using a pedagogic corpus in EFL learning and teaching have been discussed in detail in chapter 3. Our decision to create a specialized pedagogic corpus was dictated by the advantages of using such a source of raw data as the basis for the design of DDL activities; in particular, the specialized corpora are “carefully targeted” and “likely to allow the occurrence of specialized lexis and structures with more regular patterning and distribution, even with small amounts of data” (O’Keeffe et al 2007: 199). Therefore, our decision to compile a specialized corpus specifically for the purposes of the present experimental study was based on the necessity to ensure that all students, both those of the experimental and of the control groups were taught mostly through the use of B1+ material.

This specialized corpus was a pedagogic one, since it included only material taken from EFL course books used in senior high schools.

5.6.5.1 Corpus material choice

The second step in the compilation procedure of ECCo had to do with collecting the appropriate data that would fit the requirements or our experimental study. As pointed out in Palmer et al (2010: 259) “the criteria for corpus selection depend closely on the objectives for intended use”.
The material that we intended to produce had to be appropriate for use in a Greek senior high School. Therefore, its content had to conform to the curriculum dictated by the Ministry of Education. Since the present study’s target group were students in the 1st year of senior high school, our decision was that this corpus should consist of material contained in English textbooks, approved by the Greek ministry of Education as suitable for the teaching of English at the 1st year of Senior High school.12

For this reason, we made a selection of specific course book titles by various publishers, included in the Ministry’s list of approved teaching material issued in 2011. The aforementioned list contains 125 course book titles, 15 titles of optional supplementary material and 9 titles of interactive whiteboard support material. The course book titles cover all levels defined by the Common European Framework of Reference for languages (CEFR), that is from A1 to C2.

For economy reasons we decided to use five of the aforementioned course book titles by various authors. In order to decide which five titles to use for the creation of our corpus, we made an investigation on the most popular course books by State school teachers who work with students of the 1st grade of senior high school. For this reason, we asked the publishers to give certain information on the number of schools that chose one of their titles as teaching material for the school year 2010-11. However, the response was not satisfactory, as the vast majority of publishing houses we contacted were reluctant to give any details on the orders they received from state schools.

For this reason, we contacted all appointed State school advisors responsible for English language teachers in various geographical areas in Greece and asked them to provide us with information regarding which course book title was chosen by each one of the Senior High schools in the region in which they were appointed during the school year 2010-11. Again, the results were somehow disappointing. Out of 35 advisors, only five responded to our request and out of 1,197 Senior High schools in Greece, only 60

---

12 As discussed earlier in this chapter, the textbook material for the teaching of English in Senior High schools is not provided by the Greek ministry of Education. Instead the teacher selects a title from a list of freely distributed commercial material, which is provided by the ministry of Education at the beginning of every school year – usually early September.
provided information, providing thus poor data. However, even in that low sample of 5%, some clear preferences for specific titles were evident. This led us to include in our corpus material contained in the following course books: ‘The Outsiders’ (Hillside Press), ‘Journeys’ (Hillside Press), ‘What if’ (Hillside Press) and ‘Upstream’ (Express Publishing and ‘Enterprise’ (Express Publishing).

Unfortunately, once the total number of pages of the above-mentioned course books was scanned, the new list of approved teaching material for the school year 2011-12 was issued, which excluded the course book ‘What if’ (Hillside Press). Therefore, a new course book had to be selected. In early September 2011 all schools received the official list of approved material for use in Senior High schools, issued by the Ministry of Education, Lifelong learning and Religious affairs, which dictated that only those textbooks issued during the last five years should be used. I decided to replace the ‘What if’ title with the ‘Activate’ course book by Longman. This choice took into consideration state school teachers’ responses, albeit few, regarding the selection of course book they made.

Later on, it was decided that, since the actual in-class experiment would take place during the 2012-13 school year, only the titles published from 2007 onwards should be used. ‘Activate’ (Longman), ‘Outsiders’ (Hillside Press), ‘Journeys’ (Hillside Press), and ‘Upstream’ (Express Publishing) and ‘Enterprise’ (Express Publishing) were therefore eligible since they were included in the Ministry’s prescribed list.

As discussed in Section 5.4, another decision that had to be made was one concerning the CEFR level each of the abovementioned titles corresponds to, as it is common for EFL publishers to produce course books that correspond to a combination of CEFR levels. For example, the ‘Outsiders’ series includes 3 different course books covering levels B1-B2, the ‘Journeys’ series includes 2 course books covering levels B1-B1+, and the ‘Upstream’ series has six course books covering levels A1-C1. For this reason, we based our choice on the results of the Quick Placement Test. The results showed that the vast majority, 80% of the students, were either at level B1 or B2. Therefore, we decided to include in our corpus only course books of B1+ level according to CEFR.
5.6.5.2 The ECCo corpus compilation

The next step was the compilation of the corpus. Since an electronic version of the aforementioned titles is not publicly available, we decided to do it manually. Therefore, the conventional method of scanning and processing though an Optical Character Recognition (OCR) program was chosen. This was a laborious and time-consuming procedure since it was done for each page separately, and many corrections had to be made due to the lack of 100% accuracy of the OCR program.

Another decision had to be made concerning the content of my corpus, in particular which parts of the five course books were to be included. This heavily depends on our definition of a pedagogic corpus. There have been researchers who claim that a pedagogic corpus consists “of all the language a learner has been exposed to” (Hunston 2002: 16); if that is the case, ECCo is a clear case of a sample of a pedagogic corpus, its content being part of the material included in B1+ level English course books. Instead of using all the language available in the five textbooks, we excluded the cover page words, as well as a great percentage of the activities (reading comprehension, grammar, listening and writing). On the other hand, we used only full texts, which are included in the course book and constitute part of the reading and writing tasks, as well as the listening transcripts. Our decision to use only full texts was based on the requirements of the present study; the corpus material would provide the concordances we needed in order to create the DDL activities, and those concordances had to contain full sentences. If we had included in our corpus other material contained in the course books, for example the activities, we would have to deal with fragmented sentences, isolated words and erroneous sentences which wouldn’t have been useful for the purposes of the present study.

The texts we selected were saved in txt format (Figure 5.6), without any bold or italics and without indication as to paragraph change. The txt (or ‘plain text’) format was deemed appropriate for storage as it “offers the maximum flexibility of use with different software suites” (O’Keeffe et al. 2007:6).
All in all, the ECCo corpus consists of 4 subcorpora which include 104,238 tokens, 8,664 types and 7,721 sentences, as indicated in the corpus statistics figure that follows (Figure 5.7).
Finally, regarding the frequency of modal verbs, corpus statistics have shown that all modal verbs are present in ECCo. More specifically, 1,363 of all words are modal verbs (including the six modal verbs examined in the present study). The frequency of modal verbs in ECCo is demonstrated in Table 5.3.

<table>
<thead>
<tr>
<th>N</th>
<th>Overall</th>
<th>1</th>
</tr>
</thead>
<tbody>
<tr>
<td>text file</td>
<td>Overall_subcorpora.txt</td>
<td></td>
</tr>
<tr>
<td>file size</td>
<td>579,239</td>
<td>579,239</td>
</tr>
<tr>
<td>tokens (running words) in text</td>
<td>104,238</td>
<td>104,238</td>
</tr>
<tr>
<td>tokens used for word list</td>
<td>103,465</td>
<td>103,465</td>
</tr>
<tr>
<td>sum of entries</td>
<td></td>
<td></td>
</tr>
<tr>
<td>types (distinct words)</td>
<td>8,664</td>
<td>8,664</td>
</tr>
<tr>
<td>type/token ratio (TTR)</td>
<td>8.37</td>
<td>8.37</td>
</tr>
<tr>
<td>standardised TTR</td>
<td>43.72</td>
<td>43.72</td>
</tr>
<tr>
<td>standardised TTR std.dev.</td>
<td>54.75</td>
<td>54.75</td>
</tr>
<tr>
<td>standardised TTR basis</td>
<td>1,000</td>
<td>1,000</td>
</tr>
<tr>
<td>mean word length (in characters)</td>
<td>4.38</td>
<td>4.38</td>
</tr>
<tr>
<td>word length std.dev.</td>
<td>2.31</td>
<td>2.31</td>
</tr>
<tr>
<td>sentences</td>
<td>7,721</td>
<td>7,721</td>
</tr>
<tr>
<td>mean (in words)</td>
<td>13.40</td>
<td>13.40</td>
</tr>
<tr>
<td>std.dev.</td>
<td>8.34</td>
<td>8.34</td>
</tr>
<tr>
<td>paragraphs</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>mean (in words)</td>
<td>103,465.00</td>
<td>103,465.00</td>
</tr>
<tr>
<td>std.dev.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>headings</td>
<td></td>
<td></td>
</tr>
<tr>
<td>mean (in words)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>std.dev.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>sections</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>mean (in words)</td>
<td>103,465.00</td>
<td>103,465.00</td>
</tr>
</tbody>
</table>

Table 5.3: Corpus statistics for ECCo.
Table 5.3: The frequency of modal verbs in ECCo

<table>
<thead>
<tr>
<th>Modal verb</th>
<th>Occurrences</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Can</td>
</tr>
<tr>
<td>2</td>
<td>Would</td>
</tr>
<tr>
<td>3</td>
<td>Will</td>
</tr>
<tr>
<td>4</td>
<td>Could</td>
</tr>
<tr>
<td>5</td>
<td>Should</td>
</tr>
<tr>
<td>6</td>
<td>Might</td>
</tr>
<tr>
<td>7</td>
<td>May</td>
</tr>
<tr>
<td>8</td>
<td>Must</td>
</tr>
<tr>
<td>9</td>
<td>Shall</td>
</tr>
<tr>
<td>10</td>
<td>ought to</td>
</tr>
<tr>
<td>Total</td>
<td></td>
</tr>
</tbody>
</table>

5.6.5.3 Metadata information

It is important that the creation procedure of a corpus should be “documented fully with information about the contents and arguments in justification of the decisions taken” (Wynne 2005: 13). We also decided that metadata information, (also called ‘structural markup’ in Meyer 2002: 81), should be included in a separate file addressing all sub-corpora. By ‘metadata’ we mean contextual information about the ‘primary’ (digital language) data. More specifically, a ‘header’ in the form of a Text Encoding Initiative (TEI) compliant .xml file was created, containing the following information for each one of the 5 sub corpora respectively:

- General information on ECCo; date of compilation, name of author
- Course book title name, author and publication details
- Indication as to whether each sub corpus contains ‘reading/ writing practice texts’ or ‘transcripts’.
Corpus annotation has been defined as “the practice of adding interpretative linguistic information to a corpus” (Leech 2005: 17). It is in other words, the linguistic interpretation of the primary data. Annotation should be clearly specified and led by the motives and objectives of any research. For the needs of the present study which focuses on modal verbs only, a grammatical analysis of the ECCo corpus in the morphological level was deemed necessary. This was achieved with part-of-speech (POS) tagging, a decision in agreement with McEnery et al. (2006: 18) who claim that “corpora constructed for grammatical study should be POS tagged”.

Annotating a corpus can be a highly time-consuming and prone to errors process; for this reason, automatic annotation can be of great help. The present study made use of the CLAWS tagger (Constituent-Likelihood Automatic Word-tagging System) created at the University of Lancaster. This tagging system automatically attaches a tag to each word indicating its grammatical class, or part-of-speech (POS), and to each punctuation mark. The CLAWS system was successfully used in the tagging of some of the most popular large corpora such as the BNC (Leech et al. 1994) and the LOB corpus (Garside et al. 1987). The advantages of the CLAWS tagger are that (a) it offers annotation with a high level of accuracy (approximately 98%) compared with the output of other
taggers, (b) it is freely accessible, and (c) that it has a user-friendly online interface without setting a word limit. Therefore we were able to have the whole bulk of our corpus tagged in a very short time.

5.6.5.5 ECCo Concordancing

Having compiled and tagged ECCo, the next step was to design appropriate learning material. Corpora are “simply linguistic data and [...] specialized software tools required to view and analyze them” (Anthony 2013:141). This in the DDL case is done with a concordancer. Most online available corpora nowadays have a built-in concordancing program, but in the case of handmade corpora which were built for specific, research or teaching, purposes one has to choose among many effective concordancing programs that are available.

When choosing a concordancing program, we should bear in mind what information we want from our corpus (Bennett 2010). Based on this assumption, among those concordancers available either without charge or with a subscription, I chose the WordSmith tools concordancer. Its utilities include, but are not limited to, producing concordances for finding all instances of a word or phrase, keywords which help find
salient words in a text or set of texts, and frequency or alphabetical word lists\textsuperscript{13}. Its advantages are that it can produce concordances out of a corpus of a large number of words and it can also handle tags (Reppen 2001); last but not least, it has a user-friendly interface. Another feature that proved quite useful for the creation of the DDL activities was the WordSmith Tool’s ability to ‘blank out’ target words in the concordance output. Over the two school years in which the present study’s DDL-based lessons were created, use was made of versions 5 and 6 of WordSmith tools (Scott 2008 and Scott 2012).

5.6.6 The COCA

The compilation of ECCo provided the data for the creation of pedagogically appropriate teaching material that would be both easy to become understood by B1+ level students and would also be in relevance with the material used for teaching the control group the ‘traditional way’. This material would therefore provide common ground for comparable results.

However, the question of language authenticity and the DDL demand for a pure hands-on learning approach still remained. The language contained in course books has often been labeled as ‘artificial’. Therefore, a freely available online corpus of natural occurring English was deemed necessary to complement the present study’s DDL methodology.

Moreover, if the present study only made use of a small corpus such as ECCo to teach modal verbs the results would be extremely limited (Sinclair 2004). “If the corpus is small it can only provide a small window on the language phenomenon under investigation and hence, the results will only provide a partial picture of its ‘true’ complexity… a large corpus will provide a more complete view of the phenomenon” (Anthony 2013: 146). Therefore, it is imperative for the implementation of a pure DDL approach that “we go beyond […] a couple of examples given in textbooks, [and take evidence] of what a large number of proficient English users say and write in different settings” (Conrad 2016: 38).

\textsuperscript{13} Information derived from www.lexically.net
The Corpus of American Contemporary English is the largest available free corpus of English available, containing 520 million words\textsuperscript{14}. The COCA is a “powerful, reliable and dependable tool” (Huo 2014: 916), it can be accessed freely and it provides an abundance of naturally occurring English from various genres such as Spoken, Fiction, Magazine, Newspaper and Academic. It is also extremely popular, being at the moment “probably the most widely used corpus of English” with more than 130,000 unique users per month\textsuperscript{15}. Last but not least, it offers a user-friendly interface that would not deter students from trying a hands-on approach to studying collocations and looking for patterns in language. COCA has been used in several case studies such as the ones by Boulton and Landure (2016), Quan (2016), Tekin (2015), Yilmaz and Soruç (2015) to name but few of the most recent studies.

The COCA online platform was used for the purposes of the present study in two different ways. First KWIC searches were made by the teacher so as to draw examples from random corpus lines in COCA; those examples were used for the creation of handouts with DDL activities. Secondly, it was used by the students themselves. They were guided to enter the COCA online platform and form their own KWIC queries. They searched for examples of use of modal verbs and their corpus queries yielded a surprising number of occurrences, giving them thus the chance to ‘immerse’ into naturally occurring language and act as the ‘language detectives’ envisioned by Tim Johns (1997: 101).

The next chapter will discuss the DDL lessons created which were based on the data taken from ECCo and COCA.

\textsuperscript{14} It has to be noted at this point that the searches on COCA used for the purposes of the present research were carried out during 2013-15 prior to the change in the online interface- therefore some of the searches results presented here have been made using the older interface version. Moreover, since then the corpus has grown with the addition of millions of words; the frequencies shown have made use of data until before 2014.

\textsuperscript{15} Data derived from http://corpus.byu.edu/users.asp site
CHAPTER 6

THE DDL LESSONS

6.1 Introduction

This chapter discusses the application of the proposed DDL methodology in a Greek Senior High school class. Having compiled the pedagogic corpus, ECCo, and having decided on the additional use of a free online corpus, COCA, we had to design relevant DDL activities that would focus on teaching the use of modal verbs in English and make those DDL activities the building blocks of five 45-minutes DDL lessons.

The first part of this chapter presents the training sessions that were necessary for the students to familiarize themselves with the DDL methodology. The second part discusses in detail the DDL lessons during which the students were introduced to, discovered and practiced the English Modal Verbs.

6.2 The Training sessions

Previous studies, (Hatzitheodorou & Mattheoudakis 2007: 99) and Flowerdew (2012: 205) among others, have stressed the need for appropriate training on DDL, since the majority of teachers and all students are not familiar with such a methodology. In the case of the present study, I was both the designer of the proposed DDL methodology and the teacher who applied it in class. None of the students who participated in the present study had any previous experience in using corpora, as they were unfamiliar with either working with printed concordances or online corpus searches. For this reason, it was necessary that they become accustomed to being “learning detectives” (Johns: 1991). Bearing this necessity in mind, all students who belonged to the two experimental groups took some training on corpus searching.
The students’ training sessions were designed one year after the pilot study had taken place. In order to design the activities of these five training sessions, we followed the typology of induction exercises for introducing students into investigating language through corpora, provided by Lee and Swales (2006: 66). Amongst other activities Lee and Swales suggest guessing the meanings of words by studying concordances and looking at similar lexical items and puzzling pairs, activities that were included in the present study’s training sessions.

The three sessions which had been designed for the pilot study were later revised so as to become more user-friendly and were enriched with more content. Two more training sessions were added to the main study which eventually comprised five sessions in total, each one of which had a 45-minute duration. Those sessions provided training on the following; using the web as a corpus, working on printed (and edited) corpus material, and engaging on hands-on DDL activities.

6.2.1 Training session 1

The first training session’s aim was twofold. Firstly, it aimed to introduce students to KWIC searches in a non-intimidating way and secondly, to provide the students with their first ‘hands-on’ activity on the Corpus of Contemporary English (COCA).

Our objective to ‘gently’ initiate students into corpus queries was done by “taking the ‘KWIC’ term out of the equation” Boulton (2010a: 534). Instead of the concordance-based KWIC search we chose a way that is no foreign to them; namely Google searches. Using the Web as a source of texts that can be used for corpus searches has been suggested by various researchers in the past. For instance, Foucou and Kubler (2000) describe a web-based environment for teaching technical English, whereas Bennett (2010) presents the WebCorp suite of tools, which allow access to the World Wide Web as a corpus, in a list of corpus-related materials that the readers may find useful.

In the first training session, which closely resembles the pilot study’s first session (Ch. 5), all students worked in pairs on one of the school lab’s PCs with internet access. The first activity of this session involved using corpus or (corpus-like) resources to find the meaning of an unknown phrase. We chose to ‘investigate’ the phrase ‘swirl of innuendo’ after encountering it in an article presented in class. The first step of this
discovery learning procedure was to google the word ‘innuendo’ and try to guess its meaning by the hits the search gave them. The second step was to consult an online English dictionary and see the meaning of the word as well as the examples that complement each entry. This simple activity was actually very important, as it was quite easy for all students to engage in and thus gave them confidence to proceed into the unknown ‘corpus investigation territory’.

The third step of this training session involved students querying COCA. They were guided to use in pairs the built-in COCA concordancer in order to make a simple KWIC search of the word ‘innuendo’. Afterwards they were asked to examine the first 10 result lines and choose and discuss in class one of the lines containing that word.

Figure 6.1: KWIC query of ‘innuendo’ in COCA
Let’s discover what ‘innuendo’ means!

Step 1. google’ the word and guess by the ‘hits’

“Innuendo may mean………………………………………”

Step 2. Search an online English-English dictionary (e.g. word-reference.com)

“Innuendo may mean………………………………………”

Step 3. Search an online English-Greek dictionary

- Enter the website of COCA. Type the following link
  
  http://corpus.byu.edu/coca/

  The website should look like this

- Type the keyword ‘innuendo’ and press KWIC
  
  KWIC stands for ‘Key Word in Context’ and the result will show the keyword you typed within a sentence
- Have a look at the first 10 result lines
- Write one example containing the word ‘innuendo’ and explain its meaning in class

……………………………………………………………………
……………………………………………………………………
6.2.2 Training session 2

This training session was exclusively focused on the use of the built-in COCA concordancer. The aim was for the students to familiarize themselves with raw concordance output and take two steps further into corpus searching; the first one concerned engaging into more refined queries and the second one querying for chunks, rather than single words.

At the beginning of the 2nd training sessions, the phrase ‘swirl of innuendo’ was written on the board. By now, the students knew the meaning of ‘innuendo’ but not that of ‘swirl’. They were once again guided into entering the COCA platform and typing the search word ‘swirl’. In this training session, however, they were advised to narrow down the query output by choosing to investigate only ‘spoken’, ‘magazines’ and ‘newspaper’ subcorpora. The rationale behind this choice of ours was that the ‘academic’ and ‘fiction’ subcorpora would be more likely to contain either specialized terminology or dated vocabulary, whereas ‘spoken’, ‘newspaper’ and ‘magazines’ subcorpora were expected to provide mostly everyday language.
More specifically, in this training session, the students focused on the information that can be drawn from each concordance line. They paid attention to the fact that different parts of speech are highlighted in different colors, and that helped them to draw their own conclusions as to which words usually precede and follow the node word swirl.

Previous studies have stressed the importance of acquiring new vocabulary and grammar structures through chunks. O’Keeffe et al (2007: 76 & 79) claim that the use of chunks is central to fluency and that students should be encouraged to record them in order to raise awareness of their usefulness. Lewis (1994) acknowledges the identification of lexical chunks as a classroom activity which is worth all our attention, while Baigent (2004) advocates learning chunks for faster and more natural speech. In the case of the present study, the students performed a query of the chunk ‘swirl of’, rather than a single lexical item. Finally, they were guided to extract a full sentence out of a concordance line and share their findings with the rest of the class.

Figure 6.4: The 2nd training session handout

---

**So, what is a ‘swirl of innuendo’?**

*Enter the website of COCA. Type the following link*

http://corpus.byu.edu/coca/

1. type the word ‘swirl’ in the search box
   - choose KWIC (from the choices above)
   - choose sections ‘spoken’ and ‘magazines’ (from the ‘sections’ category below)
   - now, look at the first 10 result lines and answer the following questions:
     iv. which is the word in the middle? ....................................................
     v. what info is provided in each line? ....................................................
     vi. what colour is the keyword? ....................................................
     iv. why is it highlighted? ....................................................
     v. which other words are highlighted? Why? ....................................................

2. type the phrase ‘swirl of’ in the search box
   i. write down a sentence containing the phrase ‘swirl of’ and explain its meaning in class .................................................................
6.2.3 Training session 3

The 3rd training session on DDL intended to consolidate knowledge acquired in the 2nd session. The students made use of the same online corpus, COCA, and, though their queries did not differ from the ones they had engaged in before, in this case they had the relative freedom of choosing what to investigate among the choices provided. Here the focus was on clarifying the meaning and distinguishing the use of words that can be easily confused, which, according to Higgins (1991), is one of the basic objectives of classroom concordancing.

Pairs of such words were selected from Oxford Dictionary and included ‘affect/ effect’, ‘adverse/averse’, ‘born/ borne’ and ‘complement/ compliment’. As in the previous training sessions, the objective here was an actual corpus consultation. The students were also given a worksheet to fill with the findings of their ‘lexical investigation’ on COCA.
6.2.4 Training session 4

In the 4th training session, use was made of the material contained in the pedagogic corpus ECCo. This training session involved students working individually on a printed handout with edited concordances. These lines were produced by the teacher using WordSmith Tools 6 concordancer and included only full sentences containing instances
of verbs in Past Simple and Past Continuous. It has been pointed out that one of the objectives of classroom concordancing is to use queries for grammatical structures as a means for assisting students to discover grammatical rules (Higgins 1991). It was the first attempt to revise certain grammatical structures, in this case the tenses Past Simple and Past Continuous, using an –indirect- DDL approach. The students were asked to investigate the sentences contained in the worksheet for any words/ structures that are repeated. The final goal was to use their findings in order to make generalizations and form their own rules regarding the formation of the above-mentioned tenses.

Figure 6.7: The 4th training session handout

_Hidden words, revealed structures_

- Have a look at the following sentences. All of them are sentences taken from course books. Then, answer the questions

1. I realized I'd left them on the kitchen table as soon as I closed the front door. I'm so angry with myself. It's such a stupid thing to do!
2. Backing silently towards the bedroom door, she kept her eyes on the window. She expected a horrible alien face to appear there at any second. Suddenly she felt something behind her.
3. I closed my eyes and opened them again. Several Japanese tourists were admiring the cow. Some were taking photos, others were laughing!
4. Then something caught her eye - one of the dummies was moving! The dummies were coming to life and they started to walk slowly towards her.
5. While they were drinking their coffee, she showed them a leaflet.
6. During the night of 19 September, 1961, an American couple called Betty and Barney Gill were driving home to Portsmouth, New Hampshire from a holiday.
7. Rose panicked and tried to run away, but they were getting nearer. She screamed in horror.
8. E: Why did you hit my bike? Couldn't you see where you were going?
   F: Well, no, I couldn't because the sun was in my eyes.
6.2.5 Training session 5

In the final training session, the students practiced the skills acquired in the 2nd and 4th session. By now, they were quite confident in their searches and were able to conduct corpus investigation with less interference and guidance by the teacher.

In this last training session, the words ‘deny/ refuse’ were written on the board. The goal of this activity was for students to observe these two words in a real context which illustrates how they are used by native speakers. The students were asked to form the query ‘deny/ refuse’ in COCA and observe which words frequently come before and after each one of them. By noticing the patterns in this task, they formed the following two general rules for those two lexical items; they are both verbs and their meaning is similar, but, when followed by a verbal type, it is a gerund in the case of ‘deny’ and a full infinitive in the case of ‘refuse’.
**Words which confuse us**

**Step 1** Today, we’ll investigate the following words:
- deny/ refuse

**Step 2** Enter the website of COCA. Type the following link
http://corpus.byu.edu/coca/

**Step 3** type the words in the search box: ‘deny/ refuse’

i. choose COMPARE (from the choices above)

ii. choose sections ‘spoken’ (from the ‘sections’ category below)

iii. now, look at the first 10 result lines and answer the following questions:
- which words usually appear next to the words ‘deny’ and ‘refuse’? ……………………………………………………………………………
- does this gives you any hints concerning the meaning of these two words? ……………………………………………………………………………

**Step 4** type the words in the search box: ‘deny/ refuse’

iv. choose KWIC (from the choices above)

v. choose sections ‘spoken’ and ‘magazines’ (from the ‘sections’ category below)

vi. now, look at the first 10 result lines and answer the following questions:

d) write down what part of speech is each word ……………………………………………………………………………
e) write down an example of a sentence containing the word ……………………………………………………………………………
f) try to guess the meaning of the words ……………………………………………………………………………
g) What part of speech usual comes before/ after ‘deny/ refuse’? ……………………………………………………………………………
h) Can you write an example of your own? ……………………………………………………………………………
6.2.6 Feedback from the training sessions

Not all students gave an answer to the training sessions’ activities and not all of those who answered were successful. In sessions 1, 2, 3 and 5 the students worked in pairs due to the lack of an adequate number of computers. In session 4 they worked individually.

The general impression derived from short informal in-class interviews which took place after the five training sessions were concluded, was that students were not intimidated by the novelty of the DDL approach. Moreover, they seemed well-acquainted with computer technology and they didn’t complain either about difficulty or irrelevance with the ‘mainstream’ learning practices. Similar conclusions were derived from their answers given to the questionnaire administered after the end of the experiment.

A 7-item questionnaire was given to student participants in order to assess the DDL approach to EFL teaching, using two points (yes/ no) rating scale (Table 6.1). All 64 students in this study stated that the corpus-based methodology is something they had not encountered before and agreed on the usefulness of using a computer to find the
meaning of an unknown word. What is more, the majority of students agreed that on
line corpora can also be of benefit for language learning (see Figure 6.10: Questions 1, 2 and 3). Concerning questions 4 and 5 about the level of difficulty of engaging in
corpus searches, the answers appear to go both ways; although the majority of students
claim that making corpus queries is easy, more than half of them admit that those
queries look awkward to them (see Figure 6.10, Q 4 and 5). In the case of questions 6
and 7 the vast majority of the respondents underlined their preference for examples and
more than half students chose concordances over rules when learning grammar.

In brief, the questionnaire findings highlight the importance of computer technology in
language learning, and in lesser degree the usefulness of corpora too. The second major
finding of the training session’s evaluation questionnaire is that the majority of students
are in favor of learning through observation of examples rather than memorizing rules,
even though a significant number of students find it difficult to conduct corpus queries.

Table 6.1: The training sessions’ evaluation questionnaire results

<table>
<thead>
<tr>
<th>Questions</th>
<th>Students</th>
</tr>
</thead>
<tbody>
<tr>
<td>1  The past five lessons were novel for me</td>
<td>100%</td>
</tr>
<tr>
<td>• Yes</td>
<td>0%</td>
</tr>
<tr>
<td>• No</td>
<td></td>
</tr>
<tr>
<td>2  Using the computer to find the meaning of an unknown word is useful</td>
<td>100%</td>
</tr>
<tr>
<td>• Yes</td>
<td>0%</td>
</tr>
<tr>
<td>• No</td>
<td></td>
</tr>
<tr>
<td>3  Using an online corpus to find the meaning of an unknown word is useful</td>
<td>72%</td>
</tr>
<tr>
<td>• Yes</td>
<td>28%</td>
</tr>
<tr>
<td>• No</td>
<td></td>
</tr>
<tr>
<td>4  Using an online corpus to find the meaning of an unknown word is easy</td>
<td>61%</td>
</tr>
<tr>
<td>• Yes</td>
<td>39%</td>
</tr>
<tr>
<td>• No</td>
<td></td>
</tr>
</tbody>
</table>
Reading concordance lines is awkward
- Yes
- No

<table>
<thead>
<tr>
<th>Yes</th>
<th>No</th>
</tr>
</thead>
<tbody>
<tr>
<td>53%</td>
<td>47%</td>
</tr>
</tbody>
</table>

Reading examples is preferable to studying grammar rules
- Yes
- No

<table>
<thead>
<tr>
<th>Yes</th>
<th>No</th>
</tr>
</thead>
<tbody>
<tr>
<td>91%</td>
<td>9%</td>
</tr>
</tbody>
</table>

Reading concordance lines is preferable to studying grammar rules
- Yes
- No

<table>
<thead>
<tr>
<th>Yes</th>
<th>No</th>
</tr>
</thead>
<tbody>
<tr>
<td>64%</td>
<td>36%</td>
</tr>
</tbody>
</table>

Figure 6.10: The training sessions’ evaluation questionnaire results

6.3 The DDL lessons

After the training sessions took place, the learners were familiar with concordances and had some training on how to make their own queries. The results of the post-training
sessions’ questionnaire were quite encouraging to proceed with the main DDL lessons of our study.

During April and May 2015 five lessons on modal verbs took place. Those five lessons focused on the main senses of modal verbs, their grammatical properties and the use of the pairs ‘will / would’, ‘could/ should’ and ‘may/might’. Each was given a title that would attract the attention of the students and also implicitly introduce them to the content of the lesson (see Table 6.2).

The experimental group followed a DDL approach, while the control group was taught using the course book and relying solely on the material provided in it.

Table 6.2: The grammar focus of the 5 DDL Units

<table>
<thead>
<tr>
<th>The DDL lessons’ titles</th>
<th>The lessons’ focus</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Words can be powerful, inspirational, funny…</td>
<td>Main senses of Modals</td>
</tr>
<tr>
<td>2. Shall we?</td>
<td>Grammatical properties</td>
</tr>
<tr>
<td>3. Will you join us?</td>
<td>Will / would</td>
</tr>
<tr>
<td>4. You can do it!</td>
<td>Could / should</td>
</tr>
<tr>
<td>5. Shoot for the moon; you might get there!</td>
<td>May / might</td>
</tr>
</tbody>
</table>

In all five DDL lesson the students engaged in both direct and indirect DDL activities. The direct DDL approach involved students’ own querying and interpreting the output of COCA built-in concordancer, whilst the indirect activities were in essence worksheets of corpus-based tasks developed by the teacher. This is a blended DDL approach, since it involves both handouts of printed concordance output and ‘pure’ hands-on DDL. The different access configurations to data resources are illustrated in figure 6.11.
Figure 6.11: Configurations of access to the data resources used in the present study’s DDL lessons (adapted by Lew 2009: 298)

For the main study’s DDL lessons on modal verbs, appropriate DDL activities had to be designed. Very few studies on DDL have described the actual design and implementation of DDL tasks in a language classroom; Johns (1991a and 1991b) is the pioneer on DLL in-class application, whereas Tognini-Bonelli (2001), O’Keefe et al. (2007), Chujo and Oghigian (2008 and 2010), Tyne (2012) and Hirata and Hirata (2015) are among those who have published further fieldwork studies.

The rationale for the design of those DDL lessons was in accordance with the “desiderata” set by Gaskell and Cobb (2004: 303) for a corpus-based grammar resource; in our study that resource being the grammar teaching material. The first of these guidelines to corpus-base materials states that the instructional presentation should be as non-declarative as possible; in other words it should highlight repeated patterns rather than present the students with rules. The second guideline draws attention to the corpus-based material itself, arguing that raw concordance output should be complemented by concordance lines appropriately tailored by the teacher, at least initially, so as to make the approach less intimidating. It is this format we selected for the design of the DDL lessons.
6.3.1 Lesson 1: The main senses of modal verbs

The 1st lesson’s primary objective was to provide an introduction on the main senses of modal auxiliaries by setting a purpose; to look closely at popular quotes and investigate their meaning. Quotations provide authentic material which is eye-catching, entertaining and mind challenging at the same time. The usefulness of exploiting quotations in language learning has been discussed in Partington (1998) whereas Munoz & Towner (2009) discuss the benefits of using quotations for educational purposes in relevance to the presence of Facebook in teenagers’ life; quotations are very common in Facebook language and therefore young people are accustomed to them. Their presence in the EFL classroom would thus be a familiar element.

At the beginning of this 45-minute lesson, the students were presented with 6 different popular quotes, each containing a modal auxiliary. Even though the World Wide Web is full of available quotations, it sometimes lacks in quality material, since it consists of both carefully prepared texts and “‘casually’ prepared material” (Mc Enery 2001). After reading the quotes, the students were asked to comment on them. Guided by the answers they provided on the above questions, the students were able to reach the following conclusions; all sentences contain modal auxiliaries and these modal verbs express beliefs, attitudes, opinions and hopes.

Figure 6.12: The 1st DDL lesson handout

<table>
<thead>
<tr>
<th>Words can be powerful, inspirational, funny...</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Believe you can, and you’re halfway there. (T. Roosevelt)</td>
</tr>
<tr>
<td>2. When I let go of what I am, I become what I might be. (Lao Tzu)</td>
</tr>
<tr>
<td>3. You may be disappointed if you fail, but you’re doomed if you don’t try. (B. Sills)</td>
</tr>
<tr>
<td>4. Tension is who you think you should be. Relaxation is what you are (Chinese proverb)</td>
</tr>
</tbody>
</table>
5. People will forget what you said, people will forget what you did, but they will never forget how you made them feel. (M. Angelou)

6. W: If I were married to you, I would put poison in your coffee.
   M: If I were married to you, I would drink it. (1900s’ joke)

………………………………………………………………………………….

• Have you ever come across any on these?
  o Yes
  o No

• If yes, which one(s)?
  …………………………………………………………………………………….
  …………………………………………………………………………………….

• Can you explain its meaning to your fellow students?
  …………………………………………………………………………………….
  …………………………………………………………………………………….

• Which one do you like most?
  …………………………………………………………………………………….
  …………………………………………………………………………………….

• Is any of the above-mentioned quotes difficult to understand?
  …………………………………………………………………………………….
  …………………………………………………………………………………….

• Are there any grammatical structures repeated in those sentences?
  …………………………………………………………………………………….
  …………………………………………………………………………………….

• What is expressed in all these sentences?
  …………………………………………………………………………………….
  …………………………………………………………………………………….
6.3.2 Lesson 2: The grammatical properties of modal verbs

The 2nd lesson was an introduction to the main grammatical properties of modal auxiliaries. The students worked in pairs in the school pc cluster. They were asked to enter the COCA interface and were guided to search through the KWIC results of ‘will’, ‘would’, ‘should’, ‘could’, ‘may’ or ‘might’. Once more they were asked to pay attention to the node word and its surrounding items. One of their first findings was that their search word was not a modal in all sentences; for instance, the word ‘May’ in the sentence “Well, like the whole world back in April, May and June, I was watching in horror” is highlighted grey whereas the word ‘may’ in the sentences “It is tax day as you may already know” and “Is it true that low fat milk may be 2 percent but regular milk is 4 percent?” is purple. This way they learned how to recognize the modal verbs and discard homophones.

In this lesson we chose a ‘guided-inductive’ approach to corpus consultation (Johansson 2009). This approach involves assisting the students to discover the main rules concerning the usage of modal verbs by offering them specific rules-pointing questions. In the case of the present study these questions focused on the NICE properties of modal verbs, them being negation, inversion, code and emphasis (Bas Aarts 2011:68).

Each student was given a worksheet to fill in the answer to the above-mentioned questions. After they had provided an answer to all questions they discussed their findings in class and formed generalisations concerning the grammatical properties of modal verbs.

Figure 6.13: The 2nd DDL lesson handout

<table>
<thead>
<tr>
<th>Shall we?</th>
</tr>
</thead>
<tbody>
<tr>
<td>Step 1 Enter the website of COCA. Type the following link</td>
</tr>
<tr>
<td><a href="http://corpus.byu.edu/coca/">http://corpus.byu.edu/coca/</a></td>
</tr>
<tr>
<td>Step 2 type the words in the search box like that</td>
</tr>
</tbody>
</table>
6.3.3 Lesson 3: will – would

The third DDL lesson focused on the usage of the modal verbs will and would and had a 50-minute duration. Here we followed an indirect approach to DDL in which the students worked on a worksheet rather on a PC. There were three reasons behind this more teacher-centered setting. First of all, this lesson’s objective was to cover all senses of ‘will’ and ‘would’ and this was not possible in the examples provided even in the first 100 lines of the COCA output. The students would eventually tire of searching through endless result lines and would not come up with any results. Moreover, appropriate activities had to be designed and this could only be achieved with carefully selected concordance lines. Last, one could not ignore the obvious time restrictions. Since each DDL lesson had a 50-minute duration, the time was too limited for the students to engage in time consuming corpus searches; the printout provided us with
carefully designed classroom material which, as claimed by Johns (1991b), can prove invaluable in stimulating inductive learning strategies, one of them being the ability to notice similarities and differences.

The first activity was created with selected concordances from ECCo, each containing an instance of ‘will’ or ‘would’. A selection of such edited concordances was considered more appropriate so as to gently initiate the DDL lesson, as adapted materials have been claimed to be more appropriate for low level students (Bennett 2010). Here the students were asked to investigate the lines for some of the grammatical properties by paying particular attention to the words that precede and follow the modal verb.

There were also three activities (2, 3 and 8) which prompted the students to investigate the senses of ‘will’ and ‘would’. In this case we tried to overcome the drawbacks of the purely inductive DDL approach, mainly associated with a number of students disliking working inductively. This drawback has been noted in previous works in the field, for example Gavioli (2005) and Meunier (2002). Therefore we aimed for a combined approach in which we offer the senses of modals and we ask the students to match them to examples of use taken from the corpora.

Two of these tasks gave them a very limited choice of three examples taken from ECCo, while the third one asked them to search through a separate handout which contained 30 concordance lines taken from COCA. In this way, they were given the chance to work both with school text book language they feel comfortable with, and natural occurring language.

Activities 5, 6 and 7 involved deduction; the learners were called to look closely at the 10 concordance lines included in activity 5 and to investigate the text for repeated structures in the co-text of ‘would’. Their observations led them to understand the formation and use of the 2nd conditional and 3rd conditional and thus to form the relevant rules.

This lesson also included production exercises in which the students were able to practice the newly acquired knowledge of modal verbs. We checked their understanding by asking them to use the patterns they had just learnt to create a few sentences of their own. These were activities 4 and 9 and involved using fragments of concordance lines taken from COCA in order to form full sentences.
The last activity included in the lesson was a gap-filling drill, to be done at home. This activity was created with six edited concordance lines from ECCo and its objective was to consolidate the newly acquired knowledge on modals ‘will’ and ‘would’.

Figure 6.14: The 3rd DDL lesson handout

Will you join us?

1. Look at the following sentences. Can you identify any grammatical structures that 'will' and 'would' often appear next to?
   
   a. Today's recipe will serve 4 (people).
   b. The factory has been fined and will be shut until a filtering system has been installed.
   c. ...you are going to have a difficult day because other buskers will have taken the best pitchers.
   d. ...in the future automatic flying cars will be operating without human pilots.
   e. George: Hey... try some of our food!
      John: Sounds great, George. I definitely will.
   f. ...if I say I can't go, my friend will ask someone else...

2. Here you can see 3 sentences including will/ would. Read them carefully, and decide what does each one express?

   **Willingness, certainty, habit**

   a. ...the fair began yesterday, Saturday, and will run for two weeks.
   b. George: Hey... try some of our food!
      John: Sounds great, George. I definitely will.
   c. They want to appear 'cool' and will often conform to peer pressure, so not to feel isolated...

3. Now have a look at your handout and try to find at least one instance of “will' or 'would' expressing **Willingness, certainty, habit**
4. Form full sentences out of the following words taken from CBS, CNN, and ABC TV networks respectively using 'will'.
   a. The president/ address/ nation/ from / Oval Office/ at 8am eastern time.
   b. Bashar Assad/ tell Time magazine/ this week/ Syria/ withdraw/ from Lebanon/ this summer.
   c. LISA McRee: You/ come/ back/ and / talk/ to /us?

5. Can you tell if the underlined phrases refer to the present or the past?
   1. First of all would you say that unemployment was a factor in the rise in crime?
   2. It said that she needed to get away for a few days and would come back later in the week.
   3. A surfer’s greatest disappointment would be missing the opportunity to surf in the best weather conditions.
   4. Lizzie: I would have been shouting at them from day one!
   5. …it would be better if young people learnt about first aid in their free time.
   6. I was nervous at first and I’m glad I speak English or it could have been very difficult.
   7. Let me tell you, if I had known it would really hurt me, I would never have bought it!
   8. Mrs A: Would that mean that we wouldn’t have to pay at all?
   9. Interviewer: I see. Now, if you were to be offered a post, would you be able to start straight away?
   10. If I had known I couldn’t do those things here, I would have taken advantage of my last days in the city.
7. Read the sentences (5), (7), (9), (10) again and find one feature they have in common as far as the context in which 'would' is used is concerned.

..............................................................................................................................................

8. Now, look carefully at sentences (7) and (9) and say how they are different? Can you explain the difference?

..............................................................................................................................................

9. 'would' is used to express willingness and prediction. What do the following express?

  a. A lot of animals would have become extinct by now without zoos to protect them!
  b. I couldn’t understand why anyone would do it.
  c. Building a sports centre would be really expensive too.

10. Form full sentences out of the following words taken from ABC, CNN and ABC TV channels respectively, using 'would'

    a. if/ you/ can/ see / your wife/ what / you /say / to her?
    b. KING: if / you/ be/ a prosecutor now,/ you/ be/ down?
    c. I realize/ it/ be / sloppy/ of you / to be/ too / specific.

..............................................................................................................................................
..............................................................................................................................................
..............................................................................................................................................
..............................................................................................................................................
..............................................................................................................................................
**HOMEWORK**

Let’s try the following activity!

Fill in the following gaps with

a. will / would, and

b. the verb provided in the appropriate form

a. I was wondering if you could answer a few of my queries. I ……….. (be) very grateful.

b. When Frank……. (not) (wake up) Ann got some help from the family dog.

c. How ……. (get) to the theatre? My car broke down and it’s in the shop.

d. If I hadn’t taken the holiday, I suppose I ………. (become) a successful university lecturer, but fate had other plans.

e. Tom: I thought it was a fair punishment, although I ……. (be) very embarrassed if it had been me.

f. Due to the train strike, we’ll be going by coach. It …… (wait) for us here in the morning.

6.3.4 Lesson 4: could - should

The fourth DDL lesson covered the use of Modal Verbs ‘could’ and ‘should’. As in the previous lesson of ‘will’ and ‘would’ it began with a number of carefully concordance lines of ECCo which the students had to investigate in order to discover the way ‘could’ and ‘should’ are used in English and formulate the relevant rules.

The first activity simply involved them looking through 13 sentences and trying to figure out which are the modals they would be working with. The second question asked them to search for any similarities between those two lexical items, while the third and the fourth tasks guided them to discover the senses of ‘could’ and ‘should’ by offering those senses and asking them to match them in each sentence.

Activity 5 was another awareness raising task which made them realize that ‘could’ and ‘should’ can also be found in the 2nd and 3rd conditional. The sixth activity was a fill-in
task in which they could practice what they had learnt in the previous steps. That was the case of the next two activities as well; here the drill was to use fragments of concordances from COCA selected by the teacher. Those included both affirmative, interrogative and negative forms of those two modal verbs.

Last, activity 9 which was assigned for homework aimed to help them consolidate their knowledge and give them the time to reflect on what they had learnt in class.

Figure 6.15: The 4th DDL lesson handout

You can do it!

1. Look at the following sentences. Which are the modal verbs that you can see?

3. Amazing! Could you tell us something about some of the 1980 eruption? (1)
4. I could feel eyes staring at me and could hear unusual noises all around (2)
5. The newspapers are predicting that Becky could be the richest woman in England (7)
6. Instead of going skiing, they could do something less adventurous (169)
7. If these improvements were made, we could all be very proud of our block of flats. (218)
8. The whole class could hear my stomach rumbling in the afternoon! (15)
9. Could I have your passport and ticket, please? (102)
10. Those with tickets for Manchester and North should change at Crewe station. (103)
11. Furthermore, parents should set limits to the hours their children watch TV.. (105)
12. It's a four-hour flight, so they should get in around three thirty. (167)
13. If a stranger approaches, they should go to the nearest safe place.. (158)

2. What is common between could and should?

.................................................................................................................................................................
3. could may express *past ability, formal request, formal permission, general possibility or opportunity*. Which sentence shows each of the above senses?

4. should may express *advice/suggestion, obligation or deduction*. Which sentence shows each of the above senses?

5. What is common between sentences 7 and 13? What do they both express?

6. Look at the following sentences. Fill in either 'could' or 'should'

   1. I think she …..be allowed to skate.
   2. They claim to show just how greedy Shelly……be
   3. Critics insist the private sector…….bring about changes must sooner.
   4. Simon: And we …….mention the book opens with your amazing preview.
   5. Flatow: …….you describe the paintings for us?
   6. Mr Doxiadis: First I …..say, first of all, my co-author is Christos Papadimitriou.

7. Form full sentences out of the following words taken from CBS, CNN, and ABC TV networks respectively, into affirmative, interrogative and forms, using 'could'

   a. Jacob: I/ bet/ anything/ happen
   b. Students / learn/ the color of clothing/ also / mean / taking sides
   c. Aunt: My initial thought/ be/ no,/ Russel/ not/ do/ that.
8. Form full sentences out of the following words taken from ABC, CNN and ABC TV channels respectively, using 'should'

a. Nutley resident: People/ be/ responsible /for their actions
b. The law proposed is not right. It /be/ voted down.
c. I don't know what the moral is. Or how the song/ end.

9. HOMEWORK

Let’s try the following activity!

Fill in the following gaps with
g. could / should
h. a verb in the appropriate form

1. Animals .................(be allow) to go where they want, when they want.
2. Jenny: What other areas ..........biometric technology (be use) in?
3. Well, I believe that every child ........(attend) kindergarten.
4. Under no circumstances ........anyone (drive) or (be) outside once the hurricane has arrived.
5. I was having a wonderful time and ........(stay) there for ever
6. TV Pr: Professor, ........ (I ask) for your views on what makes a good teacher?
7. Mrs A: When do you think Sarah ........ (go) – I mean- if that’s what we decide?
8. I knew that she ........ (give) me detailed directions, and she did.
9. So, ........... (we/ use) our microscopes now?

10. There are many who believe that the story is true. In fact, they ...........(be label) as 'mad' and might even have lost their jobs.

6.3.5 Lesson 5: may - might

In the final DDL lesson, the students were once again presented with concordance lines from ECCo, and then they were asked a series of questions that helped them to identify and generalize the rules that govern the usage of ‘may’ and ‘might’. There were also two activities whose objective was to practice the rules the learners had ‘discovered’ through questions 1-4. The first activity involved matching the two halves of concordances which were derived from ECCo. The second activity was the now familiar production of sentences with fragments taken from COCA spoken section. Finally, there was also a consolidation activity in the form of a gap-filling task to be done at home.

Figure 6.16: The 5th DDL lesson handout

**Shoot for the moon; you might get there!**

1. Have a look at the following sentences. Which are the modal verbs that today’s lesson focuses on?

a) ..and I ’d like, if I may, to give a few handy tips.. (115)

b) May you have a long a healthy life! (176)

c) Dr Porentz: Can I answer?
a. King: You may.

d) Thank you, God bless you, and may God bless the United States of America

e) Female 1: Hi, I'm from a local radio station and Jill thought you might be able
to talk to me about your work... (5)
f) When you administer it to human patients, you might as well call it patient
necrosis factor.
g) Senator: Well, if I might say, I want to repeat...the turtle eggs may be
accidentally crushed by tourists... (1)
h) Doctor: If you give in to phobia, then you may let it rule your life. (9)
i) Excuse me miss, I'm Julia Marquez for WRSB, may I ask your name and
what are you doing in this line? (31)
j) ..Teams may win prizes too (62)

2. ‘may’ is a modal auxiliary verb commonly found in English.
What does it refer to past, present or future?

3. ‘might’ is a modal auxiliary verb commonly found in
English. What does it refer to the past, present or future?

4. May is used to express specific functions.
Can you identify what do the sentences above express?
- general truth
- wish
- permission
- possibility

5. Match the fragments to make full sentences
a. It is important to stay calm, as panic leads to rash actions which..
b. In what ways..
c. Indeed, as our understanding of them grows, they..
d. James and his family started to worry that they..
e. Clouseau is worried that he too.
i. ...may result in injury
ii. ...might be killed
iii. ...may have even bigger part to play in society
iv. ...might a job like this help students
v. ...might have a ghost up there

6. Form full sentences out of the following words taken from CBS, CNN and ABC TV channels respectively, using ‘may / might’.

- That pipe/ leak / deadly carbon dioxide/ into your home
  ……………………………………………………………………….
- And /whatever /New York really /be /matter /less than/ what we believe it is.
  ……………………………………………………………………….
- But for many/ the right treatment/ mean/ they finally/ be /able /to recover.
  ……………………………………………………………………….
- I hope it/ not come /to that/ but /it ().
  ……………………………………………………………………….

HOMEWORK

Let’s try the following activity!
Fill in the following gaps with
- may / might
- the verb in the appropriate form

1) Steve: I like them like that! Still, I think you....(be) right about the smart shirt.
2) You …..(care) about environment at home, but what happens on holiday?
3) If you put some fruit in a glass bowl, you ….. (dream) you are wearing the glass bowl on your head.

4) They ……..(be) right, but I want to give it a go.

5) Mount Vesuvius …..(be) one of the most dangerous volcanoes in the world.

6.4 The control group’s learning sessions

The control group was taught the modal verbs during four learning sessions of 50-minute each. The methodology used was the one explicitly presented in the course book ‘Take Off! B1+’ (Hillside Press 2013) which had been chosen for this class at that time.

In ‘Take Off!’ the English Modal Verbs are presented in the 8th Unit, entitled ‘At the scene of Crime’. Each unit is consistent in its layout which includes eight distinct sections. The first section involves an introduction to the common topic on which each unit is based, usually with the form of a short text, relevant photos and a short activity. The second section targets reading comprehension, and the third one involves vocabulary learning. The fourth part of each unit focuses on one or more grammatical phenomena, while the fifth section is listening practice. The sixth part focuses on oral production with the seventh one being written production. The last section of each unit comprises revision activities aiming to consolidate the newly acquired vocabulary and grammar.

The grammar section of the eighth unit comprised seven activities. There were six additional tasks contained in the workbook which accompanies the ‘Take Off!’ course book, and two more in the revision section of unit 8 in the course book and workbook respectively, in total 17 activities. The teacher’s book offers explicit guidelines regarding the teaching methodology of the phenomena each unit focuses on. Before engaging in any grammar activities teachers are advised to guide their students into consulting the relevant grammatical rules appendix at the end of the book.

The first learning session with the control group focused thus solely on reading the rules regarding the formation and use of Modal Verbs which are located at the appendix of
the course book (Hillside Press 2013: 199) and discussing them in class. The teacher was the central figure in this process, as she was the one to present the rules, explain them and provide examples and clarifications. The students read excerpts from the appendix but apart from that they remained silent and were generally passive.

The next three sessions focused on the activities offered in both the course book and the workbook. Those activities included a variety of multiple choice, gap filling, matching and rephrasing tasks, all designed to provide them with practice and to test the knowledge they had received by studying the grammar appendix. The tasks did not focus on any modal verb in particular; each activity treated some or all of them at the same time.

The present chapter discussed in detail the DDL lessons designed for the experimental groups and the traditional methodology lesson followed by the control group. A comparison in the improvement of those two groups through appropriate pre and post-tests is to be presented in chapter 7.
CHAPTER 7

THE TEST RESULTS

7.1 Introduction

The present chapter discusses the results of the pre, immediate-post and delayed post-tests (see appendix E) whose design has been described in detail in chapter 5. The three tests were identical and were administered to the three classes of senior high school students who participated in the present case study. The students sat for the aforementioned tests before (pre-test) and after (immediate post-test/ delayed post-test) they had been taught the modal verbs in class, incorporating for the classes B1 and B2 (experimental groups 1 and 2) the data-driven learning (DDL) intervention, and for class B3 (control group) the traditional course-based method. The data collection time is presented in Table 7.1.

Table 7.1: The data collection timeline

<table>
<thead>
<tr>
<th>Apr 2016</th>
<th>May 2016</th>
<th>Feb 2017</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pre-test</td>
<td>Immediate post-test</td>
<td>Delayed post-test</td>
</tr>
</tbody>
</table>

The three experimental tests contained in total 34 multiple choice questions, each one offering a selection of four lexical items, out of which 21 questions covered the main senses of the modal verbs taught in class. More specifically, 2 items covered the main senses of will, 3 items examined the senses of would, 7 those of could, 4 referred to should, 3 addressed the senses of may and 2 those of might (see Table 7.2). Each correct answer earned one point. Therefore, the maximum number of points each participant
could earn was 34 per test, but for the purposes of the present study we only took into account the answers provided to the 21 questions concerning the modal verbs.

Table 7.2: The senses of will, would, could, should, may, might in the pre, post and delayed post-experimental tests

<table>
<thead>
<tr>
<th>No of occurrences</th>
<th>Modal</th>
<th>Senses</th>
</tr>
</thead>
</table>
| 2                 | Will  | a. opinion  
                      b. Request |
| 3                 | Would | a. 2nd conditional  
                      b. invitation  
                      c. past of will in indirect speech |
| 7                 | Could | a. polite request  
                      b. couldn't (negative deduction – past)  
                      c. ability based on a certain condition  
                      d. suggestion  
                      e. strong possibility  
                      f. past ability  
                      g. ability that was not used in the past |
| 4                 | Should| a. obligation  
                      b. advice – present  
                      c. advice – past  
                      d. expectation |
| 3                 | May   | a. weak possibility about a present condition  
                      b. weak possibility about a future action  
                      c. permission |
| 2                 | Might | a. weak possibility – past  
                      b. past of indirect speech |
a. ‘Does a DDL learning methodology positively affect the learning of Modal Verbs by Greek Senior High school students?’, and
b. ‘Is any improvement in the competence of the learners who have been taught the Modal Verbs following a DDL methodology statistically significant?’

7.2 Results with number right scoring
Initially, the learners’ performance was assessed using simple counting or to borrow a term from Lesage et al (2013:188) “number right (NR) scoring”\textsuperscript{16}. After gathering the data provided by the three tests, the calculation of the mean scores (see table 7.3) revealed that the mean (M) of the pre-test scores of the experimental group 1 (M= 14.13) and of the experimental group 2 (M=13.70) were lower than that of the control group (M= 14.45).

Another finding is that the two experimental groups improved their scores in the immediate post-test; B1 class raised their score to post-test mean score (M= 15.08) and class B2 to post-test mean (M= 15.15). On the contrary, the control group showed a slight decrease in its performance and that is reflected in the post-test mean score (M= 14.30).

All three groups’ scores in the delayed post-test are higher than in the pre-test. The experimental group 1 and the control group achieved higher scores in the delayed post-test as compared to the immediate post-test; experimental group 2 scored lower on the delayed post-test than on the immediate post-test.

Table 7.3: Mean scores of Pre, Immediate Post and Delayed Post-tests in all three groups using simple counting

<table>
<thead>
<tr>
<th></th>
<th>Pre-test</th>
<th>Immediate post-test</th>
<th>Delayed post-test</th>
</tr>
</thead>
<tbody>
<tr>
<td>Experimental group 1 (B1)</td>
<td>14.13</td>
<td>15.08</td>
<td>15.83</td>
</tr>
<tr>
<td>Experimental group 2 (B2)</td>
<td>13.70</td>
<td>15.15</td>
<td>14.65</td>
</tr>
<tr>
<td>Control Group (B3)</td>
<td>14.45</td>
<td>14.30</td>
<td>14.75</td>
</tr>
</tbody>
</table>

\textsuperscript{16}For a detailed presentation of each students’ scores see appendix E.1
7.3 Factors affecting students’ performance in the tests

The actual knowledge of the senses and grammatical properties of Modal Verbs may not be the only factor affecting each student’s answer in every item of the multiple-choice test. Other reasons rather than actual knowledge may be at work, something that is justified by the fact that there were students who received better scores in the pre-test rather than in the post-test regardless of the teaching method followed.

Classroom observation indicates that the students’ performance could have been affected by psychological reasons; for instance, the fact that they had to re-sit the test caused some complaints and may have led to unwillingness to take it seriously.

Also, due to school timetable limitations, the examinees would sit for the pre, immediate post and delayed post-tests during inconvenient times, for example early in the morning, or early afternoon, or right before or after school performance rehearsals. Furthermore, one should not overlook the possibility of a student being sick, absent minded, sleepy or bored and thus unable to focus on the test. Previous studies on the field have also stressed other factors affecting test performance; Lesage et al (2013: 189) suggest risk-taking behavior and test anxiety, something that is also mentioned in Liao (1999), while Chan and Wu (2004: 287) focus on ‘foreign language anxiety’.

Last but not least, we have to bear in mind the randomly made choices. The ‘lucky guess’ factor is one that we believe can be taken into account when assessing the findings of such tests. The students were told to provide an answer for all questions whether they were confident about the answer or not.

7.4 Minimizing the ‘lucky guesses’ factor

7.4.1 First normalization correction

As mentioned in the previous paragraph, multiple choice test results may contain irregularities due to blind guessing; however, this can be corrected. The rationale is actually quite simple; if a student has provided ‘a’ wrong answers we assume that these wrong answers constitute $\frac{3}{4}$ of the total number of questions they answered by chance, since it is only natural that if a student made a lucky guess, (s)he had $\frac{1}{4}$ chances to give the correct answer. Therefore, the total number of answers ‘x’ provided by chance are
3/4x=a wrong and 1/4x correct. Those correct answers cannot be considered proof of knowledge and they should be deducted from the total number of correct answers. The above-mentioned rationale, led to the calculation of a normalization of the tests’ scores illustrated by the formula \[ NS = R - W^{\frac{1}{3}} \] 17.

This has been taken into account in our analysis of the results of both pre and post test scores and can be seen in the ‘level of knowledge’ table and graphs. What we want to achieve is a more accurate estimation of the actual knowledge of each student. These observations agree with previous studies on multiple-choice tests scoring, in which it has been suggested that it is important to deal with the phenomenon of examinees producing results which may be sheer guesses due to their obligation not to leave any question in the test unanswered. This may produce, according to Reid (1977:335), an “upward bias” in the scores especially of poorer students as a result of guessing. Frary (1988) also stresses the need to reduce multiple-choice test score irregularities due to guessing. Both aforementioned studies provide negative (or, as we call it, ‘normalization’) scoring, using the formula \[ FS = \frac{R - W}{(C-1)} \] 18. It is thus obvious that the Formula Score by Frary (1988) agrees with the Normalized Score formula we designed for the purposes of the present study.

7.4.2 Second normalization correction

There were students who received better scores in the pre-test rather than the post –test regardless of the teaching method followed. In those cases, where a student’s score at the pre-test was higher or equal to that of the post test, we made a correction. The pre-test grade was corrected so that it equals the post-test grade since it is not possible that the student knowledge was lower in the post-test; we rather assumed that the lucky guess factor was the cause behind the higher pre-test score.

---

17 Where \( NS \) stands for ‘Normalized Score’, \( R \) for number of items answered right and \( W \) for number of items answered wrong

18 Where \( FS \) stands for ‘Corrected Score’, \( R \) for number of items answered right, \( W \) for number of items answered wrong and \( C \) for number of choices per item
Table 7.4: B1 class sample knowledge level score (after the normalization)

<table>
<thead>
<tr>
<th>Experimental Group 1 (B1) knowledge level</th>
<th>Pre-test</th>
<th>Immediate Post-test</th>
<th>Delayed Post-test</th>
</tr>
</thead>
<tbody>
<tr>
<td>(minimizing luck factor)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>11.69</td>
<td>14.35</td>
<td>14.35</td>
</tr>
<tr>
<td>2</td>
<td>11.69</td>
<td>17.01</td>
<td>15.68</td>
</tr>
<tr>
<td>3</td>
<td>6.37</td>
<td>6.37</td>
<td>6.37</td>
</tr>
<tr>
<td>4</td>
<td>18.34</td>
<td>18.34</td>
<td>15.68</td>
</tr>
<tr>
<td>5</td>
<td>11.69</td>
<td>19.67</td>
<td>17.01</td>
</tr>
</tbody>
</table>

Figure 7.1: Distribution of original and normalized scores for the experimental group 1 pre-test results
7.5  Descriptive statistics for pre, immediate post and delayed post-tests

7.5.1  The test scores

In the statistical analysis of the pre- and post-test results, we calculated each student’s mean score (μ) for each one of the three tests (s)he sat. Furthermore, the population variance σ² was calculated using the mathematical formula \( \sigma^2 = \frac{1}{N} \sum (X_i - \mu)^2 \), where \( N \) is the number of students and \( \mu \) the mean ‘knowledge level’ score. The square root of \( \sigma^2 \) called standard deviation (σ or s), which shows the dispersion of deviations of a group of scores from the average, was also calculated. Standard deviation measurement was very useful for the present study, since it showed how much each ‘knowledge level’ score of our population deviates from the average knowledge level of each group and thus we were able to compare the variance between two groups of observations.

The results from each group’s pre, immediate post and delayed post test scores are shown in tables 7.5, 7.6 and 7.7. The examination of the mean score shows that the group which originally had the lowest ‘level of knowledge’ was the experimental group 2 (B2). Surprisingly though, this group also had the highest increase in score in the immediate post-test. Another observation is that both the experimental and control groups showed a growth in the delayed post-test scores as compared to the pre-test, however these scores are lower than those of the immediate post-test.

Table 7.5:  Descriptive statistics for experimental group 1 (B1)

<table>
<thead>
<tr>
<th></th>
<th>N</th>
<th>Range</th>
<th>Minimum</th>
<th>Maximum</th>
<th>Mean</th>
<th>Std Deviation</th>
<th>Variance</th>
</tr>
</thead>
<tbody>
<tr>
<td>B1pre</td>
<td>24</td>
<td>14,63</td>
<td>3,71</td>
<td>18,34</td>
<td>11,0250</td>
<td>3,69997</td>
<td>13,800</td>
</tr>
<tr>
<td>B1imppost</td>
<td>24</td>
<td>13,36</td>
<td>6,37</td>
<td>19,67</td>
<td>13,1308</td>
<td>4,07424</td>
<td>16,599</td>
</tr>
<tr>
<td>B1depst</td>
<td>24</td>
<td>10,64</td>
<td>6,37</td>
<td>17,91</td>
<td>12,5767</td>
<td>3,75499</td>
<td>14,100</td>
</tr>
<tr>
<td>Valid N (listwise)</td>
<td>24</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Figure 7.2: Experimental group 1 (class B1) scores

Table 7.6: Descriptive statistics for experimental group 2 (B2)

<table>
<thead>
<tr>
<th></th>
<th>N</th>
<th>Range</th>
<th>Minimum</th>
<th>Maximum</th>
<th>Mean</th>
<th>Std. Deviation</th>
<th>Variance</th>
</tr>
</thead>
<tbody>
<tr>
<td>B2pre</td>
<td>20</td>
<td>15.96</td>
<td>3.71</td>
<td>19.67</td>
<td>10.22</td>
<td>5.06725</td>
<td>25.677</td>
</tr>
<tr>
<td>B2immpost</td>
<td>20</td>
<td>17.29</td>
<td>3.71</td>
<td>21.00</td>
<td>13.21</td>
<td>6.20098</td>
<td>27.050</td>
</tr>
<tr>
<td>Valid N (listwise)</td>
<td>20</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Figure 7.3: Experimental group 2 (class B2) scores

Table 7.7: Descriptive statistics for control group (B3)

<table>
<thead>
<tr>
<th></th>
<th>N</th>
<th>Range</th>
<th>Minimum</th>
<th>Maximum</th>
<th>Mean</th>
<th>Std. Deviation</th>
<th>Variance</th>
</tr>
</thead>
<tbody>
<tr>
<td>B3pre</td>
<td>20</td>
<td>15.96</td>
<td>1.05</td>
<td>17.01</td>
<td>10.6255</td>
<td>4.44842</td>
<td>19.708</td>
</tr>
<tr>
<td>B3immepost</td>
<td>20</td>
<td>17.29</td>
<td>1.05</td>
<td>18.34</td>
<td>12.0890</td>
<td>5.08925</td>
<td>25.900</td>
</tr>
<tr>
<td>B3delpost</td>
<td>20</td>
<td>17.29</td>
<td>1.05</td>
<td>18.34</td>
<td>11.2245</td>
<td>4.65300</td>
<td>21.650</td>
</tr>
<tr>
<td>Valid N (listwise)</td>
<td>20</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
At this point, it is important to note that in the case of our study, the standard deviation values of the pre- and post-test scores of all three groups of students (experimental groups 1, 2 and control group) do not share the same average ‘knowledge level’ score. Therefore, it was deemed necessary that we calculated the relative standard deviation ($RSD$ or $CV$), that is $\sigma/\mu \Rightarrow CV=S/X$, so as to have values comparable across all three groups (see Table 7.8).

Table 7.8: Relative standard deviation ($CV$) values for the pre and immediate post-tests

<table>
<thead>
<tr>
<th></th>
<th>Pre-test</th>
<th>Immediate Post-test</th>
<th>Delayed Post-test</th>
</tr>
</thead>
<tbody>
<tr>
<td>B1 Exp. group 1</td>
<td>0.33</td>
<td>0.30</td>
<td>0.36</td>
</tr>
<tr>
<td>B2 Exp. group 2</td>
<td>0.48</td>
<td>0.38</td>
<td>0.54</td>
</tr>
</tbody>
</table>
The calculation of the relative standard deviation (CV) was a useful tool for us to calculate, first of all, how homogeneous each group initially was, regarding knowledge level. Indeed, the calculation of CV (see Figure 7.5) shows that all three groups had similar CV values and more interesting is the fact that Control group rates exactly between experimental group 1 and experimental group 2. That provides an excellent basis to assess the post-test results, and in this way the success or failure of each methodology.

Moreover, we were able to reach a conclusion regarding how much the relative standard deviation of each group’s scores increased or reduced after the DDL or the traditional teaching method. In other words, an answer was provided to the question ‘did the class respond to the teaching method followed or were there some students that improved their knowledge of modal verbs and others that did not respond positively?’
In this case, the calculation of relative standard deviation (CV) in the immediate and delayed post-tests (see Figure 7.6) reveals that the relative standard deviation average of both classes that followed the DDL method was reduced. This shows that students who belonged to the two experimental groups had more or less the same progress, they responded to the DDL method with homogeneity, whereas in the control group the standard deviation increased in the post-test results. This fact may lead us to reach the conclusion that some students responded to the traditional methodology and improved their scores but they were many others who failed to do so.

Figure 7.6: Immediate post-test knowledge level for all groups

Less homogenous is the students’ score in the delayed post-test. Here, we note a slight raise in the relative standard deviation values, as far as the experimental groups are concerned, whereas there is no change in the CV value for the control group (see Figure 7.7).
7.5.2 Analysis of the score increase between pre and post-tests

The next set of results relates to the improvement (if any) in the students’ performance in the post test, by summing the extra points gained in the post test. It has been previously mentioned that the total points awarded in each test for all modal verb-related items are 21. In the present study we added the additional points gained by each student in the post test. In other words, we have investigated how great the improvement in the test performance is among those students who showed improvement in the post test. Here, we have to note that the three groups had different population. Therefore, in order to have comparable results we adopted a reference population of 100 students for all three groups.
Figure 7.8: Sum of points difference between pre-test and post-test scores of students listed according to their pre-test scores

Figure 7.9: Sum of points difference between pre-test and post-test scores of students listed according to their post-test scores
What becomes evident from the data provided in Figures 7.8 and 7.9 is that there is a growth in the knowledge level of the students in the two experimental groups, a fact underlined by the high curves of the test scores’ sum of points for groups experimental 1 and experimental 2. More interesting is the fact that the DDL method was clearly very helpful to students with low pre-test knowledge level, as can be clearly seen in Figure 7.8. Here, the experimental groups’ improvement curves are definitely more dramatic as compared to the control group. More specifically, as regards experimental group 1, there has been an increase in the scores of the students who had got around 9-12 correct items in the pre test, whereas regarding experimental group 2, the DDL methodology had a marked impact on the post-test performance of students who had performed even more poorly at the pre test.

The comparison of the data illustrated in Figures 7.8 and 7.9, reveals similar tendencies in the pre/post test performance of the two experimental groups and a completely different behaviour in the control group. The latter shows almost identical curves in those two figures. The gain scores of the control group are definitely lower and the traditional methodology seems to favour students who have already mastered the use of modal verbs to a certain degree, since the curve peaks at the score of 12-15 out of a total of 21 points.

7.6 Analysis of the statistical significance

After the scores between the three groups’ results in the pre, immediate post and delayed post-test were calculated, a statistical analysis of the significance of the results was calculated. Statistical significance reflects how far a given association exceeds those which would be expected by chance. More specifically, a paired t-test to each pre- and immediate post, pre and delayed- post and immediate and delayed post-test score for the two experimental groups and the control group was applied. The merits of a t-test as a statistical technique used to determine if a significant difference exists between two groups have been often stated in previous studies such as in Nunan (1992) and Biber et al (1998). Furthermore, a one-way analysis of variance (ANOVA) was also conducted between the pretest, immediate posttest and delayed posttests of the two experimental groups and the control group.
7.6.1 Within-group statistical analysis - paired samples tests

The aim of the paired t-test was to confirm whether the results presented in paragraph 7.5 are statistically significant. The null hypothesis was that ‘there is no statistical significance in either the raise or drop in score results between the pre- and immediate post, pre and delayed-post and immediate and delayed post-tests respectively’.

More specifically, in the case of Experimental group 1 (B1) the students made significant gains where pre and immediate post and pre and delayed post-tests are paired; in the former case, with a t value of -3.616 which is statistically significant since \( p < 0.05 \), and in the latter, with a t value of -3.077 which is statistically significant since \( p = 0.05 \). However, there was no statistical significant difference between the B1 group’s results in immediate post and delayed post-test (\( t = 2.632 \), \( p > 0.05 \). Therefore, we fail to reject the null hypothesis that there is no significance in either the increase or drop in score results between the immediate and delayed post-test.

Table 7.9: Paired Samples Test for the experimental group 1

<table>
<thead>
<tr>
<th>Paired Samples Test</th>
<th>Paired Differences</th>
<th>95% Confidence Interval of the Difference</th>
<th>t</th>
<th>df</th>
<th>Sig (2-tailed)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean</td>
<td>Std. Deviation</td>
<td>Std. Error Mean</td>
<td>Lower</td>
<td>Upper</td>
<td>t</td>
</tr>
<tr>
<td>Pair 1 B1pre-B1immediatepost</td>
<td>-2.10693</td>
<td>2.65288</td>
<td>.90236</td>
<td>-3.01054</td>
<td>-3.01054</td>
</tr>
<tr>
<td>Pair 2 B1pre-B1delayedpost</td>
<td>-1.55167</td>
<td>2.47011</td>
<td>.59421</td>
<td>-3.59470</td>
<td>-3.59470</td>
</tr>
<tr>
<td>Pair 3 B1immediatepost-B1delayedpost</td>
<td>.04417</td>
<td>1.03106</td>
<td>.21055</td>
<td>.19802</td>
<td>.19802</td>
</tr>
</tbody>
</table>

As far as the experimental group 2 (B2) test score is concerned, the results are presented in Table 7.10. Regarding the pre and immediate post-test, with \( t(19)= -3.707, p = .001 \) the mean difference was significant. That can also be claimed in the case of immediate post- and delayed post-test, with \( t(19)= 3.358, p = .003 \). On the other hand, concerning the paired t-test for B2 group’s immediate post and delayed post-test, the \( p \) value exceeds the significance level of 0.05. This result shows that the mean difference between immediate post and delayed post-test is not significant.
In the case of the Control group (class B3), the interpretation of the statistical analysis showed that there is a statistical significance in the mean difference between pre- and immediate post-test, with $t(19) = -3.707, p = .001$. On the other hand, the null hypothesis was confirmed in the comparison of results for the pre- and delayed post-test and the immediate and the delayed post-test, the former having $t(19) = -1.371, p = .186$ and the latter resulting in $t(19) = 2.942, p = .008$.

### Table 7.11: Paired Samples Test for the control group

<table>
<thead>
<tr>
<th>Paired Differences</th>
<th>Mean</th>
<th>Std Deviation</th>
<th>Std Error</th>
<th>95% Confidence Interval of the Difference</th>
<th>Lower</th>
<th>Upper</th>
<th>t</th>
<th>df</th>
<th>Sig. (2-tailed)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pair 1 B3pre - B3impost</td>
<td>$-1.2030$</td>
<td>$1.5230$</td>
<td>$0.4039$</td>
<td>$-1.9760$ to $-0.4400$</td>
<td>$-1.9760$</td>
<td>$-0.4400$</td>
<td>$19$</td>
<td>$0.001$</td>
<td></td>
</tr>
<tr>
<td>Pair 2 B3pre - B3delaypost</td>
<td>$-3.9091$</td>
<td>$1.3017$</td>
<td>$0.2107$</td>
<td>$-4.3621$ to $-3.4561$</td>
<td>$-4.3621$</td>
<td>$-3.4561$</td>
<td>$19$</td>
<td>$0.000$</td>
<td></td>
</tr>
<tr>
<td>Pair 3 B3impost - B3delaypost</td>
<td>$0.6458$</td>
<td>$1.3141$</td>
<td>$0.2485$</td>
<td>$0.2246$ to $1.0670$</td>
<td>$0.2246$</td>
<td>$1.0670$</td>
<td>$19$</td>
<td>$0.000$</td>
<td></td>
</tr>
</tbody>
</table>

#### 7.6.2 Between-groups statistical analysis - ANOVA

A one-way analysis of variance (ANOVA) was calculated on the pretest scores of all students belonging to the experimental groups 1 and 2 as well as the control group. ANOVA allows one to determine whether the differences between the samples are simply due to random error or whether there are systematic treatment effects that cause the mean in one group to differ from the mean in another.
The ANOVA was conducted to compare the effect of existing knowledge of modal verbs on the pretests taken by the students in experimental groups 1 and 2 and the control group. The analysis was not significant, $F(2,61) = 0.189, p = .828$. Comparisons indicated that the experimental group’s 1 pretest scores were not significantly different from those of the experimental group 2 and the control group.

Table 7.12: ANOVA for the pre-test

<table>
<thead>
<tr>
<th></th>
<th>Sum of Squares</th>
<th>df</th>
<th>Mean Square</th>
<th>F</th>
<th>Sig</th>
</tr>
</thead>
<tbody>
<tr>
<td>Between Groups</td>
<td>7,313</td>
<td>2</td>
<td>3,657</td>
<td>.188</td>
<td>.828</td>
</tr>
<tr>
<td>Within Groups</td>
<td>1176,707</td>
<td>61</td>
<td>19,323</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>1186,020</td>
<td>63</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Regarding the immediate posttest results, the one-way ANOVA determined that there were no statistically significant differences between group means ($F(2,61) = 0.356, p = .702$).

Table 7.13: ANOVA for the immediate post-test

<table>
<thead>
<tr>
<th></th>
<th>Sum of Squares</th>
<th>df</th>
<th>Mean Square</th>
<th>F</th>
<th>Sig</th>
</tr>
</thead>
<tbody>
<tr>
<td>Between Groups</td>
<td>16,187</td>
<td>2</td>
<td>8,094</td>
<td>.356</td>
<td>.702</td>
</tr>
<tr>
<td>Within Groups</td>
<td>1387,849</td>
<td>61</td>
<td>22,752</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>1404,037</td>
<td>63</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

A one-way analysis of variance was also calculated on the delayed posttest scores of the students belonging to the experimental groups 1 and 2 and the control group. Again, the analysis was not significant $F(2,61) = 0.801, p = .454$. 143
Table 7.14: ANOVA for the delayed post-test

<table>
<thead>
<tr>
<th></th>
<th>Sum of Squares</th>
<th>df</th>
<th>Mean Square</th>
<th>F</th>
<th>Sig.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Between Groups</td>
<td>28,835</td>
<td>2</td>
<td>14,417</td>
<td>0.801</td>
<td>0.454</td>
</tr>
<tr>
<td>Within Groups</td>
<td>1097,927</td>
<td>61</td>
<td>17,999</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>1126,762</td>
<td>63</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

7.7 Analysis of the test results for each modal verb

A further statistical analysis was conducted, this time focusing on the scores the students of the two experimental groups and the control group achieved on each of the modal verbs included in the test. The items testing the various senses of each modal were 21, but the same modals were used as distractors in various items. Our analysis was twofold; firstly, it focused on the success rate of each modal and secondly it calculated the relevance of choice which is the number of times each modal was successfully chosen.

The success rate for each modal verb of the students who belong in the experimental group one is presented in figure 7.10. Here, we can see that the group’s average score was raised in the immediate post and the delayed posttest. Moreover the score received for each modal in both the immediate and the delayed posttests was higher than that of the pretest, with ‘may’ being the only exception; in the case of ‘may’, the students dropped their scores in the posttest, but in the delayed posttest the scores are higher than those of the pre and the immediate posttest. The students gave the most successful answers in the items testing the knowledge of ‘would’; the items testing ‘might’ received the least successful answers. The statistical analysis’ results regarding the relevance of choice, in other words the number of times each modal verb was selected, successfully or not, were equally revealing. The participants improved their successful hits in almost all modal verbs, in the immediate post and delayed post tests. The only exception is that of ‘will’ which seems to be over selected.
Figure 7.10: Experimental group 1 (class B1) success rate in every modal verb

Figure 7.11: Experimental group 1 (class B1) successful hits for every modal verb
As far as experimental group 2 is concerned, the results or the success rate for the six modal verbs, illustrated in figure 7.12, clearly show that students’ choices were more successful in the immediate post (69%) and delayed post-test (70.79%) rather than in the pre-test (64.00% success rate).

Moreover, the learners seemed to have comprehended the senses of ‘would’ better than the other modal verbs, whereas ‘might’ seems to rate below average in all three tests.

Figure 7.12: Experimental group 2 (class B2) success rate in every modal verb

The statistical analysis’ results of the times each modal verb was selected successfully or not, yielded further interesting results. Figure 7.13 shows that ‘will’ and ‘would’ tended to be rather over-selected. This finding is similar to the data obtained by the analysis of the successful hits made by the students who belonged in the experimental group 1. On the other hand, ‘could’ and ‘may’ had definitely more successful hits, and this is again in accordance with the behavior demonstrated by experimental group 1 students.
If one compares the results illustrated in Figures 7.12 and 7.13, they can clearly see that the average successful response in each test item and the average relevant selection of the same modal verbs seem to match, since the percentages are very close for the pre and immediate post-test. What is interesting though, is the definite increase in the average calculated for the successful hits made for each modal verb in the delayed post-test (73.26% as compared to 70.29% in the case of successful response rating).

Figure 7.13: Experimental group 2 (class B2) successful hits for every modal verb

In the case of the control group, the statistical analysis of the test results for each modal verb (see Figures 7.14 and 7.15) offered us some very interesting findings. The students’ success rate in the immediate posttest is only slightly higher than that of the pretest (65.59% in the immediate posttest as compared to 65.35% in the pretest), while there is an improvement of the scores in the delayed posttest (69.24%).
Figure 7.14: Control group (class B3) success rate in every modal verb

Figure 7.15: Control group (class B3) successful hits for every modal verb
The two experimental groups’ tendency to have mastered ‘would’ best is also evident in the control group. Likewise, the lowest success rate can be traced in the results for the modal verb ‘might’. It seems thus that the students in all three groups have understood the senses of ‘would’, and to a lesser degree those of ‘could’, ‘should’ and ‘may’; on the other hand they face difficulties in mastering the use of ‘might’.

The statistical analysis of the relevance of choice in the various items included in the test by the students of the control group was the source of one of the most interesting findings of the present research. The control group students seem to have made the most relevant choices in the case of modal verbs ‘could’ and ‘may’, a fact that closely resembles the results of experimental group 2.

However, what differentiates the control group from the two experimental groups is that here the average successful hits score seems to have dropped in the immediate posttest (with a 65.99% successful hits rate in the immediate posttest and opposed to the 67.36% successful hits rate in the pretest). In other words, regarding the immediate posttest, the students of the control group seem not to aim so well in their answers; for example they tend to over select certain modal verbs and avoid others.

7.8 Conclusion

Chapter 7 focused on the data that was collected by the students’ scores in the experimental pre, immediate post and delayed post-tests. The calculation of score means showed that there is a definite improvement in the knowledge level of all three groups, even though the means were higher in the case of the two experimental groups. Paired samples analysis of the results revealed that there is significant difference in the pre- and immediate post-test scores of all three groups, as well as between the immediate and delayed post-test scores of all three groups; there is significant difference between pre- and delayed post-test scores only in the case of experimental group 1. Regarding the between groups statistical analysis, one-way analysis of variance revealed that the difference in the score means obtained by the students in the pre-test, immediate post-test and delayed post-test are not statistically significant. The hard evidence which came to light after the relevant tables and graphs were created,
will be examined in detail in Chapter 8 and relevant conclusions will be drawn which will support the research hypotheses set by the present study.
CHAPTER 8

DISCUSSION OF THE RESULTS

8.1 Introduction

The following chapter will discuss the findings of the study. The data gathered from the pre-test, the immediate post-test and the delayed post-test was analyzed using the statistical measurements of Distribution, Standard Deviation, Relative Standard Deviation, paired-samples t-test and one-way analysis of variance (ANOVA). This analysis was conducted using the Microsoft Office Excel and the Statistical Package for Social Science (SPSS) version 20. The data and analytics presented in the previous chapter will be leveraged to make informed decisions. The results of the results and performances in this study are discussed in the order of the research questions presented in chapter 5. This chapter also discusses the learners’ attitudes towards DDL, based on in-class observations and informal interviews throughout the duration of the project.

8.2 Discussion of the results

8.2.1 Research question 1

“Can a DDL teaching methodology be applied – either directly or indirectly - to a Greek high school classroom setting?”
The answer to the first research question is affirmative, as shown by the results of our experimental study. Our study first and foremost revealed that although the existing educational framework in Greek state schools tends to favor a traditional course book-based language learning methodology, the introduction of a corpus-based (or corpus-driven) DDL approach is something implicitly encouraged by the students themselves, because of their familiarity with computers in general and the adolescents’ inquisitive nature and their own intuition to search for meaning in the context. This was a conclusion we reached by interpreting the students’ responses in the computer-related questionnaire and their attitudes towards grammar learning as these were expressed in the relevant questionnaire. The data collected from the administration of the aforementioned questionnaires showed the students hold generally favorable perceptions towards an approach to learning that is more technologically-friendly and allows them a degree of autonomy by encouraging discovery learning rather than providing them with rules.

Furthermore, one cannot overlook the fact that a corpus-based methodology is implicitly recommended by the national curriculum for EFL teaching provided by the Greek ministry of education, since the lexical items and functions to be mastered are presented within examples of communication instances, rather than as solitary units.

Regarding the actual application of a DDL methodology, our study proved that, given the appropriate material, it could be both indirect and direct. In fact, ours is a combined inductive and deductive approach, the theoretical underpinning of which has been discussed in chapter 3. As we discussed in chapter 5, such a dual approach is mostly beneficial, since the indirect DDL, which makes use of printed concordances and handouts with relevant activities, under the clearly defined guidelines provided by the teacher, is more appropriate for training the students on how to deal with relevant activities and corpora, eliminating any fears or reservations they may have as much as possible. On the other hand, a direct or inductive, DDL approach, in which the student queries online corpora and makes use of appropriate concordancing tools, thus benefiting fully from the exposure to and investigation of authentic language, is also applicable, even in medium-to-low level classes such as ours.

In fact, this dual, or blended as we have previously called it, methodology was most appropriate for the Greek high school environment for practical reasons; the school’s
computer laboratory is not always available to the foreign language teacher and internet access is not always reliable. Moreover, shortage of available funding renders the use of commercial concordancers by students impossible. In order to account for this kind of problems, also mentioned by scholars such as Boulton (2010) and Warren (2016), the practitioner must improvise and search for the optimum way to provide his/ her students with access to corpora and relevant activities.

Here, we have to note that this was one of the advantages of the pilot sessions; not only did they provide valuable insight to the target population’s reaction towards the novel teaching approach, and the shortcomings in lesson planning and the design of the DDL activities, but they also tested the strengths and limitations of the setting in which our empirical study would take place, a typical state senior high school. In this way we modified the content of the DDL lessons appropriately, taking into account such parameters as school time-table limitations and availability of equipment, in order for them to be realistically applicable in the Greek high school environment.

As was extensively discussed in paragraph 5.4.2, any attempt to apply a new learning methodology, albeit an attractive one, cannot be successful without the appropriate training, a fact also mentioned in Sripicharn (2010). In the case of the present study, teacher training was redundant, since the teacher was also the researcher and designer of the DDL course. In the case of student training, the three DDL-designed learning sessions provided our learner target population with necessary skills in order to interpret corpus data. The benefits of the training sessions were twofold; first and foremost they helped the learners overcome the fear of the uncharted area of corpus- aided learning, by providing them with a “gentle way in” (Boulton 2001: 571) through relatively easy tasks in which they had to take one step at a time. Secondly, they offered the students practical skills so as to be able to read concordances and examine the linguistic structures contained in them with the aim to make their own pedagogically useful generalizations. Training sessions were the factor which triggered the student’s ability to notice and form hypotheses with regards to the linguistic items they came across in each step of the learning process. In other words, in order to become ‘language researchers’, the ultimate goal of any DDL approach, they first acted as ‘apprentice researchers’, and later on, during the DDL lessons they became ‘junior researchers’ aided by the guidance of the teacher.
8.2.2 Research question 2

“How can materials and methodology be appropriately adapted so as to have a data-driven approach to EFL teaching?”

In order for any indirect or direct DDL methodology to be applicable in a state school class in Greece, it is crucial that any material, corpora and activities be carefully adapted. This prerequisite is set by the demand to conform to the guidelines of the national curriculum. However it is a claim also highlighted by scholars abroad, such as in Braun (2007) who states that any material “needs to fit in with the overall language curriculum” (p.2). This was taken into account in the compilation of our pedagogic corpus, the design of the activities, as well as the selection of concordances from COCA which were appropriately adapted to match the points covered by the textbook. The requirement to comply with the syllabus, set by the Ministry of Education’s guidelines for the teaching of English in Senior High schools, serves another purpose too. It is important that both the experimental and the control groups focus on learning the same senses of modal verbs, regardless of the methodology, so as to get comparable results of their performance; this ‘curricular integration’ (Braun 2010) ensures the comparability of the results.

Our study proved that it is feasible to adapt the standard language learning procedure in Senior High school in order to make it data-driven, so as to comply with the directives set by the Ministry of Education and, at the same time, incorporate the principles which govern DDL. This was achieved by adopting a blended DDL approach, which involved compiling a pedagogic corpus (ECCo) which consisted of textbooks approved for use in class by the Ministry of Education and Religious Affairs and also using the Web as Corpus (WaC), as well as exploiting COCA, an online free corpus.

More specifically, the process of creating a data-driven learning environment in our research involved three steps; the students making use of the Web as a corpus, the teacher and the learners using existing online corpus resources and the creation of a custom made (“ad hoc”) pedagogic corpus and its exploitation for the design and administration in class of useful learning material.

As far as the WaC approach is concerned, the findings of this study were in line with the observation by Robb (2003) that linguists tend to use the web as a ‘quick and dirty’ source of language data for everyday concerns, while, at the same time, carefully
compiled corpora attempt to create a more principled and orderly bank of texts for linguistic analysis and, in our case, for effective EFL learning. Lastly, the use of a free online corpus with a friendly user interface (COCA), ensured the ‘pure’ hands-on DDL application.

Careful design of both the learning activities and the DDL lessons as a whole was also deemed crucial for the successful application of a DDL methodology. Since our initial literature review (see chapter 3) revealed the scarcity of relevant studies in the field we could only design the DDL lessons the way any novel approach is designed; by trial and error. This provided an answer to the second part of the 2nd research question; the existing learning methodology can be adapted so as to become a DDL one with the running of one or more pilot sessions (discussed in chapter 5), which will test the methodology to be used and will provide us with activities that are both syllabus appropriate and student-friendly.

The pilot study involved learning different linguistic items, both vocabulary and grammar, following both a paper based and a pure hands-on DDL approach. The students’ responses to the tasks as well as their comments on this novel approach led us to alter and enrich the training sessions and the main study’s lessons with a greater variety of tasks, as well as more detailed guidelines.

Lastly the design of activities had to be in accordance with the DDL framework which replaced the old ‘presentation, practice, production’ paradigm with the –more inductive- “identification, classification, generalization” (Johns 1991: 29), or as other scholars (Carter and McCarthy 2006: 155) have named it “illustration, interaction, induction” paradigm.

The type of tasks was not totally alien to the ones found in traditional course books; the students are familiar with fill-in-the-gap, matching and multiple choice exercises. What was different though is the lesson plan design and the order of the above-mentioned tasks; the lesson began with the investigation of corpus-derived examples which illustrated linguistic features, then proceeded to formulate rules by exchanging opinions in class and, lastly, consolidating the newly-found knowledge through activities based on data taken from corpora.

Granath (2009) and Whiste (1999) have argued in favor of using printed-concordances rather than having the students conducting the queries themselves, based on the
opinions expressed by their students after the DDL treatment, as this could reduce the
danger of being ‘overwhelmed’ (Johns 1986) by the abundance of data. However, we
felt that if we only made use of printed handouts containing edited concordances we
would water down the ‘pure’ DDL approach to a degree that it might be unrecognizable
and eventually lose those unique characteristics that are its strength; the motivation it
offers the learner to act as a researcher, the degree of autonomy and the immersion into
real ‘raw’ data. This was a parameter that we took into consideration when designing
our DDL methodology, with the aim to achieve a balance between ‘purity’ and
‘feasibility’. The positive results of the post tests and the equally positive reactions of
the students during the administration of the DDL lessons, testify to the success of the
middle-way which incorporated both printed handouts of edited concordances and a
hard-core, hands on approach; we made it possible to offer the students a somewhat
familiar, non-intimidating learning experience without compromising the ‘data-
driveness’ of our approach.

Looking closely at the design of the DDL lessons, one easily notices that they actually
involve three steps: 1) awareness raising and noticing of collocations 2) formation of
generalizations and 3) practice and recycling, also mentioned in Moreno Jaén (2008:
232). What our empirical study lacked was the fourth step, learner’s production, in a
sufficient number of production activities, such as creative writing exercises. This was
due to time limitations, since every DDL lesson was restricted by the school timetable
which dictates that each hour lasts for 50 minutes. We should note here that the 50
minutes time span is in paper only, because several minutes are needed for the class to
gather and the computers in the PC lab to be activated. There was an initial thought that
every lesson would occupy two 50 min lessons. However, this thought was quickly
dismissed for the following two reasons. There was no possibility for the EFL teacher
to be with the same class for two consecutive hours, because the school timetable could
not be modified in order to accommodate this requirement; Furthermore, even if there
was such a possibility, the hours assigned for the DDL lessons would exceed those
assigned for the lessons delivered with the traditional course book methodology, and
we feel that, if that was the case, the results would have to be filtered and interpreted
taking into account this parameter as well.
8.2.3 Research question 3

“Is DDL more effective than traditional grammar teaching in learning the English modal verbs?”

a) “Does a DDL learning methodology positively affect the learning of modal verbs by Greek senior high school students?”

The calculation of the students’ scores in the pre, immediate post and delayed post-test using number right (NR) scoring revealed that the two experimental groups improved their mean scores in the immediate post-test; on the contrary, the control group showed a slight decline in its performance.

A second examination of the experimental post-test scores, calculating this time the Normalized Score formula we designed for the purposes of the present study in order to account for the lucky guess factor which is normally found in multiple choice tests, delivered similar results. Once more, the two experimental groups’ learners scored better in the post test than those belonging in the control group. These improved scores provided the hard evidence deemed necessary to reach the conclusion that the proposed DDL methodology positively affects the learning of modal verbs by Senior high school students. This finding is in agreement with previous research on the application of a DDL methodology in foreign language grammar learning, by such scholars as those listed in the meta-analysis of DDL empirical studies by Cobb and Boulton (2015).

The quantitative analysis of the pre- and post-tests’ results of the two experimental groups (classes B1 and B2) and the control group (class B3) also offered valuable insight as regards learners’ knowledge of each one of the modal verbs that were taught. Two values were calculated; the success rate for each one of the six modal verbs and the relevance of choice of each modal verb among all distractors offered in the four choices for each question item, in other words, the number of times each modal verb was selected, successfully or not. The analysis of the pre, immediate post and delayed post-test results led to the following observations:
i. All three groups raised their average success rate in the immediate post and
delayed posttests; the control group had the lowest growth in the immediate
posttest between the three groups,

ii. All three groups had a higher success rate in ‘would’ and the lowest in ‘might’,

iii. The two experimental groups showed an improvement in the relevance of
choice for each modal verb in both the immediate post and delayed posttests,
while the control group’s average successful hits dropped in the immediate
posttest.

iv. The calculation of relevance of choice in the immediate post-test showed that
the regarding all three groups learners’ choices were more successful in the case
of ‘could’, followed in the order of successful ‘hits’, by ‘may’, and the least
successful hits involved ‘will’.

The writer of the present study was also the EFL practitioner and thus useful
observations of the application of the DDL methodology in the EFL classroom were
made. Close observation of the work the students did on the DDL material revealed that
some students adapted faster to the DDL rationale and showed a superior ability to
notice and interpret patterns in texts and chunks of language, either in printed
concordance output or an online concordancer, a fact also mentioned in Marinov
(2016). In fact, despite the awkwardness of the first contact with the new methodology,
the students quickly adapted to it and showed promising progress. This may be
attributed to the use of Web and googling as a gentle initiation to corpus searching.
While it is true that none of the students were familiar with KWIC searches, or even
with the term ‘corpus’, they conducted their first searches for unknown terms with a
certain confidence, since they were already familiar with querying various search
ingines. In other words, the Web as Corpus approach was crucial in introducing DDL,
least in its ‘light’ version, not as a radical new technique but as a well-trodden path
and assisted the practitioner to ‘demystify’ corpus use (Gabrielatos 2005).

At this point, we have to point out another interesting fact of our empirical study.
During the first sessions of both the pilot and the training lessons, the steps which
involved querying the web for unknown words were outlined on the printed handouts
which were created by the teacher. For example, the students were asked to use the
google search engine, or a specific online dictionary, or a thesaurus. During those
sessions we were impressed by the students’ active participation in the learning process.
They did not follow blindly the activities dictated in the handout; they questioned the choice of the online tools mentioned in the activities, exchanged views and past experiences of searching online for unknown vocabulary and suggested alternative ways of querying the web.

One further remark concerning the effectiveness of the proposed DDL methodology is that the students participating in the DDL lessons were students who specialized in either language subjects or science subjects. This means that not all of them shared the same metalinguistic awareness. There have been past studies, for instance the one described in Dobrushina (2008), which only involved high school students specializing in linguistics. Ours may have been a less homogenous group, but we believe that our choice to set the usage English modal verbs as a learning goal and to keep the linguistic terminology to a minimum suited all students. Whether and how much students specializing in either linguistics or science benefit from the DDL approach is left for future research.

b) “In the case of the present study, was the DDL methodology more beneficial than the traditional grammar teaching methodology?”

Regarding solely the results of the present study, the answer to this research question was provided by the comparison of the pre, post and delayed posttests scores between the two experimental groups and the control group. The examination of the results started from square one, which is the pre-test. Incidentally, the group which had received the best scores in the pre-test was the control group. One might have thought that this could offer those students the leverage to score better in the post test as well, but the results show that was not the case. Even though there was an increase in the post-test scores in all three groups, the two experimental groups’ scores were higher than those of the control group.

The examination of the mean score in the pre-test results shows that the group which originally had the lowest ‘level of knowledge’ was the experimental group 2 (class B2). Surprisingly though, this group also had the highest growth in the immediate post-test score. This may lead to the conclusion that, as far as our study is concerned, DDL suits best low-level students. At this point we have to note that all students had come in
contact with modal verbs at some time in the past. The low knowledge level revealed in the pretest scores indicates that many students, especially those belonging in experimental group 2, had not been positively influenced by the lessons following the traditional methodology which they had taken during the previous years. This low achievement may contradict students’ answers regarding their attitude towards grammar learning in the questionnaire which was administered before the pilot sessions and DDL lessons took place. In this questionnaire, the majority of students had answered that they considered grammar learning important in order to be able to communicate in English. One may wonder then, if they were aware of the importance of grammar, why they had failed to reach a satisfactory level in grammar learning. A closer look at the same questionnaire, fully justifies the pre-test’s low scores and provides an answer to the aforementioned question; most students stated that they dislike learning a foreign language by studying grammar rules, even though they admitted the usefulness of knowing grammar rules. It is no wonder then that the traditional deductive learning methodology which introduces each lesson with rules dictated by the teacher failed to motivate them and fully take advantage of their skills and learning abilities.

The observation that the DDL approach possibly suits best low level students was consolidated by the calculation of Relative Standard Deviation (RSD). The initial calculation of RSD in the pre-test scores showed close values in all three groups which indicated that all groups were more or less equally homogeneous. This was a useful value because starting from that point we could later obtain clear results concerning the homogeneity of each group; would they act as a homogenous group and be equally affected by the treatment or would some students respond to it while others stay behind? The drop of RSD values in the two experimental groups’ post-test shows that the learners belonging to those two groups’ response to the treatment was homogenous. That is, despite the awkwardness of the first contact with a methodology in many aspects alien to what they had experienced in the past, the students benefited in equal terms. This could be interpreted as follows; on the one hand, the learners who had failed to benefit much from the traditional course-book based lessons they had attended in the past, were given the advantage of a fresh start, based on the advantage of working on a computer, a skill they all possessed. On the other hand, those students who had already grasped the use of modal verbs to a certain satisfactory degree in the past, also had to
start anew, dealing with novel methods and materials. In a sense, they also had to start from square one, and that was the catalyst which bridged the gap between pre-test high and low achievers. In other words, DDL was a much-needed change of tactics which proved to be fruitful.

Another observation derived from the post-test results is the following. While calculating the post-test scores, we did not only count the correct vs wrong answers, but also the total number of times a modal verb was chosen among the four answer choices following each multiple-choice question item. The data collected by the analysis of the post –test results reveals a tendency to overuse the modal verbs after the end of the DDL lessons. It is our firm belief that this is only natural since the knowledge was very recent and they tended to choose them instead of the other lexical items provided in the multiple-choice test.

Apart from the promising results in the calculation of the post-tests, other advantages of the DDL were also revealed. The participants in the DDL group, due to the nature of the learning tasks, dealt with the language learning activities and the sample language in ways that the other learners did not. They searched online corpora and printed texts for examples of language use and they discussed these examples with their classmates in an effort to discover patterns in language and solve problems. This confirms Johns' (1997) assumption that in a DDL environment learners act as ‘language detectives’ (p. 101).

Another observation made by the practitioner/ researcher is that while searching the online corpora and interpreting the concordance lines either online or in printed form, students often came across unknown vocabulary. They had to find the meaning of those unknown words or phrases in order to comprehend the concordance output and do the DDL activities. Therefore, it was highly likely for them to look up for words in a dictionary or discuss the meaning of unknown words with their classmates. In this way, our DDL methodology not only improved their grammar skills, but also acquired new vocabulary and furthered their reading competence, a fact that echoes the observation by Fischer – Starcke (2008: 281). Vocabulary expansion and heightened pattern awareness are benefits of DDL also mentioned in Gilquin and Granger (2010). This was evident in the present empirical study as well, and it was the result of actively engaging in the investigation of authentic language.
The findings of our study, regarding the prominence of the DDL approach over traditional methods agree with the results presented in most empirical case studies on DDL. This fact is illustrated in Cobb and Boulton’s (2015) meta-analysis of empirical DDL studies which stress that almost always the DDL group outperforms the control group. Only a few scholars have reported that the control group showed a greater improvement than the experimental one (Smart 2014, Hadley and Charles 2016) and even fewer empirical studies, such as the one by Crosthwaite (2017) have stated that the usefulness of DDL for grammar learning was less than that for vocabulary and the learning of phrases.

c) ‘Is any improvement in the competence of the learners who have been taught the Modal Verbs following a DDL methodology statistically significant? In contrast to past studies who reported that DDL groups’ performance wasn’t significantly improved (Boulton 2010, Tian 2005, Vyatkina 2016), the data collected by the pre-tests, post-tests and delayed post-tests of the present experimental study indicated the opposite. After getting the results of the pre-test, immediate post-test and delayed post-test scores, paired-samples t-tests were conducted with the use of SPSS 20 in order to examine whether there were significant differences in the improvement of the experimental groups and the control group as well. The study findings showed that there was a significant difference in the post-test scores of the two experimental groups and these are in agreement with the findings of previous studies, such as the ones presented in Luo (2016) and Paker and Özcan (2017). More specifically, the experimental group 1 showed significant results in the comparisons between pre-test and immediate post-test, immediate and delayed post-test and pre-test and delayed post-tests, while the experimental group 2 showed significant improvement in pre-test-immediate post-test and immediate post-test - delayed post-test comparisons.

This definite improvement in the language knowledge of the learners can be explained by what Yunus and Awab (2014: 92) described as “the power of contexts”. It is a fact that DDL allows greater opportunities for learners to observe recurring patterns in language use, as it is in essence a process of discovery which encourages exploration and exploitation of any available resources. This echoes the statement by Gilquin and Granger (2010) that one advantage of DDL is the motivation it offers to students to think actively about the grammar they were learning. Our experimental study’s
findings, regarding the definite benefits of employing corpora in the teaching of grammar are in agreement with scholars such as Meunier (2002: 137) who argues in favor of making use of the methods and tools of corpus linguistics in grammar teaching.

It is our belief that the statistical significance of the present study’s results provides solid evidence of the usefulness of DDL and could persuade foreign language teachers to adopt either a pure or a weaker form of DDL in classroom. The choice of the degree of ‘purity’ of DDL will depend on the learners’ age, their level of knowledge of the target language, their familiarization with corpora and KWIC searches and last, but certainly not least, with the availability of computers, internet access and relevant software in the school computer laboratory.

A paired-samples t-test was also conducted for the control group’s scores in the pre, immediate post and delayed posttests. The findings show that this group’s improvement was also statistically significant in the case of pretest - immediate posttest, and immediate posttest- delayed posttest, but not in the case of pretest – delayed posttest.

Another finding of the present study is that there were no statistically significant differences between the experimental 1, experimental 2 and control group means in the pre-experimental test as determined by the one-way ANOVA. The ANOVA test results indicate that before the teaching of modal verbs there was not any difference among the three groups. ANOVA tests were also conducted to compare the three groups’ immediate post-test and delayed post-test scores. Again, tests showed no statistical significance in the student’s scores in terms of the different teaching methodology they had been subjected to.

A limitation of the present empirical study is the lack of evaluative feedback which would provide additional evidence regarding the effectiveness of our approach. Such feedback is suggested in several studies; Braun (2007) and Moreno Jaén (2008) used evaluation questionnaires to measure the learners’ attitudes towards DDL, Vyatkina (2016) has proposed the use of learner open-ended commentaries, while Paker and Özcan (2017) made use of interviews and learner diaries in order to find out the effectiveness of corpus-based activities. Nevertheless, our study had the advantage of continuous feedback from the teacher, who was also the researcher/designer of the pedagogic corpus and the DDL material, an approach recommended by scholars such as Lin and Lee (2015).
d) “Are the gains of a DDL methodology retained eight months later?”

Both the experimental and control groups showed an increase in the delayed post test scores as compared to the pre-test, however these scores are lower than those of the immediate post-test. The delayed post test results of the experimental group 1 show a rise in the score only in those items which involve the use of ‘will’. As concerns the other five modal verbs, in which those students’ scores were lower than in the post-test, the modal verb could receive the largest number of corrects answers, followed by ‘might’, ‘may’, ‘should’, ‘would’ and ‘will’. This overall decline in the delayed post-test scores is only natural, since 8 months intervened between the administration of immediate post and delayed post-tests including two months of summer holidays, during which the students had no English classes. The eight-month interval between the immediate and the delayed post-test, was a decision based on our wish to discover how much knowledge was retained after a long period of time, during which, even though the students were taught English for two hours per week, they had no further lessons on modal verbs. Regarding the decline in the delayed post-test scores, one should not overlook the fact that, when they took the delayed post-test, those students were in 3rd grade of senior high school, a year during which the students in Greece tend to focus on their preparation for the end-of-school year University entrance examinations. Foreign language learning is not a priority for them unless they plan to study abroad or pursue further education in a foreign language teaching or translation and interpreting related university school.

Notwithstanding these facts, the delayed post-test scores outmatch those of the pre-test and this clearly indicates that new knowledge was gained. Therefore, the decline in delayed post-test scores could not be interpreted as a failure to master the formation and use of modal verbs.

What is interesting though, is the definite increase in the mean score calculated for the successful hits made for each modal verb in the delayed post-test, 73.26% as compared to 70.29 in the case of successful response rating. More successful hits were found in those items examining the use of ‘could’, followed by ‘may’, ‘might’, ‘should’, ‘would’ and, lastly, ‘will’. The case of ‘will’ is actually the most interesting of all; it received the highest scores in the delayed post-test albeit the fewer ‘successful hits’. One might
interpret this finding as follows; the students were overly confident of their knowledge of the senses of ‘will’ and tended to overuse it, making thus mistakes in questions which examined the use of other modal verbs.

From the above we reach the conclusion that the students generally made fewer irrelevant choices in the delayed post-test. This fact could be interpreted as a further success of the DDL methodology in the case of the present study. There may be the case, that learners’ immersion into raw data, and their struggle to cope with both advanced language and ‘alien’ tools made them more aware of grammatical structures. We believe that the fact they actively participated in searching and uncovered linguistic patterns and formulated rules on their own, could have imprinted those patterns in their memory. Therefore, this new methodology allowed the students to comprehend the use of modal verbs and to use them more successfully.

In-class observations by the researcher/practitioner showed that DDL turned a teacher-centered learning context into a student-centered classroom (Lin and Lee 2015: 9), where the teacher’s role was that of a facilitator and ‘middle-man’ between the learner and the language. This student-centered classroom is a setting, in which knowledge is constructed in the mind of the learners through exploration, interpretation and autonomy, according to the cognitive-constructivist theories (Braun 2007). To this we will add the advantage of collaborative and communicative learning. The students may have acted autonomously while conducting their searches, but the lack of sufficient number of computers also urged them to work in pairs, or in small groups and then to share their findings. This was perhaps the major advantage of our approach. Not only was knowledge retained eight months later, but also the students acquired new skills and strategies concerning foreign language learning; they were implicitly taught how to assume responsibility for their own learning, how to search, hypothesize, work in pairs and groups, comment on others’ findings and form their own generalizations.

8.3 Major findings of the study

The main objective of the present empirical research was to design, apply and prove the usefulness of a DDL approach in foreign language teaching. This study has found that:
1) The use of ICT for the learning of foreign languages is most welcome by students since they are already familiar with working on computers either for recreational or educational purposes;

2) The use of corpora and concordancers is not something that students are familiar with, but given time and appropriate material, they could learn how to effectively employ them for learning;

3) Suitable DDL material for use in the EFL class could be designed and compiled by EFL teachers given that they take into consideration the curriculum for the teaching of EFL in schools;

4) A blended DDL methodology which incorporates paper-based and online learning, based on both raw and edited corpus data, is most suitable for teenage students as it slowly and gently introduces them into corpus driven learning;

5) The web-as-a-corpus approach, is a novel methodology which can especially facilitate the students’ introduction to learning with corpora in a non-threatening way;

6) A DDL methodology can positively affect the learning of grammar, and the improvement in the immediate posttest and delayed posttest scores is both statistically and practically significant;

7) There is a statistically significant relationship between the traditional grammar teaching methodology and the students’ scores in the immediate and delayed posttests. This shows that the traditional, course book-base methodology can also affect positively the learning of modal verbs;

8) Our findings show that the average mean score growth between pretest and immediate post-test and between pre-test and delayed post-test was higher in the two experimental groups. One might interpret this a success of the DDL methodology over the traditional teaching methodology, as far as the present study is concerned;

8) Low level students, who had received the lowest grades in the pre-test, are those who showed the greatest progress. This concerned individual cases of students, as well as the experimental group 2 as a whole, which had the lowest mean score in the pretest and also had the highest mean score difference between the pretest and immediate posttest. As practitioners, we would have lower expectations from the so called ‘low
level’ learners, but the findings proved us wrong. This finding might be considered as a possible advantage of the DDL methodology – to offer an alternative way of learning to students who seem not to have benefited in the past by the traditional teaching methodology, at least, as much as others;

9) The knowledge gained through the participation in the DDL lessons is retained long after the lessons have been concluded.

8.4 Conclusion

The main findings of the present study can be summarized as follows. Both the traditional and DDL methodology yielded positive results and there is an evident longitudinal effect for both the experimental groups and the control group. However, this is more striking in the case of the experimental groups and even more so in the case of those students who, as shown in the pre-test results, had a low knowledge level of modal verbs. The following chapter will suggest improvements or alterations which could be made to the DDL methodology proposed in the present study, including the compilation of the pedagogical corpus, the structure of the DDL lessons and the design of the activities. It will also discuss the theoretical and methodological implications of a DDL approach and will also suggest areas in which our research could expand to in the future.
CHAPTER 9

CONCLUSION

9.1 Introduction

As already stated, this exploratory study aimed at examining whether a DDL methodology can be applied in the teaching of English as a foreign language in Greek state senior high schools and how this could be achieved. In other words, it investigated in detail the means and conditions that could facilitate the application of a DDL methodology in a Greek state school EFL class. Its primary goal was to shed light to the usefulness of such an approach to grammar teaching as compared to the traditional course book-based teaching methods. This last chapter will provide the implications for pedagogy and also some suggestions for future research.

9.2 Implications for pedagogy

The evaluation of the learning outcomes showed in this course design led to useful observations concerning foreign language pedagogy. In particular, the quantitative evidence drawn from the pre, post and delayed posttests, presented in chapter 7, as well as classroom observation and informal interviews with students, during and after the DDL lessons provided insight as to how the highly versatile DDL methodology can be adapted so as to be of value for the foreign language teaching.

9.2.1 The case for richness and diversity of materials

The overall rationale of our research agrees with previous statements by scholars such as McCarthy, Matthiessen and Slade (2002) that producers of pedagogical materials can make those materials more representative of language use. Grammar as it is taught
in textbooks is “exceptionally incomplete, partial and misleading” (Johns 1991: 3) and therefore an immersion into data derived from corpora perfectly complements the knowledge one could receive from traditional course books. Our approach showed that a ‘blended’ strategy to materials design serves best the interests of foreign language learning for all students. The combination of a tailor-made pedagogic corpus which contains language found in course books, whose usefulness has also been discussed in Aston 1997, and online corpora which contain genuine communicative situations, advocated -among others- by Bernardini 2000, de Schryver 2002 and Krishnamurthy 2001, may help moderate the degree of difficulty of the DDL activities. Our materials are made up from traditional course book language that students are familiar with and real, natural occurring language which is in accordance with the demands of modern pedagogy for communicative language teaching. The overall rationale of our research agrees with researchers such as Braun (2007: 308) that language learning should be based on a rich repertoire of materials.

9.2.2 The middle-way: the case for blended Data-driven language learning

The findings of the present study have added to past reports which claim that a DDL methodology can be equally, or more, effective to deductive approaches to language learning. The principle of learner autonomy has to be introduced gradually progressing from less to more autonomous tasks (Heift & Vyatkina 2017: 39). This is mainly due to the amount of stress adolescents may experience when moving from a traditional teacher-centered learning methodology to a novel DDL methodology which urges the learner to be a leader instead of a follower in the learning procedure.

Ours is a blended approach, using a combined inductive and deductive approach, since to a certain degree it is teacher-guided and leads the students to formulate grammar rules. This approach is in agreement with previous studies, such as the one by Rapti (2010) stating that students may face difficulties while copying with complex data and therefore need teacher’s guidance and prompting.

Teenage students are on the threshold of cognitive maturity and metalinguistic awareness. Since successful learning draws on different mental abilities (Lightbrown & Spada 2006: 31) the pedagogical material and activities created for senior high school
students should be suitably modified and provide the learners with teacher feedback and guidance.

9.2.3 The case for a combined DDL and Traditional learning methodology

Quantitative evidence drawn from the present study’s findings leads us to the conclusion that both DDL and traditional learning methodologies can positively affect the learning of modal verbs. The immediate posttest and delayed posttest means scores may have been higher in the case of the experimental groups and a statistically significant improvement in the immediate and delayed posttest scores may have been made, but one should not oversee the fact that, a) in the case of the control group there is also significant improvement in the immediate and delayed posttests, and b) that there was no statistically significant difference between the three groups of students, in the pre, immediate post and delayed posttests respectively.

Therefore, one would welcome a combined DDL and traditional learning methodology, at least as far as grammar is concerned. This would offer the best of both worlds; on the one hand, the readily available material and the familiarity of the traditional approach, and, on the other hand, the abundance of authentic language and the excitement of discovery that DDL offers.

9.2.4 What does DDL have to offer to EFL teaching in Greek state schools?

Research on DDL application to foreign language learning is not yet at a point where it is possible to say to teachers: “Here is a list of what you should include in the design and application of your DDL lessons what you should avoid”. This echoes the statement by Lightbrown and Spada (2006: 177) that “the learner’s age, metalinguistic sophistication, prior educational experiences, motivation and goals need to be taken into account” when designing and implementing a new teaching methodology.

To the aforementioned conditions, we would add that, in the case of DDL implementation in Greek senior high school setting, it is inevitable that one also bear in mind the inherent limitations of the Greek educational reality, the curricula set by the
Ministry of Education and Religious Affairs, the time limitations imposed by the school timetable and the sad reality of the poorly equipped computer laboratories.

One would say thus, that DDL is in essence a tailor-made methodology, a versatile tool and pliable clay to the hands of the practitioner; he will use it and shape it according to the needs of the learners and the demands of local curricula. However, the power and dynamics of the raw material, data, and the means to exploit it, corpora and concordancers, are ever present. After 30 years of relevant research in the field, there have been numerous empirical studies suggesting different ways of employing corpora in the foreign language learning classroom. They do not dictate a single methodology; they merely suggest ways to turn the lesson into a data-driven one. They may exploit tailor-made or general online corpora, they may involve paper based or hands on online searches, they may make use of mediated or raw concordance output, they may involve much or no teacher intervention, and any combination of the above. The choices are endless. The training sessions, pilot lessons and final DDL lessons of the present study showed that DDL is in essence a trial-and-error process through which the practitioner could design and implement his/her own ideal lessons.

9.3 Directions for future research

We see numerous opportunities for further improvement of our framework. One could be to complement our DDL lessons by enriching them with more ‘production’ or ‘creation’ activities, adding thus to the third step of the presentation-practice-production process. Our DDL lessons consisted mainly of presentation and practice, but few production activities, due to time limitations, with them mainly asking the students to synthesize chunks taken from concordance output. This allowed little to the creativity of the learners. Ideally, a future endeavor would complement them with role playing, creative writing and other, more communicative, activities.

The present empirical study drew its conclusions from the quantitative results of tests. Another future action could be to receive more feedback by the students after the end of the course with the administration of appropriate questionnaires and interviews. This way we could achieve qualitative as well as quantitative evaluation of our methodology. The questionnaires, as well the interviews could provide valuable information as to the
appeal of the DDL approach, the problems they may encounter and -even – possible suggestions for the improvement of the approach.

It is true that the study case of the present research endeavor comprised only a small-scale experiment of 60 students. We believe that a larger scale experiment with a greater number of students could validate the positive results attained by our target group. Another point that could be covered with future actions is the introduction of unknown material, such as for example, general academic vocabulary for Senior high school students. We also agree with Sun, Y-C. & Wang, L-Y. (2003: 91) that a longitudinal study which would track student performance over a longer period of time would be of much interest. This could be a way to get “holistic benefits that can accrue from the regular practice of DDL” (Kerr 2013: 17). Further research could also involve an approach employing a blend of DDL and Grammar Translation, as suggested by Lin and Lee (2015). This could appeal to a wider spectrum of students, those who favor the traditional teaching approach and those who are openminded to novel approaches to language learning; one should not forget that an ordinary class includes both types of students.

9.4 Conclusion

Data-driven learning is a bottom-up approach which once the teachers and students have been trained in working with, they gain the relevant bottom-up skills; they learn how to deal with two amazing things, technology and data. This is of extreme importance nowadays, because the amount of technology and tools we have at our fingertips is amazing; we have the ability to track and monitor, to evaluate and revise our decisions, goals and methodologies. Data, either big or little, may still look like a mountain to the average teacher. One should not forget though, that mountains offer the best views.

Data-driven learning is ever present in a plethora of everyday applications, in science, marketing, even politics and the military and last, but certainly not least, in education and learning. The present thesis corroborates the findings of previous studies that data and analytics can allow us to continue to improve our knowledge of how language is used, modify existing methodologies, create teaching material and eventually modernize EFL learning. After almost three decades of work done on DDL, one can
safely reach the conclusion that there exists indeed an elective affinity, i.e. a correspondence or mutual attraction, between corpora and foreign language teaching. Nevertheless, our study has shown that any successful formula should be carefully suited to the needs of the learner target group. This is definitely the optimal way to plant the seed of curiosity for learning into the minds of our students, which will eventually turn them into language detectives who take responsibility for and make informed decisions regarding their learning needs and goals. There is no other way for us to achieve the results that the needs of 21st century dictate and our students rightly deserve.
BIBLIOGRAPHY


Alruwaili, A. (2016). *Have intermediate level students learned how to use the corpus approach for learning general verbs in the classroom?* In the proceedings of TaLC 12 -12th Teaching and Language Corpora conference in 2016, Giessen, Germany

Angouri, J., M. Mattheoudakis & M. Zigrika, (2010). “*Then how will they get the ’much-wanted paper’? A multifaceted study of English as a foreign language in Greece.*” GALA – Advances in research on Language Acquisition and Teaching: Selected Papers, 179-194


Rohdenburg, G. (2007). Determinants of grammatical variation in English and the formation / confirmation of linguistic hypotheses by means of internet data. In M.
Hundt, N. Nesselhauf & C. Biewer (Eds.), *Corpus linguistics and the web* (pp. 191-209). Amsterdam: Rodopi.


Ruppenhofer, J., & Rehbein, I. (2016). *Yes we can!? Annotating the senses of English modal verbs*, (pp. 1538-1545). Institut für Deutsche Sprache, Bibliothek.


APPENDICES

A. The Pedagogic corpus ECCo

A.1 The header

This is the header containing the metalinguistic information on the pedagogic corpus

<teiHeader>
<fileDesc>
<titleStmt>
<title>ECCo: Corpus of English Coursebooks</title>
</title>
<respStmt>
<resp>compiled by</resp>
<name>Vasiliki Papaioannou</name>
</respStmt>
</titleStmt>
<editionStmt>
<edition>
<date>August 2011-July 2012</date>
</edition>
</editionStmt>
</publicationStmt>
<distributor>Aristotle University of Thessaloniki</distributor>
</fileDesc>
</teiHeader>
<sourceDesc>

<bibl>

<book category="COURSEBOOK">

<title>Traveller B1_plus students' book. MM Publications. 2008- reading parts</title>
<title>Journeys B1_plus students' book. Hillside Press. 2010- listening transcripts</title>
<title>Traveller B1_plus students' book. MM Publications. 2008- listening transcripts</title>

<p>only reading parts and listening transcripts have been included</p>

</bibl>
</sourceDesc>

<encodingDesc>

<samplingDecl>

</samplingDecl>

</encodingDesc>
A whole new wardrobe changes Alex’s image. Look good, feel good! Before - after. Alexandria is a 15-year-old schoolgirl. She's very sporty and fit and enjoys life. She isn't really interested in following the fashion; she just chooses clothes that she feels comfortable in. She doesn't believe that people judge her by what she looks like. Her sister Charlotte, though, disagrees with this and wants to replace some of her clothes with a few smarter outfits. So, with the help of a fashion mentor, Charlotte aims to give Alexandria a makeover! Jane, the stylist, turns up at Alexandria's house. While they are chatting, she draws attention to the fact that Alexandria's clothes have one thing in common: 'There's not a skirt in sight!' It's true, Alexandria's wardrobe is full of baggy tracksuit's T-shirts and some denim jeans. She avoids wearing brightly-coloured clothes but Jane is going to bring in a bit of variety! Alexandria now seems keen to face this new challenge and create a new image for herself. Jane takes the two sisters to the shops where Charlotte has the chance to play a being stylist. However, the items she chooses for Alexandria to try on are not quite her style, although they're getting near. Finally, it's Jane's turn and she chooses three new outfits for Alexandria. When Alexandria comes out of the changing room with them on, it's difficult to recognize her as the same person. In her skirts, fancy tops and heels and with a new funky hairstyle, she could be a professional model! Charlotte is hugely impressed with her sister's chic new look and Alexandria herself looks confident and happy. Now she will stand out in a crowd, whether she likes it or not! Zac, 17, wants to do a bit of modelling in his spare time, but he has to overcome some personality problems first. He gets so shy and nervous that he can't talk to people and he can't cope with the stress of casting sessions for photo shoots. 'My hands get all sweaty and I mix up my words if I feel I'm the centre of attention,' he says. He finds it difficult to make eye contact with people he doesn't know, so he often
comes across as an unfriendly or aggressive character. He feels he should keep up with the fashion but he always wants to blend in, so he tends to wear the same urban-style clothes as everyone else. A psychologist helps him to loosen up and get over his fears. She tells him that his body language is all wrong: 'You're always sending out the wrong signals!' Over the course of a few weekends, Zac learns to express himself properly and he even manages to show up for an interview for a modelling job! Since he had missed several similar appointments in the past, this is a major achievement. The process of changing isn't easy but if he can get it right, nothing will hold him back anymore! Zac's height and build are just right for modelling so, after getting some advice from a stylist about his clothes, he soon manages to get some catalogue work and offers of future jobs. Everyone is delighted with his progress, and his mentors and friends hope that he'll move on to fashion shows and the catwalk.

A.3 The UCREL CLAWS7 tagset

APPGE  possessive pronoun, pre-nominal (e.g. my, your, our)
AT    article (e.g. the, no)
AT1   singular article (e.g. a, an, every)
BCL   before-clause marker (e.g. in order (that), in order (to))
CC    coordinating conjunction (e.g. and, or)
CCB   adversative coordinating conjunction (but)
CS    subordinating conjunction (e.g. if, because, unless, so, for)
CSA   as (as conjunction)
CSN   than (as conjunction)
CST   that (as conjunction)
CSW   whether (as conjunction)
DA    after-determiner or post-determiner capable of pronominal function (e.g. such, former, same)
DA1   singular after-determiner (e.g. little, much)
DA2   plural after-determiner (e.g. few, several, many)
<table>
<thead>
<tr>
<th>Code</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>DAR</td>
<td>comparative after-determiner (e.g. more, less, fewer)</td>
</tr>
<tr>
<td>DAT</td>
<td>superlative after-determiner (e.g. most, least, fewest)</td>
</tr>
<tr>
<td>DB</td>
<td>before determiner or pre-determiner capable of pronominal function (all, half)</td>
</tr>
<tr>
<td>DB2</td>
<td>plural before-determiner (both)</td>
</tr>
<tr>
<td>DD</td>
<td>determiner (capable of pronominal function) (e.g. any, some)</td>
</tr>
<tr>
<td>DD1</td>
<td>singular determiner (e.g. this, that, another)</td>
</tr>
<tr>
<td>DD2</td>
<td>plural determiner (these, those)</td>
</tr>
<tr>
<td>DDQ</td>
<td>wh-determiner (which, what)</td>
</tr>
<tr>
<td>DDQGE</td>
<td>wh-determiner, genitive (whose)</td>
</tr>
<tr>
<td>DDQV</td>
<td>wh-ever determiner, (whichever, whatever)</td>
</tr>
<tr>
<td>EX</td>
<td>existential there</td>
</tr>
<tr>
<td>FO</td>
<td>Formula</td>
</tr>
<tr>
<td>FU</td>
<td>unclassified word</td>
</tr>
<tr>
<td>FW</td>
<td>foreign word</td>
</tr>
<tr>
<td>GE</td>
<td>germanic genitive marker - (’ or’s)</td>
</tr>
<tr>
<td>IF</td>
<td>for (as preposition)</td>
</tr>
<tr>
<td>II</td>
<td>general preposition</td>
</tr>
<tr>
<td>IO</td>
<td>of (as preposition)</td>
</tr>
<tr>
<td>IW</td>
<td>with, without (as prepositions)</td>
</tr>
<tr>
<td>JJ</td>
<td>general adjective</td>
</tr>
<tr>
<td>JJR</td>
<td>general comparative adjective (e.g. older, better, stronger)</td>
</tr>
<tr>
<td>JJT</td>
<td>general superlative adjective (e.g. oldest, best, strongest)</td>
</tr>
<tr>
<td>JK</td>
<td>catenative adjective (able in be able to, willing in be willing to)</td>
</tr>
<tr>
<td>MC</td>
<td>cardinal number, neutral for number (two, three..)</td>
</tr>
<tr>
<td>MC1</td>
<td>singular cardinal number (one)</td>
</tr>
<tr>
<td>MC2</td>
<td>plural cardinal number (e.g. sixes, sevens)</td>
</tr>
<tr>
<td>MCGE</td>
<td>genitive cardinal number, neutral for number (two's, 100's)</td>
</tr>
<tr>
<td>MCMC</td>
<td>hyphenated number (40-50, 1770-1827)</td>
</tr>
<tr>
<td>MD</td>
<td>ordinal number (e.g. first, second, next, last)</td>
</tr>
<tr>
<td>-----</td>
<td>-----------------------------------------------</td>
</tr>
<tr>
<td>MF</td>
<td>fraction, neutral for number (e.g. quarters, two-thirds)</td>
</tr>
<tr>
<td>ND1</td>
<td>singular noun of direction (e.g. north, southeast)</td>
</tr>
<tr>
<td>NN</td>
<td>common noun, neutral for number (e.g. sheep, cod, headquarters)</td>
</tr>
<tr>
<td>NN1</td>
<td>singular common noun (e.g. book, girl)</td>
</tr>
<tr>
<td>NN2</td>
<td>plural common noun (e.g. books, girls)</td>
</tr>
<tr>
<td>NNA</td>
<td>following noun of title (e.g. M.A.)</td>
</tr>
<tr>
<td>NNB</td>
<td>preceding noun of title (e.g. Mr., Prof.)</td>
</tr>
<tr>
<td>NNL1</td>
<td>singular locative noun (e.g. Island, Street)</td>
</tr>
<tr>
<td>NNL2</td>
<td>plural locative noun (e.g. Islands, Streets)</td>
</tr>
<tr>
<td>NNO</td>
<td>numeral noun, neutral for number (e.g. dozen, hundred)</td>
</tr>
<tr>
<td>NNO2</td>
<td>numeral noun, plural (e.g. hundreds, thousands)</td>
</tr>
<tr>
<td>NNT1</td>
<td>temporal noun, singular (e.g. day, week, year)</td>
</tr>
<tr>
<td>NNT2</td>
<td>temporal noun, plural (e.g. days, weeks, years)</td>
</tr>
<tr>
<td>NNU</td>
<td>unit of measurement, neutral for number (e.g. in, cc)</td>
</tr>
<tr>
<td>NNU1</td>
<td>singular unit of measurement (e.g. inch, centimetre)</td>
</tr>
<tr>
<td>NNU2</td>
<td>plural unit of measurement (e.g. ins., feet)</td>
</tr>
<tr>
<td>NP</td>
<td>proper noun, neutral for number (e.g. IBM, Andes)</td>
</tr>
<tr>
<td>NP1</td>
<td>singular proper noun (e.g. London, Jane, Frederick)</td>
</tr>
<tr>
<td>NP2</td>
<td>plural proper noun (e.g. Browns, Reagans, Koreas)</td>
</tr>
<tr>
<td>NPD1</td>
<td>singular weekday noun (e.g. Sunday)</td>
</tr>
<tr>
<td>NPD2</td>
<td>plural weekday noun (e.g. Sundays)</td>
</tr>
<tr>
<td>NPM1</td>
<td>singular month noun (e.g. October)</td>
</tr>
<tr>
<td>NPM2</td>
<td>plural month noun (e.g. Octobers)</td>
</tr>
<tr>
<td>PN</td>
<td>indefinite pronoun, neutral for number (none)</td>
</tr>
<tr>
<td>PN1</td>
<td>indefinite pronoun, singular (e.g. anyone, everything, nobody, one)</td>
</tr>
<tr>
<td>PNQO</td>
<td>objective wh-pronoun (whom)</td>
</tr>
<tr>
<td>PNQS</td>
<td>subjective wh-pronoun (who)</td>
</tr>
<tr>
<td>PNQV</td>
<td>wh-ever pronoun (whoever)</td>
</tr>
</tbody>
</table>
PNX1    reflexive indefinite pronoun (oneself)
PPGE    nominal possessive personal pronoun (e.g. mine, yours)
PPH1    3rd person sing. neuter personal pronoun (it)
PPHO1   3rd person sing. objective personal pronoun (him, her)
PPHO2   3rd person plural objective personal pronoun (them)
PPHS1   3rd person sing. subjective personal pronoun (he, she)
PPHS2   3rd person plural subjective personal pronoun (they)
PPIO1   1st person sing. objective personal pronoun (me)
PPIO2   1st person plural objective personal pronoun (us)
PPIS1   1st person sing. subjective personal pronoun (I)
PPIS2   1st person plural subjective personal pronoun (we)
PPX1    singular reflexive personal pronoun (e.g. yourself, itself)
PPX2    plural reflexive personal pronoun (e.g. yourselves, themselves)
PPY     2nd person personal pronoun (you)
RA      adverb, after nominal head (e.g. else, galore)
REX     adverb introducing appositional constructions (namely, e.g.)
RG      degree adverb (very, so, too)
RGQ     wh- degree adverb (how)
RGQV    wh-ever degree adverb (however)
RGR     comparative degree adverb (more, less)
RGT     superlative degree adverb (most, least)
RL      locative adverb (e.g. alongside, forward)
RP      prep. adverb, particle (e.g about, in)
RPK     prep. adv., catenative (about in be about to)
RR      general adverb
RRQ     wh- general adverb (where, when, why, how)
RRQV    wh-ever general adverb (wherever, whenever)
RRR     comparative general adverb (e.g. better, longer)
RRT     superlative general adverb (e.g. best, longest)
<table>
<thead>
<tr>
<th>Code</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>RT</td>
<td>quasi-nominal adverb of time (e.g. now, tomorrow)</td>
</tr>
<tr>
<td>TO</td>
<td>infinitive marker (to)</td>
</tr>
<tr>
<td>UH</td>
<td>interjection (e.g. oh, yes, um)</td>
</tr>
<tr>
<td>VB0</td>
<td>be, base form (finite i.e. imperative, subjunctive)</td>
</tr>
<tr>
<td>VBDR</td>
<td>were</td>
</tr>
<tr>
<td>VBDZ</td>
<td>was</td>
</tr>
<tr>
<td>VBG</td>
<td>being</td>
</tr>
<tr>
<td>VBI</td>
<td>be, infinitive (To be or not... It will be ..)</td>
</tr>
<tr>
<td>VBM</td>
<td>am</td>
</tr>
<tr>
<td>VBN</td>
<td>been</td>
</tr>
<tr>
<td>VBR</td>
<td>are</td>
</tr>
<tr>
<td>VBZ</td>
<td>is</td>
</tr>
<tr>
<td>VD0</td>
<td>do, base form (finite)</td>
</tr>
<tr>
<td>VDD</td>
<td>did</td>
</tr>
<tr>
<td>VDG</td>
<td>doing</td>
</tr>
<tr>
<td>VDI</td>
<td>do, infinitive (I may do... To do...)</td>
</tr>
<tr>
<td>VDN</td>
<td>done</td>
</tr>
<tr>
<td>VDZ</td>
<td>does</td>
</tr>
<tr>
<td>VH0</td>
<td>have, base form (finite)</td>
</tr>
<tr>
<td>VHD</td>
<td>had (past tense)</td>
</tr>
<tr>
<td>VHG</td>
<td>having</td>
</tr>
<tr>
<td>VHI</td>
<td>have, infinitive</td>
</tr>
<tr>
<td>VHN</td>
<td>had (past participle)</td>
</tr>
<tr>
<td>VHZ</td>
<td>has</td>
</tr>
<tr>
<td>VM</td>
<td>modal auxiliary (can, will, would, etc.)</td>
</tr>
<tr>
<td>VMK</td>
<td>modal catenative (ought, used)</td>
</tr>
<tr>
<td>VV0</td>
<td>base form of lexical verb (e.g. give, work)</td>
</tr>
<tr>
<td>VVD</td>
<td>past tense of lexical verb (e.g. gave, worked)</td>
</tr>
<tr>
<td>VVG</td>
<td>-ing participle of lexical verb (e.g. giving, working)</td>
</tr>
</tbody>
</table>
A.4 Sample CLAWS7 tagger output

Look_VVB good_AJ0 !_PUN feel_VVB good_AJ0 !_PUN before_AV0 after_PRP._SENT

Alexandria_NP0 is_VBZ a_AT0 15-year-old_AJ0 schoolgirl_NN1 ._SENT

She_PNP 's_VBZ very_AV0 sporty_AJ0 and_CJC fit_AJ0 and_CJC enjoys_VVZ life_NN1

She_PNP is_VBZ n't_XX0 really_AV0 interested_AJ0 in_PRP following_VVG the_AT0 fashion_NN1 ;_PUN she_PNP just_AV0 chooses_VVZ clothes_NN2 that_CJT she_PNP

VVGK  -ing participle catenative (going in be going to)
VVI  infinitive (e.g. to give... It will work...)
VVN  past participle of lexical verb (e.g. given, worked)
VVNK  past participle catenative (e.g. bound in be bound to)
VVZ  -s form of lexical verb (e.g. gives, works)
XX  not, n't
ZZ1  singular letter of the alphabet (e.g. A,b)
ZZ2  plural letter of the alphabet (e.g. A’s, b’s)
feels_VVZ comfortable_AJ0 in_PRP _SENT ------_PUN

A.5 WordSmith Tools 6 concordancer output

A.5.1 Full sentences (keyword: past perfect verbs)

N Concordance
2 ed across the desert. This was the path that so many adventurers had taken in the hope of finding water. But what became of them w
N Concordance
2 ed across the desert. This was the path that so many adventurers had taken in the hope of finding water. But what became of them w
N Concordance
19 tic child's watch! It belonged to Bert's grandson. However, Bert had put it in the cupboard and forgotten about it. The alarm was
N Concordance
22 Lloyds Bank in Shellsby. One witness told reporters that Briggs had run into the bank holding a gun and threatened to shoot every
N Concordance
26 00 feet over the western United States. Fortunately the captain had been on one of our courses, he able to stay calm even though
N Concordance
27 turday, I attended a Chinese wedding at temple. My friend Chang, had given me the invitation in a red envelope. I found out that a Chi

A.5.2 Blanking the keyword (past perfect tense verbs)

N Concordance
1 ing they'd shared together. Darren had come into her life unexpectedly and ___ her world completely. And now, just as suddenly, it looked a
N Concordance
2 s later. Alexander himself never saw the city. He left before the building ______ , and later died in Babylon (in today's Iraq). Alexandria was
Concordance
3 ur rows of seats, so Rosa sat in the fifth row. After a few stops, the bus _____ up, and one white man didn't have a seat. The bus driver
Concordance
4 re studying the seismic activity. However, one man did report that his dog _____ to howl at the time the quake was said to have started. The
Concordance
5 ch intelligent friendly creatures but, at the same time, I was glad that I ____ the dolphin. The Rio Carnival in Brazil. This festival is by
Concordance
6 there was nice, a little cold but the fresh made us feel good. I was glad I____ persuade my friends to join me. We were walking along, admir

B. The Questionnaires

B.1 The Questionnaire investigating the student’s attitudes towards Computer Assisted Language Learning

1. Do you use a computer?
   o Yes
   o No

2. How often do you use a computer?
   o Once a week
   o Twice a week
   o Three times a week
   o Every day

3. What do you do when you're using the computer?
   Surf the net
   Do homework
   Use Social Media
Send emails/ make calls to friends and relatives

4. Which e-learning medium do you find most interesting?
   o Online exercises
   o Dictionaries
   o Reading online articles
   o Email communication
   o Social media
   o Videos/ movies
   o Other……

5. Which of the above mentioned e-learning media do you use?
   i. Online exercises
   j. Dictionaries
   k. Reading online articles
   l. Email communication
   m. Social media
   n. Videos/ movies
   o. Other……

6. Which of the above-mentioned e-learning media do you use to learn foreign languages?
   o Online exercises
   o Dictionaries
   o Reading online articles
   o Email communication
   o Social media
   o Videos/ movies
   o Other……

7. Does your school provide computers for use by students?
   o Yes
8. Are (some of) the computer facilities available to students after IT/other classes?
   - Yes
   - No

9. Are the computers connected to the internet?
   - Yes
   - No

10. Do you use internet at home for learning a foreign language?
    - Yes
    - No

11. Would you be interested in taking part in a computer-assisted language learning lesson?
    - Yes
    - No

12. Do you think taking part in computer-assisted language learning lesson would be more effective than traditional teaching?
    - Yes
    - No

13. Can you see any advantages in taking part in computer-assisted language learning lesson?
    - Yes
    - No

Which are those?..........................
14. Can you see any disadvantages in taking part in computer-assisted language learning lesson?
   
   o Yes
   o No

Which are those?..............................

15. Do you think that taking part in computer-assisted language learning lesson would be more difficult than traditional learning?
   
   o Yes
   o No

16. Do you think that taking part in computer-assisted language learning lesson might raise the interest of students to learn foreign language?
   
   o Yes
   o No
### B.2 The Questionnaire investigating the student’s attitudes towards Grammar Learning

1=strongly disagree  
2= disagree  
3= neither agree nor disagree  
4= agree  
5= strongly agree

<table>
<thead>
<tr>
<th>Questions</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 I like to study English Grammar</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2 The best way for me to learn is by studying rules</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3 The best way for me to learn is study is to read texts</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4 The best way for me to learn is study rules and read texts</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5 Grammar and vocabulary are two completely separate parts of English</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6 Grammar and vocabulary should be taught separately</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7 I must know grammar in order to speak/ write in English</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>8 I must know grammar rules in order to speak/ write in English</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>9 I must know many individual words in order to speak/ write in English</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>10 I must know many word chunks in order to speak/ write in English</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
On Staten Island, Relentless Bullying Is Blamed for a Teenage Girl’s Suicide

Robert Stolarik for The New York Times

Tottenville High School students gathered after class at a makeshift memorial for Felicia Garcia, 15, who leapt to her death at the Huguenot train station on Staten Island this week.

By Vivian

Felicia Garcia rarely cried, so when her friend saw her sobbing alone in a hallway between classes at Tottenville High School on Wednesday, she knew something was wrong.

The friend, Briana Torres, at 16 a year older and a grade ahead of Felicia, hugged her and walked her to sixth-period English class, the girl’s arms clasped around each other’s shoulders. On the way, Felicia cheered up enough to laugh at a joke, and make a joke of her own.

But there were signs of unraveling. Late Monday night, she had posted a brief Twitter message: “I cant, im done, I give up.” After school Wednesday, Felicia walked to the Staten Island Railway station where many students board trains home. She waited impatiently for the train, and as it approached, she hurled herself backward onto the tracks. A friend grabbed her arm, but she twisted free. She was pronounced dead that evening.

By the time her friends began to congregate in the hospital waiting room, posting messages on Twitter and Facebook in what would become a flurry of online speculation about her death, most had pinpointed a cause: Felicia had been bullied, they said, tormented by football players on Tottenville’s undefeated team. Some said she was teased because she had piercings and lived in foster care. Others said players had spread sexual boasts about her over the weekend, after Tottenville’s 16-8 victory over Port Richmond High School.

To many friends, she appeared to weather the swirl of innuendo with her usual confidence. “She never really reached out for help, she was a really tough person,” Briana said Thursday.
wearing a small tribute on her left wrist — an “RIP Felicia” inked in purple. “When I dropped her off at class, I wasn’t really worried about her.”

Felicia had reported the taunts to an administrator, who arranged mediation sessions between Felicia and the boys she said were harassing her. Police are now investigating her death. Neither they nor the Education Department nor the school would comment on the bullying allegations.

There was already little that was easy in Felicia’s life. Friends described her childhood as a patchwork of loss and instability: both her parents died when she was young, and she disliked living with her aunt, said Kaitlyn Antonomarchi, 15, who said she had been Felicia’s best friend since eighth grade. At one point, Felicia ran away from her aunt’s house with an older man. After she entered the foster system, she bounced in and out of different homes, dyed her dark hair red and sprouted a cluster of piercings.

With her latest foster parents, Felicia finally seemed happy and stable, Kaitlyn said. Moving to the other side of Staten Island, she started high school at Tottenville, improved her grades, let the dye wash out and eliminated most piercings. At Friday’s football game, Kaitlyn said: “She looked happy. She was laughing. It didn’t look like anything was upsetting her at all.”

Bullying is common at the school, classmates said, but administrators usually acted to stop it, and rarely reached the level that Felicia experienced. Tease Felicia, and she would come back with a quick, witty retort, said Alissa Compatello, 17, a senior.

“If you tried to bully her, she’d laugh at you,” she said. “Somebody must’ve said something pretty bad about her for this to happen. They just wouldn’t stop.”

On Wednesday, Felicia had asked Karl Geiling, 15, a sophomore at Tottenville, about how his test had gone. He saw her at the train station later. “I was way down, away from her,” he said. “All I heard was screams, and then everybody went silent.”

At school on Thursday, many students wore black and purple, colors often associated with anti-bullying campaigns, and met with grief counselors. A crowd of about 500 gathered at the station in the evening, many holding candles. Someone had tied purple and black balloons to a chain-link fence overlooking the tracks, with notes and a photo fluttering alongside them.

As their classmates created anti-bullying Facebook pages in Felicia’s honor Wednesday night, several football players took to Twitter to protest what they saw as the wholesale tearing of the team, which is a perennial favorite to win the Public School Athletic League championship. At least two seniors have been offered scholarships to play Division I college football.

“None of you even half the story so stop pointing fingers at the football team,” wrote James Munson, a safety on the team and the son of the team’s coach, Jim Munson. Another player, Richy Lam, a senior, said Thursday that many members of the team had not even known Felicia.

In New York, an anti-bullying statute signed in 2010, one of numerous laws passed around the country in the wake of teenage suicides, requires schools to develop policies to deter harassment of students by other students, including education programs and disciplinary procedures.

It is not clear whether anyone will be disciplined in Felicia’s case. For some students, the school’s next challenge is Friday’s football game against the rival Curtis High School team, the last of the season, which may be pushed to Sunday. Felicia was a fan. When Kaitlyn last saw her, she said, she had been planning to cheer Tottenville this weekend.

“She said, ‘Yeah, I’m going.’” Kaitlyn said. “And I said, ‘I’ll see you there.’”

Al Baker and Christopher Maag contributed reporting.
D. The main study
D.1 The pre, post and delayed post-test

Choose the correct word

1. Did you ........ anywhere interesting last weekend?
   a. go
   b. going
   c. visit
   d. went

2. I work as a teacher, and my wife........, too.
   a. do
   b. is
   c. works
   d. does

3. That all teachers .......... have patience is essential.
   a. may
   b. might
   c. could
   d. should

4. Would you like … help?
   a. a
   b. some
5. She ............ have stolen the car as she doesn’t know how to drive.
   a. can
   b. could
   c. couldn’t
   d. shouldn't

6. If you want to lose weight, you ........ run every day.
   a. may
   b. would
   c. shall
   d. should

7. He hasn’t played since he ........ the accident.
   a. had
   b. has had
   c. has
   d. had had

8. You ........ have stayed in bed yesterday. Today you look even worse.
   a. may
   b. might
   c. would
9. When I spoke to them they indicated they ……… find time to come, but they didn’t.
   a. can
   b. will
   c. might
   d. may

10. Mum gave ……… her job when I was born.
    a. in
    b. up
    c. off
    d. away

11. The loudspeakers won't work unless you ……… those cables.
    a. connected
    b. connect
    c. connects
    d. connecting

12. If I had money I ……… travel around the world.
    a. will
    b. would
    c. shall
13. .......... you like to come to dinner with us?
   a. will
   b. would
   c. could
   d. should

14. H promised he .......... help me, but he didn't
   a. will
   b. would
   c. may
   d. should

15. It's a pity you ........ here last night.
   a. weren’t
   b. didn’t
   c. hadn’t
   d. aren’t

16. The TV's too loud. Please, ........
   a. put it down
   b. switch it down
   c. turn it down
   d. close it down
17. Jim................ be homesick, he seems very depressed lately.

   a. can
   b. would
   c. may
   d. should

18. I ............. go home during the next semester break, but I'm not sure yet.

   a. can
   b. shall
   c. will
   d. may

19. Have you had ........ to eat?

   a. too many
   b. some more
   c. everything
   d. enough

20. Mum, ............... I use your mobile, please?

   a. Should
   b. will
   c. may
   d. do

21. You ............ be a better pianist if you practiced more.
22. He didn’t study at all so without a doubt he …………. fail the test.

   a. can  
   b. would  
   c. should  
   d. will

23. ……… is it from here to Berlin?

   a. How long way  
   b. How much time  
   c. How far  
   d. How many

24. He said he ………. go home for the semester break.

   a. might  
   b. can  
   c. will  
   d. shall

25. The financial director ……….. for almost an hour.

   a. kept us to wait
b. kept us waiting  
c. made us to wait  
d. made us waiting 

26. You ……… get your father a tablet for his birthday.  
   a. will  
   b. could  
   c. do  
   d. would 

27. Be careful with those matches. You ……… start a fire.  
   a. should  
   b. could  
   c. would  
   d. shall 

28. Do you think I should move to Ireland? You shouldn't do anything ……… you think it's the right thing to do.  
   a. when  
   b. unless  
   c. in case  
   d. if 

29. Hubert is an uncle of ……… .  
   a. Kim  
   b. Kims  
   c. Kim's  
   d. Kims' 

30. Two years ago I …………. play tennis like a professional.
31. The helpers at the information kiosk............. know the way to the bank.
   a. could
   b. would
   c. should
   d. might

32. ...... you please open the door?
   a. could
   b. would
   c. should
   d. may

33. I ...... have taken my English test yesterday. (but I wasn’t ready)
   a. could
   b. would
   c. should
   d. might

34. ......you bring me some coffee please?
   a. will
   b. may
   c. might
   d. would
D.2 Main study lesson plan 1: “Words and ideas can change the world!”

Class: 2nd grade senior high school – experimental groups 1 & 2

Lesson duration: 45 minutes

Lesson topic: modal verbs (will, would, could, may, might, should)

Introduction to Modal Auxiliaries; learning their grammatical properties

<table>
<thead>
<tr>
<th>Stage</th>
<th>Procedure</th>
<th>Resources</th>
</tr>
</thead>
<tbody>
<tr>
<td>Introduction – setting</td>
<td>The students are presented with 6 different quotes by famous people each containing a modal auxiliary.</td>
<td>Power point presentation</td>
</tr>
</tbody>
</table>
| a purpose               | 11. Believe you can, and you’re halfway there. (T. Roosevelt)  
12. When I let go of what I am, I become what I might be. (Lao Tzu)  
13. You may be disappointed if you fail, but you’re doomed if you don’t try. (B. Sills)  
14. Tension is who you think you should be. Relaxation is what you are (Chinese proverb)  
15. People will forget what you said, people will forget what you did, but they will never forget how you made them feel. (M. Angelou)  
16. W: If I were married to you, I would put poison in your coffee.  
   M: If I were married to you, I would drink it. (1900s’ joke) |
|                         | After reading the quotes the students are asked to comment on them;                                                                                                                                       | Sentences taken from the Web     |
|                         | • Have you ever come across any on these?  
  • Which one do you like most?  
  • Is there anyone difficult to understand?  
  • What is expressed in all these sentences? |                                 |
The answer to the last question should be ‘beliefs, attitudes, moods’

Then the students will be presented another ppt slide where all Modal Auxiliaries in the quotes will be highlighted. They will reach the conclusion that the modal auxiliaries are associated with the expression of beliefs, attitudes and moods.

### D.3 Main study lesson plan 2: “Shall we?”

**Class:** 2nd grade senior high school – experimental groups 1 & 2

**Lesson duration:** 45 minutes

**Lesson topic:** modal verbs (will, would, could, may, might, should)

**Modal Auxiliaries:** learning their grammatical properties

<table>
<thead>
<tr>
<th>The grammatical characteristics of modal auxiliaries</th>
<th>The students are asked to work in pairs in the school pc cluster. They are guided to:</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>• enter the COCA corpus</td>
</tr>
<tr>
<td></td>
<td>• form the following query</td>
</tr>
<tr>
<td></td>
<td>➢ display: KWIC</td>
</tr>
<tr>
<td></td>
<td>➢ word(s): will/would/ should/could/may/might</td>
</tr>
<tr>
<td></td>
<td>➢ sections: spoken + newspaper</td>
</tr>
<tr>
<td></td>
<td>• go through the results. All modal verbs are in the center of each result line and they are all highlighted in the same color. This is why the word ‘may’ in the sentence</td>
</tr>
</tbody>
</table>
“Well, like the whole world back in April, May and June, I was watching in horror” is in grey whereas the word ‘may’ in the sentences “It is tax day as you may already know” and “Is it true that low fat milk may be 2 percent but regular milk is 4 percent?” is in purple.

- Search for patterns in the result lines.
  - What words do MAux usually come after?
  - What words usually follow a MAux?
  - How to they form a question/ negation?
  - How do they form 3rd person singular?

- Make their own generalizations as far as the grammatical properties of MAux are concerned.

- To decide as a team on the four most important grammatical traits of MAux, write them down and present their findings in class.

D.4 Main study lesson plan 3: “Will you join us?”

Class: 2nd grade senior high school – experimental groups 1 & 2

Lesson duration: 50 minutes

Lesson topic: modal verbs (will, would)

Learning the functions of modal auxiliaries will/ would

<table>
<thead>
<tr>
<th>stage</th>
<th>Procedure</th>
<th>Resources</th>
</tr>
</thead>
<tbody>
<tr>
<td>Introducing the modal auxiliaries</td>
<td>Each student is given a handout of concordance output of will/ would instances from ECCo.</td>
<td>Printed handouts from ECCo. Edited, full sentences only.</td>
</tr>
</tbody>
</table>
### ‘will’ and ‘would’

They are guided to guess the functions of will/ would in each sentence.

**Questions:**

1. ‘will’ is a modal auxiliary verb commonly found in English. What does it refer to / is associated with? “future”
2. ‘would’ is a modal auxiliary verb commonly found in English. What does it refer to / is associated with? “past”

### Grammatical characteristics of will

‘Will’ is mainly used as future marker.

Can you identify what tenses / grammatical structures is it associated with?

“Future simple, future continuous, future perfect, future in passive voice, 1st conditional, short answers”

### Sentences including both ‘will’ and ‘would’

<p>| | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>a. Today’s recipe will serve 4. (8)</td>
<td></td>
</tr>
<tr>
<td>b. The factory...will be shut. (11)</td>
<td></td>
</tr>
<tr>
<td>c. …other buskers will have taken the best pitches. (26)</td>
<td></td>
</tr>
<tr>
<td>d. …in the future automatic flying cars will be operating without human pilots. (31)</td>
<td></td>
</tr>
<tr>
<td>e. ..try some of our food! John: sounds great George.I definitely will (49)</td>
<td></td>
</tr>
<tr>
<td>f. ..if I say I can’t go, my friend will ask someone else.. (76)</td>
<td></td>
</tr>
</tbody>
</table>
### Senses of ‘will’

‘Will’ is used to express specific future functions. Can you identify what do the following 3 sentences express? (willingness, certainty, habit)

<table>
<thead>
<tr>
<th>Sentence</th>
<th>Expression</th>
</tr>
</thead>
<tbody>
<tr>
<td>a. the fair began yesterday, Saturday, and will run for two weeks (4)</td>
<td>WILLINGNESS</td>
</tr>
<tr>
<td>b. try some of our food! John: sounds great George. I definitely will. (49)</td>
<td>CERTAINTY</td>
</tr>
<tr>
<td>c. They want to appear ‘cool’ and will often conform to peer pressure, so as not to feel isolated… (10)</td>
<td>HABIT</td>
</tr>
</tbody>
</table>

### Practice

Now try to find in the COCA corpus at least three sentences with instances of ‘will’ expressing either ‘certainty’, ‘willingness’ or ‘habit’

Then try to find one sentence with ‘will’ in 1st conditional

Form full sentences out of the following words taken from CBS, CNN and ABC TV channels respectively, into affirmative, interrogative and negative forms using ‘will’.

<table>
<thead>
<tr>
<th>Sentence</th>
<th>Action</th>
<th>Form</th>
</tr>
</thead>
<tbody>
<tr>
<td>a. The president/address/nation/from/Oval office/at 8am eastern time.</td>
<td>enter the COCA corpus</td>
<td>affirmative</td>
</tr>
<tr>
<td>b. Bashar Assad/tell Time magazine/this week/Syria/withdraw/from Lebanon/this summer.</td>
<td>form the following query</td>
<td>interrogative</td>
</tr>
<tr>
<td>c. LISA McREE: You/come/back/and/talk/to/us?</td>
<td>display: KWIC</td>
<td>negative</td>
</tr>
<tr>
<td></td>
<td>word(s): will</td>
<td></td>
</tr>
<tr>
<td></td>
<td>sections: spoken + newspaper</td>
<td></td>
</tr>
</tbody>
</table>
### Stage

<table>
<thead>
<tr>
<th>Procedure</th>
<th>Resources</th>
</tr>
</thead>
<tbody>
<tr>
<td>Grammatical characteristics of ‘would’</td>
<td>Printed handouts from ECo.</td>
</tr>
<tr>
<td>‘Would’ is mainly used as the past tense of ‘will’, but it is also used to refer to the present.</td>
<td>Edited, full sentences only.</td>
</tr>
<tr>
<td>Can you tell whether the following sentences refer to the past or present and what is the one feature that differentiates them?</td>
<td>Sentences including both ‘will’ and ‘would’</td>
</tr>
<tr>
<td>“the infinitive that follows”</td>
<td></td>
</tr>
<tr>
<td>Read the sentences (e), (g), (i), (j) again and find one feature they have in common as far as the context in which ‘would’ is used is concerned.</td>
<td></td>
</tr>
<tr>
<td>Then, look carefully at sentences (g) and (i) and say how are they different? Can you explain the difference?</td>
<td></td>
</tr>
</tbody>
</table>

#### Example Sentences

- **a.** First of all would you say that unemployment was a factor in the rise in crime? (3)
- **b.** It said that she needed to get away for a few days and would come back later in the week. (5)
- **c.** A surfer’s greatest disappointment would be missing the opportunity to surf in the best weather conditions. (52)
- **d.** Lizzie: I would have been shouting at them from day one! (147)
- **e.** ….it would be better if young people learnt about first aid in their free time. (185)
- **f.** I was nervous at first and I’m glad I speak English or it could have been very difficult. (202)
| Senses of ‘would’ | Would is used to express specific future/present functions. Can you identify what do the following sentences express? (willingness, prediction) | a. A lot of animals would have become extinct by now without zoos to protect them! (15)  

b. I couldn’t understand why anyone would do it. (16)  

c. Building a sports centre would be really expensive too. (33) |
| --- | --- | --- |
| Practice | Now try to find in the COCA corpus at least three sentences with instances of ‘would’ expressing either ‘willingness’ or ‘prediction’ | • enter the COCA corpus  
• form the following query  
  ➢ display: KWIC  
  ➢ word(s): would  
  ➢ sections: spoken + newspaper |
Then try to find two sentences with ‘would’ in 2nd / 3rd conditional respectively

Form full sentences out of the following words taken from ABC, CNN and ABC TV channels respectively, using ‘would’.

a. if / you / can / see / your wife / what / you / say / to her?

b. KING: if / you / be / a prosecutor now / you / be / down?

c. I realize / it / be / sloppy / of you / to be / too / specific.

---

<table>
<thead>
<tr>
<th>Home practice/consolidation</th>
<th>Do the following activity (edited concordances from ECCo)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1. Fill in the following gaps with</td>
</tr>
<tr>
<td></td>
<td>a. Will / would</td>
</tr>
<tr>
<td></td>
<td>b. a verb in the appropriate form</td>
</tr>
<tr>
<td>a. I was wondering if you could answer a few of my queries. I .......... (be) very grateful.</td>
<td></td>
</tr>
<tr>
<td>b. When Frank .......... (not) (wake up) Ann got some help from the family dog.</td>
<td></td>
</tr>
<tr>
<td>c. How .......... (get) to the theatre? My car broke down and it’s in the shop.</td>
<td></td>
</tr>
<tr>
<td>d. If I hadn’t taken the holiday, I suppose I .......... (become) a successful university lecturer, but fate had other plans.</td>
<td></td>
</tr>
<tr>
<td>e. Tom: I thought it was a fair punishment, although I .......... (be) very embarrassed if it had been me.</td>
<td></td>
</tr>
<tr>
<td>f. Due to the train strike, we’ll be going by coach. It .......... (wait) for us here in the morning.</td>
<td></td>
</tr>
</tbody>
</table>

---

D.5 Main study lesson plan 4: “You can do it!”

Class: 2nd grade senior high school – experimental groups 1 & 2
Lesson duration: 50 minutes

Lesson topic: modal verbs (could, should)

Learning the functions of modal auxiliaries could/ should

<table>
<thead>
<tr>
<th>Stage</th>
<th>procedure</th>
<th>Resources</th>
</tr>
</thead>
<tbody>
<tr>
<td>Introducing the modal</td>
<td>Each student is given a handout of concordance output of could/should</td>
<td>Printed handouts from ECCo.</td>
</tr>
<tr>
<td>auxiliaries could and</td>
<td>instances from ECCo.</td>
<td>Edited, full sentences only.</td>
</tr>
<tr>
<td>should</td>
<td>They are guided to guess the functions of could/should in each sentence.</td>
<td>Sentences including both ‘could’ and ‘should’</td>
</tr>
<tr>
<td>Questions:</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1. What is common between</td>
<td></td>
<td></td>
</tr>
<tr>
<td>could and should?</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2. Could may express past</td>
<td></td>
<td></td>
</tr>
<tr>
<td>ability, formal request,</td>
<td></td>
<td></td>
</tr>
<tr>
<td>formal permission, general</td>
<td></td>
<td></td>
</tr>
<tr>
<td>possibility or opportunity.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Which sentence shows each</td>
<td></td>
<td></td>
</tr>
<tr>
<td>of the above senses?</td>
<td></td>
<td></td>
</tr>
<tr>
<td>characteristics</td>
<td></td>
<td></td>
</tr>
<tr>
<td>of ‘could’</td>
<td></td>
<td></td>
</tr>
<tr>
<td>a. Amazing! Could you tell</td>
<td></td>
<td></td>
</tr>
<tr>
<td>us something about some of</td>
<td></td>
<td></td>
</tr>
<tr>
<td>the 1980 eruption? (1)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>b. I could feel eyes staring</td>
<td></td>
<td></td>
</tr>
<tr>
<td>at me and could hear unusual</td>
<td></td>
<td></td>
</tr>
<tr>
<td>noises all around (2)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>c. The newspapers are</td>
<td></td>
<td></td>
</tr>
<tr>
<td>predicting that Becky</td>
<td></td>
<td></td>
</tr>
<tr>
<td>could be the richest woman</td>
<td></td>
<td></td>
</tr>
<tr>
<td>in England (7)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>d. Instead of going skiing.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sentence</td>
<td></td>
<td></td>
</tr>
<tr>
<td>----------</td>
<td></td>
<td></td>
</tr>
<tr>
<td>they could do something less adventurous (169)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>e. If these improvements were made, we could all be very proud of our block of flats. (218)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>f. The whole class could hear my stomach rumbling in the afternoon! (15)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>g. Could I have your passport and ticket, please? (102)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Senses of ‘should’</th>
</tr>
</thead>
<tbody>
<tr>
<td>3. <strong>should</strong> may express <strong>advice/suggestion, obligation or deduction</strong>. Which sentence shows each of the above senses?</td>
</tr>
<tr>
<td>h. Those with tickets for Manchester and North should change at Crewe station. (103)</td>
</tr>
<tr>
<td>i. Furthermore, parents should set limits to the hours their children watch TV.. (105)</td>
</tr>
<tr>
<td>j. It's a four-hour flight, so they should get in around three thirty. (167)</td>
</tr>
<tr>
<td>k. If a stranger approaches, they should go to the nearest safe place.. (158)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Senses of <strong>could</strong> and <strong>should</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td>4. <strong>what is common between sentences 5 and 11? What do they both express?</strong></td>
</tr>
<tr>
<td>Previously given printed concordances</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Previously given printed concordances</th>
</tr>
</thead>
<tbody>
<tr>
<td>235</td>
</tr>
<tr>
<td>Stage</td>
</tr>
<tr>
<td>-------</td>
</tr>
<tr>
<td>Practice: The senses of ‘could’ and ‘should’</td>
</tr>
<tr>
<td>Practice: The grammatical properties of ‘could’ and ‘should’</td>
</tr>
</tbody>
</table>

- a. I think she …..be allowed to skate.  
- b. They claim to show just how greedy Shelly……be  
- c. Critics insist the private sector……bring about changes must sooner.  
- d. Simon: And we ……..mention the book opens with your amazing preview.  
- e. Flatow: ……you describe the paintings for us?  
- f. Mr Doxiadis: First I …..say, first of all, my co-author is Christos Papadimitriou.  

- a. Jacob: I/ bet/ anything/ happen  
- b. Students / learn/ the color of clothing/ also / mean / taking sides
**Home practice/ consolidation**

Do the following activity (edited concordances from ECCo)

2. Fill in the following gaps with
   a. could / should
   b. a verb in the appropriate form

1. Animals...............( be allow) to go where they want, when they want.
2. Jenny: What other areas ........biometric technology ( be use) in?
3. Well, I believe that every child ........(attend) kindergarten.
4. Under no circumstances ........anyone (drive) or (be) outside once the hurricane has arrived.
5. I was having a wonderful time and ........(stay) there for ever
6. TV Pr: Professor, ........ (I ask) for your views on what makes a good teacher?
7. Mrs A: When do you think Sarah ........ (go) – I mean- if that's what we decide?
8. I knew that she ........ (give) me detailed directions, and she did.
9. So, ........ (we/ use) our microscopes now?
10. There are many who believe that the story is true. In fact, they ..........(be label) as 'mad' and might even have lost their jobs.

D.6 Main study lesson plan 5: “Shoot for the moon: you might get there”

Class: 2nd grade senior high school – experimental groups 1 & 2

Lesson duration: 50 minutes

Lesson topic: modal verbs (may, might)

Learning the functions of modal auxiliaries ‘may’/ ‘might’

<table>
<thead>
<tr>
<th>stage</th>
<th>Procedure</th>
<th>resources</th>
</tr>
</thead>
<tbody>
<tr>
<td>Introducing the modal auxiliaries ‘may’ and ‘might’</td>
<td>Each student is given a handout of concordance output of ‘may’/ ‘might’ instances from ECCo. They are guided to guess the functions of ‘may’/ ‘might’ in each sentence. Questions: 11. ‘may’ is a modal auxiliary verb commonly found in English. What does it refer to / is associated with? “present/ future”</td>
<td>Printed handouts from ECCo. Edited, full sentences only. Sentences including both ‘may’ and ‘might’ a) ..and I’d like, if I may, to give a few handy tips. (115) b) May you have a long a healthy life! (176) c) Dr Porentz: Can I answer? a. King: You may.</td>
</tr>
</tbody>
</table>
12. ‘might’ is a modal auxiliary verb commonly found in English. What does it refer to / is associated with? “past”

d) Thank you, God bless you, and may God bless the United States of America

e) Female 1: Hi, I'm from a local radio station and Jill thought you might be able to talk to me about your work... (5)

f) When you administer it to human patients, you might as well call it patient necrosis factor.

g) Senator: Well, if I might say, I want to repeat...the turtle eggs may be accidentally crushed by tourists... (1)

h) Doctor: If you give in to phobia, then you may let it rule your life. (9)

i) Excuse me miss, I'm Julia Marquez for WRSB, may I ask your name and what are you doing in this line? (31)

j) ..Teams may win prizes too (62)

---

3. ‘May’ is used to express specific functions.

Can you identify what ‘may’ expresses in the sentences a-j?

- general truth
- wish
- permission
- possibility
4. Might is used to express specific functions.
Can you identify what ‘might’ expresses in the sentences a-j

- (very formal) permission
- low possibility

| Practice | 1. Match the fragments to make full sentences | a. It is important to stay calm, as panic leads to rash actions which...
|          |                                             | b. In what ways...
|          |                                             | c. Indeed, as our understanding of them grows, they..
|          |                                             | d. James and his family started to worry that they..
|          |                                             | e. Clouseau is worried that he too..
|          |                                             | i.....may result in injury
|          |                                             | ii....might by killed
|          |                                             | iii...may have even bigger part to play in society
|          |                                             | iv.....might a job like this help students
<p>|          |                                             | v....might have a ghost up there |</p>
<table>
<thead>
<tr>
<th>Home practice/ consolidation</th>
<th>Do the following activity (edited concordances from ECCo)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>a. Fill in the following gaps with</td>
</tr>
<tr>
<td></td>
<td>a. may / might</td>
</tr>
<tr>
<td></td>
<td>b. the verb in the appropriate form</td>
</tr>
<tr>
<td>Steve: I like them like that! Still, I think you.....(be) right about the smart shirt.</td>
<td></td>
</tr>
<tr>
<td>You …... (care) about environment at home, but what happens on holiday?</td>
<td></td>
</tr>
<tr>
<td>If you put some fruit in a glass bowl, you ….. (dream) you are wearing the glass bowl on your head.</td>
<td></td>
</tr>
<tr>
<td>They …....(be) right, but I want to give it a go.</td>
<td></td>
</tr>
<tr>
<td>Mount Vesuvius …...(be) one of the most dangerous volcanoes in the world.</td>
<td></td>
</tr>
</tbody>
</table>
D.7  The traditional teaching methodology learning material

‘Take Off!’ B1+ - Student’s book, p.110
Complete each sentence with an item from the list.

- could • couldn’t have • didn’t have it
- had better not • must have • should
- wasn’t able • weren’t supposed

1. You were driving way too fast. You [could] have had an accident!
2. You [had better not] to tell anyone about that. It was a secret.
3. You [must have] forget to phone Liz.
4. I’m sorry, but I [wasn’t able to] to find the book you want.
5. They [should have] known what she was going to do. She’s done it before.
6. She [had better not] have told us what she wanted.
7. Bob [must have] driven there; he doesn’t know how to drive.
8. He [had better not] do much. I only asked for a glass of water!

Complete the second sentence so that it has a similar meaning to the first sentence, using the word given. Use between two and five words.

1. I couldn’t persuade Alice to come with us. WAS
   I [wasn’t able to] persuade Alice to come with us.
2. You oughtn’t to do that! HAD
   You [had better not] do that!
3. There was no need for you to say anything to Carl. TOLD
   You [shouldn’t have] told Carl anything.
4. You shouldn’t let other people borrow our camera. SUPPOSED
   You [mustn’t have] let other people borrow our camera.
5. Adam must still be too young to drive. ENOUGH
   Adam [isn’t supposed] to drive yet.

‘Take Off!’ B1+ - Student’s book, p.111
Grammar

Choose a, b or c.

1. You  ____ better not tell Anthony.
   a. would       b. had       c. must

2. You  ____ to invite Shelley. You know she’ll want to go.
   a. ought       b. had       c. should

3.  ____ you please give Julie this message?
   a. Must       b. May       c. Would

4. She  ____ be right, but I’m not sure.
   a. would       b. could       c. must

5. Have you ever thought that she  ____ not like your “jokes”?
   a. could       b. must       c. might

6.  ____ we order a pizza for dinner?
   a. Would       b. Needn’t       c. Shall

Change only the word(s) in colour so the sentences have the opposite meaning.

1. That can’t be the right answer!
   That  _________ be the right answer!

2. We had to get up early today.
   We  _________ get up early today.

3. Actually, I couldn’t find what he wanted!
   Actually, I  _________ find what he wanted!

4. You needn’t have told us that.
   You  _________ have told us that.
<table>
<thead>
<tr>
<th>Μορφή</th>
<th>Χρήση</th>
<th>Παράδειγμα</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>can</strong></td>
<td>1 ικανότητα/έλλειψη ικανότητας (παρών)</td>
<td><strong>Amy can/can’t/cannot</strong> drive. H Άμυ μπορεί/δεν μπορεί να οδηγάει.</td>
</tr>
<tr>
<td></td>
<td>2 δυνατότητα ή πιθανότητα/έλλειψη δυνατότητας ή πιθανότητας</td>
<td>We can/can’t/park there. Μπορούμε/δεν μπορούμε να παρκάρουμε εκεί.</td>
</tr>
<tr>
<td></td>
<td>3 άδεια</td>
<td>Can I go now? / Yes, you can. Μπορώ να πάω τώρα; / Ναι, μπορείς.</td>
</tr>
<tr>
<td></td>
<td>4 για να ζητήσουμε κάτι</td>
<td>Can you save me a seat? Μπορείς να μου φυλάξεις μια θέση;</td>
</tr>
<tr>
<td></td>
<td>5 για να προσφέρουμε για κάτι</td>
<td>Can I help? Μπορώ να βοηθήσω;</td>
</tr>
<tr>
<td><strong>can’t</strong></td>
<td>Βεβαιότητα (αρνητική) μορφή του must</td>
<td>Anne can’t be 30 years old! Η Αννη δεν μπορεί να είναι 30 ετών!</td>
</tr>
<tr>
<td><strong>could</strong></td>
<td>1 ικανότητα ή έλλειψη ικανότητας στο παρελθόν*</td>
<td>They could/couldn’t speak Italian. Μπορούσαν/δεν μπορούσαν να μιλάνεν Ιταλικά.</td>
</tr>
<tr>
<td></td>
<td>2 δυνατότητα ή πιθανότητα/έλλειψη δυνατότητας ή πιθανότητας</td>
<td>We could/couldn’t park there. Μπορούσαμε/δεν μπορούσαμε να παρκάρουμε εκεί.</td>
</tr>
<tr>
<td></td>
<td>3 άδεια**</td>
<td>Could I go now? / No, you can’t. Θα μπορούσα να πάω τώρα; / Όχι, δεν μπορείς.</td>
</tr>
<tr>
<td></td>
<td>4 για να ζητήσουμε κάτι ή να προσφέρουμε για κάτι</td>
<td>Could you help me? / Could I help? Θα μπορούσατε να με βοηθήσετε; / Θα μπορούσαν να βοηθήσουν.</td>
</tr>
<tr>
<td></td>
<td>5 για να προτείνουμε κάτι</td>
<td>You could try the fish. Θα μπορούσατε να τραυματίσετε το ψάρι.</td>
</tr>
<tr>
<td><strong>may</strong></td>
<td>1 δυνατότητα ή πιθανότητα (παρών/ μέλλον)</td>
<td>He may (not) come with us. Ο επάνω να (μη) έρθει μαζί μας.</td>
</tr>
<tr>
<td></td>
<td>2 άδεια (επίπονος λόγος)</td>
<td>May I come with you? / No, you may not. Μπορώ να έρθω μαζί σου; / Όχι, δεν μπορείς.</td>
</tr>
</tbody>
</table>

'‘Take Off!’ B1+ - study pack, p.75
<table>
<thead>
<tr>
<th>Морф</th>
<th>Χρήση</th>
<th>Παράδειγμα</th>
</tr>
</thead>
<tbody>
<tr>
<td>might</td>
<td>δυνατότητα ή πιθανότητα (παρόν/μέλλον)</td>
<td>He might (not) win the race. Τόσος να [μην] κερδίσει τον αγώνα.</td>
</tr>
<tr>
<td>must</td>
<td>1 υποχρέωση</td>
<td>You must write in pen. Πρέπει να γράφετε/γράψετε με στυλό.</td>
</tr>
<tr>
<td></td>
<td>2 βεβαιότητα</td>
<td>Ann must be at least 20 years old. Η Αν χρειάζεται να είναι πουλάγκοντας 20 χρόνια.</td>
</tr>
<tr>
<td></td>
<td>3 συμβουλή (άντονη)</td>
<td>You must go to the dentist. Πρέπει να πάει στον οδοντίατρο.</td>
</tr>
<tr>
<td>mustn’t</td>
<td>απαγόρευση</td>
<td>You mustn’t write in pencil. Δεν πρέπει να γράφετε/γράψετε με μολύβι.</td>
</tr>
<tr>
<td>should</td>
<td>1 συμβουλή</td>
<td>You should shouldn’t tell Susan. Θα πρέπει/θα πρέπει να το πεις στην Σουσίν.</td>
</tr>
<tr>
<td></td>
<td>2 υποχρέωση</td>
<td>You should stop writing now. Θα πρέπει να σταματήσετε να γράψετε τώρα.</td>
</tr>
<tr>
<td></td>
<td>3 πρόβλεψη</td>
<td>They should be arriving soon. Θα πρέπει να φτάνουν σύντομα.</td>
</tr>
<tr>
<td>ought to</td>
<td>1 συμβουλή</td>
<td>You ought oughtn’t to eat that egg. Θα πρέπει/δεν θα πρέπει να φας εκείνο το αβγό.</td>
</tr>
<tr>
<td></td>
<td>2 υποχρέωση</td>
<td>You ought to write in pen. Θα πρέπει να γράψετε/γράψετε με στυλό.</td>
</tr>
<tr>
<td>will</td>
<td>1 πρόβλεψη</td>
<td>The lim will won’t be boring. Η ταξία θα είναι/δεν θα είναι βορειά.</td>
</tr>
<tr>
<td></td>
<td>2 βεβαιότητα</td>
<td>Pam will be home soon. Η Πάμ θα είναι σύντομα σπίτι.</td>
</tr>
<tr>
<td></td>
<td>3 για να ζητήσουμε κάτι</td>
<td>Will you give this to Mark, please? Θα διάλεξες αυτό στον Μάρκ, σε παρακαλώ;</td>
</tr>
<tr>
<td>would</td>
<td>1 για να ζητήσουμε κάτι</td>
<td>Would you please come in? Θα μπορούσες να πας και να περάσετε μέσα;</td>
</tr>
<tr>
<td></td>
<td>2 συνήθεις του παρελθόντος</td>
<td>I would often play football as a kid. Συχνά έπαιζα ποδοσφαίρο, ως παιδί.</td>
</tr>
<tr>
<td></td>
<td>3 αόριστος του will</td>
<td>I knew she would pass the exam. Ηξέρα ότι θα περνούσε τις εξάμηνες.</td>
</tr>
<tr>
<td>shall</td>
<td>για να προσφεύγουμε για κάτι</td>
<td>Shall I get you some water? Να σου ζητήσουμε φέρει λίγο νερό;</td>
</tr>
</tbody>
</table>
* Χρησιμοποιούμε was/were able to όταν μπόταμε για ικανότητα που έχουμε σε μια συγκεκριμένη περίπτωση στο παρελθόν.
We were able to climb in through the window. Καταφέραμε να ακορμάσουμε και να μπούμε από το παράθυρο. (Δεν μπορούσαμε να χρησιμοποιήσουμε could στην περίπτωση.)

** Άδεια στο παρελθόν:
We could use the gym whenever we wanted to. Μας επιτρέπονταν να χρησιμοποιούμε το γυμναστήριο όσο έθελαμε.
Χρησιμοποιούμε could, όταν μπόταμε για άδεια που έχουμε γενικά στο παρελθόν.
We were allowed to use the gym last Friday. Μας επιτράπηκε να χρησιμοποιήσουμε το γυμναστήριο την προηγούμενη Παρασκευή.
Χρησιμοποιούμε was allowed to/ were allowed to, όταν μπόταμε για άδεια που έχουμε σε συγκεκριμένη περίπτωση στο παρελθόν.

**Other verb forms with similar functions** Άλλα ρήματα με παρόμοιες λειτουργίες

<table>
<thead>
<tr>
<th>Μορφή</th>
<th>Χρήση</th>
<th>Παράδειγμα</th>
</tr>
</thead>
<tbody>
<tr>
<td>have to</td>
<td>υποχρέωση/αναγκαίατη</td>
<td>You have to stop writing now. Πρέπει να σταματήσεις να γράφεις τώρα.</td>
</tr>
<tr>
<td>need to</td>
<td>υποχρέωση/αναγκαίατη</td>
<td>You need to stop writing now. Πρέπει να σταματήσεις να γράφεις τώρα.</td>
</tr>
<tr>
<td>don’t have to</td>
<td>ελλείψει υποχρέωσης/ αναγκαίατης</td>
<td>He doesn’t have to tell us everything. Δεν χρειάζεται να μας τα πει όλα.</td>
</tr>
<tr>
<td>don’t need to</td>
<td>ελλείψει υποχρέωσης/ αναγκαίατης</td>
<td>He doesn’t need to tell us everything. Δεν χρειάζεται να μας τα πει όλα.</td>
</tr>
<tr>
<td>needn’t</td>
<td>ελλείψει υποχρέωσης/ αναγκαίατης</td>
<td>He needn’t tell us everything. Δεν χρειάζεται να μας τα πει όλα.</td>
</tr>
<tr>
<td>had better</td>
<td>συμβουλή (όντως)</td>
<td>You had better tell us everything. Καλύτερα να μας τα πεις όλα.</td>
</tr>
<tr>
<td>Μορφή</td>
<td>Χρώμα</td>
<td>Παράδειγμα</td>
</tr>
<tr>
<td>-------------------------------</td>
<td>--------------------------------------------</td>
<td>------------------------------------------------</td>
</tr>
<tr>
<td>must have + μετοχή ασφαλτ.</td>
<td>βιβλιοπόθητα για κάτι που συνέβη στο παρελθόν</td>
<td>He must have been in school when it happened.</td>
</tr>
<tr>
<td>couldn't have + μετοχή ασφαλτ.</td>
<td>βιβλιοπόθητα για κάτι που δεν έχει συμβεί στο παρελθόν</td>
<td>They can't/couldn't have stolen the car.</td>
</tr>
<tr>
<td>may/might have + μετοχή ασφαλτ.</td>
<td>πιθανότητα στο παρελθόν</td>
<td>He may/might (not) have gone to Nick's house.</td>
</tr>
<tr>
<td>could have + μετοχή ασφαλτ.</td>
<td>πιθανότητα στο παρελθόν</td>
<td>We could have taken an earlier train.</td>
</tr>
<tr>
<td>ought to/should have + μετοχή ασφαλτ.</td>
<td>1 συμβουλή για κάτι που δεν έγινε</td>
<td>You ought to/should have tried the coat on before you bought it.</td>
</tr>
<tr>
<td></td>
<td>2 υποχώρηση που δεν εκπληρώθηκε</td>
<td>You ought to/should have tried the jacket on before you bought it.</td>
</tr>
<tr>
<td>should not have + μετοχή ασφαλτ.</td>
<td>για πράξη που δεν έπρεπε να έγινε</td>
<td>They ought to/should have called their teacher.</td>
</tr>
<tr>
<td>ought not to have + μετοχή ασφαλτ.</td>
<td>για πράξη που δεν έπρεπε να έγινε</td>
<td>They ought to/should have asked their teacher.</td>
</tr>
<tr>
<td>needn't have + μετοχή ασφαλτ.</td>
<td>για κάτι που δεν έπρεπε να συμβεί/υποχώρησε</td>
<td>She shouldn't have eaten your sandwiches.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Den χαίρεται να έχει φάει το ανάλογο σου.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>She ought not to have eaten your sandwiches.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Den χαίρεται να έχει φάει το ανάλογο σου.</td>
</tr>
</tbody>
</table>

Giving advice/Making suggestions Συμβουλεύοντες/Προτείνοντες

You must/could/should/ought to return that book to the library tomorrow.

Σήμερα έχετε να επιστρέψετε το βιβλίο στη βιβλιοθήκη αύριο.

Σήμερα, και θα πρέπει να επιστρέψετε το βιβλίο στη βιβλιοθήκη αύριο.

Σήμερα, θα πρέπει να επιστρέψετε το βιβλίο στη βιβλιοθήκη αύριο.

Σήμερα, θα πρέπει να επιστρέψετε το βιβλίο στη βιβλιοθήκη αύριο.
8  Grammar

1 Choose.

The lesson must/should be very boring!

2 Choose a or b.

1 I _often go fishing with my friends when I was young._
   a should  b would

2 My teacher reminded me that I still _her_ essay.
   a had to give  b must have given

3 You _do the next exercise unless you really want to._
   a don't need  b needn't

4 I think you _ask June for help._
   a ought  b should

5 If you didn't have a ticket, you _go in._
   a may not  b could not

6 You _be hungry. You've just eaten a big meal._
   a can't  b may not

7 You could ask Jim, but he _not know._
   a need  b might

8 The train was already leaving the station, but I _jump on board._
   a could  b was able to

3 Fill each gap with an item from the list.

- couldn't  - had  - must  - needn't
- should  - ought  - shall  - would

1 We're going to the park soon. _Shall_ we wait for you to join us?

2 He _be exhausted. He's just finished running a marathon._

3 I think you _to apologise to Fred for what you did._

4 I couldn't go out because I _to prepare our lunch._

5 You _help them. They can do it without you._

6 If you see Eva _you ask her to phone me this evening?_

7 She _have seen George. He's been out of town all day._

8 You _have asked Charlie to take you. He was going there anyway._

4 Match to make sentences.

1 Helen can play
   a seen your dog, but I'm not sure.

2 He'll be
   b supposed to say anything to Harold.

3 Douglas might have
   c to buy so many tickets.

4 You ought
   d go by coach if you like.

5 Ted can't have
   e both the flute and the harp.

6 They didn't need
   f not to do anything yet.

7 We could
   g known that since I hadn't told anyone!

8 She wasn't
   h able to get help from Wilma.
5. Which sentence in each group of three does not match the meaning of the other two?

1. a) You needn’t have written it out again.  
   b) You mustn’t have written it out again.  
   c) You didn’t have to write it out again.

2. a) He may be thinking of finding another job.  
   b) He could be thinking of finding another job.  
   c) He should be thinking of finding another job.

3. a) I must have locked the door.  
   b) I was supposed to lock the door.  
   c) I ought to have locked the door.

4. a) You ought to leave something for Mick to eat.  
   b) You’d better leave something for Mick to eat.  
   c) You could leave something for Mick to eat.

5. a) Can I watch the late film tonight?  
   b) Should I watch the late film tonight?  
   c) May I watch the late film tonight?

6. Rewrite the sentences correctly. Change only the words in colour.

1. Sue’s unhappy now; you oughtn’t have told her that.  
   Sue’s unhappy now; you shouldn’t have told her that.

2. Luckily, I could catch it before it fell.

3. You hadn’t better say anything to Joseph.

4. No, you don’t need to do any more.

5. I couldn’t keep the bird; I must have let it go.

7. Tick (✓) the sentences that contain the causative form.

1. He hasn’t had his hair cut recently.  
   ✓

2. I haven’t started work yet.  

3. She rarely gets annoyed.  

4. I should have the windows cleaned.  

5. I get my teeth checked twice a year.  

6. The emails have been sent.

8. Match to make sentences.

   a) of having the garden wall rebuilt.
   b) had the garden wall rebuilt.
   c) rebuild the garden wall.
   d) have the garden wall rebuilt.
   e) to be rebuilt.
   f) having the garden wall rebuilt.
   g) needs rebuilding.
   h) to rebuild the garden wall.

   1. I’m going to have them  
   2. I’m considering
   3. I’ll get them
   4. I’m thinking
   5. My garden wall
   6. They said I should
   7. I regret having
   8. My garden wall needs

‘Take Off!’ B1+ - workbook, p.67
E. Pre-, immediate post- and delayed post-test statistics

E1. Pre, immediate post and delayed post-test scores using simple counting

<table>
<thead>
<tr>
<th>Student</th>
<th>Pre-test</th>
<th>Immediate Post-test</th>
<th>Delayed Post-test</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>14</td>
<td>16</td>
<td>16</td>
</tr>
<tr>
<td>2</td>
<td>14</td>
<td>18</td>
<td>17</td>
</tr>
<tr>
<td>3</td>
<td>10</td>
<td>10</td>
<td>10</td>
</tr>
<tr>
<td>4</td>
<td>20</td>
<td>19</td>
<td>17</td>
</tr>
<tr>
<td>5</td>
<td>14</td>
<td>20</td>
<td>18</td>
</tr>
<tr>
<td>6</td>
<td>14</td>
<td>18</td>
<td>18</td>
</tr>
<tr>
<td>7</td>
<td>15</td>
<td>13</td>
<td>14</td>
</tr>
<tr>
<td>8</td>
<td>13</td>
<td>15</td>
<td>13</td>
</tr>
<tr>
<td>9</td>
<td>15</td>
<td>12</td>
<td>15</td>
</tr>
<tr>
<td>10</td>
<td>17</td>
<td>15</td>
<td>17</td>
</tr>
<tr>
<td>11</td>
<td>14</td>
<td>17</td>
<td>17</td>
</tr>
<tr>
<td>12</td>
<td>13</td>
<td>10</td>
<td>13</td>
</tr>
<tr>
<td>13</td>
<td>16</td>
<td>17</td>
<td>18</td>
</tr>
<tr>
<td>14</td>
<td>14</td>
<td>15</td>
<td>15</td>
</tr>
<tr>
<td>15</td>
<td>18</td>
<td>18</td>
<td>20</td>
</tr>
<tr>
<td>16</td>
<td>8</td>
<td>10</td>
<td>11</td>
</tr>
<tr>
<td>17</td>
<td>10</td>
<td>18</td>
<td>16</td>
</tr>
<tr>
<td>18</td>
<td>12</td>
<td>12</td>
<td>11</td>
</tr>
<tr>
<td>19</td>
<td>13</td>
<td>14</td>
<td>14</td>
</tr>
<tr>
<td>20</td>
<td>17</td>
<td>17</td>
<td>18</td>
</tr>
<tr>
<td>21</td>
<td>18</td>
<td>16</td>
<td>18</td>
</tr>
<tr>
<td>22</td>
<td>18</td>
<td>16</td>
<td>18</td>
</tr>
<tr>
<td>23</td>
<td>13</td>
<td>16</td>
<td>20</td>
</tr>
<tr>
<td>24</td>
<td>9</td>
<td>10</td>
<td>16</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>------------</td>
<td>--------</td>
<td>--------</td>
<td>--------</td>
</tr>
<tr>
<td>sum of points</td>
<td>339</td>
<td>362</td>
<td>380</td>
</tr>
<tr>
<td>Mean</td>
<td>14,13</td>
<td>15,08</td>
<td>15,83</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>B2- Exp.2</th>
<th>Scores</th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Student</td>
<td>Pre-test</td>
<td>Immediate</td>
<td>Post-test</td>
<td>Delayed post-test</td>
</tr>
<tr>
<td>1</td>
<td>16</td>
<td>21</td>
<td></td>
<td>19</td>
</tr>
<tr>
<td>2</td>
<td>20</td>
<td>21</td>
<td></td>
<td>17</td>
</tr>
<tr>
<td>3</td>
<td>18</td>
<td>20</td>
<td></td>
<td>18</td>
</tr>
<tr>
<td>4</td>
<td>18</td>
<td>17</td>
<td></td>
<td>16</td>
</tr>
<tr>
<td>5</td>
<td>14</td>
<td>8</td>
<td></td>
<td>17</td>
</tr>
<tr>
<td>6</td>
<td>15</td>
<td>16</td>
<td></td>
<td>18</td>
</tr>
<tr>
<td>7</td>
<td>8</td>
<td>11</td>
<td></td>
<td>12</td>
</tr>
<tr>
<td>8</td>
<td>15</td>
<td>18</td>
<td></td>
<td>17</td>
</tr>
<tr>
<td>9</td>
<td>15</td>
<td>16</td>
<td></td>
<td>13</td>
</tr>
<tr>
<td>10</td>
<td>12</td>
<td>9</td>
<td></td>
<td>15</td>
</tr>
<tr>
<td>11</td>
<td>15</td>
<td>15</td>
<td></td>
<td>15</td>
</tr>
<tr>
<td>12</td>
<td>9</td>
<td>16</td>
<td></td>
<td>11</td>
</tr>
<tr>
<td>13</td>
<td>8</td>
<td>17</td>
<td></td>
<td>13</td>
</tr>
<tr>
<td>14</td>
<td>12</td>
<td>12</td>
<td></td>
<td>12</td>
</tr>
<tr>
<td>15</td>
<td>12</td>
<td>18</td>
<td></td>
<td>11</td>
</tr>
<tr>
<td>16</td>
<td>17</td>
<td>17</td>
<td></td>
<td>17</td>
</tr>
<tr>
<td>17</td>
<td>10</td>
<td>15</td>
<td></td>
<td>13</td>
</tr>
<tr>
<td>18</td>
<td>12</td>
<td>14</td>
<td></td>
<td>15</td>
</tr>
<tr>
<td>19</td>
<td>14</td>
<td>14</td>
<td></td>
<td>14</td>
</tr>
<tr>
<td>20</td>
<td>14</td>
<td>8</td>
<td></td>
<td>10</td>
</tr>
</tbody>
</table>

<p>| sum of points | 274    | 303    | 293    |
| Mean          | 13,7   | 15,15  | 14,65  |</p>
<table>
<thead>
<tr>
<th>Student</th>
<th>Pre-test</th>
<th>Immediate post-test</th>
<th>Post-test</th>
<th>Delayed post-test</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>15</td>
<td>14</td>
<td>16</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>17</td>
<td>15</td>
<td>14</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>16</td>
<td>14</td>
<td>15</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>9</td>
<td>6</td>
<td>8</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>13</td>
<td>16</td>
<td>15</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>15</td>
<td>18</td>
<td>15</td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>16</td>
<td>18</td>
<td>15</td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>12</td>
<td>10</td>
<td>12</td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>14</td>
<td>16</td>
<td>14</td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>18</td>
<td>18</td>
<td>19</td>
<td></td>
</tr>
<tr>
<td>11</td>
<td>16</td>
<td>13</td>
<td>14</td>
<td></td>
</tr>
<tr>
<td>12</td>
<td>11</td>
<td>6</td>
<td>9</td>
<td></td>
</tr>
<tr>
<td>13</td>
<td>17</td>
<td>17</td>
<td>16</td>
<td></td>
</tr>
<tr>
<td>14</td>
<td>11</td>
<td>13</td>
<td>12</td>
<td></td>
</tr>
<tr>
<td>15</td>
<td>14</td>
<td>10</td>
<td>13</td>
<td></td>
</tr>
<tr>
<td>16</td>
<td>13</td>
<td>13</td>
<td>16</td>
<td></td>
</tr>
<tr>
<td>17</td>
<td>13</td>
<td>15</td>
<td>18</td>
<td></td>
</tr>
<tr>
<td>18</td>
<td>16</td>
<td>17</td>
<td>16</td>
<td></td>
</tr>
<tr>
<td>19</td>
<td>16</td>
<td>18</td>
<td>19</td>
<td></td>
</tr>
<tr>
<td>20</td>
<td>17</td>
<td>19</td>
<td>19</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Sum of points</th>
<th>289</th>
<th>286</th>
<th>295</th>
</tr>
</thead>
<tbody>
<tr>
<td>mean</td>
<td>14,45</td>
<td>14,3</td>
<td>14,75</td>
</tr>
</tbody>
</table>
E.2 Pre, immediate post and delayed post-test scores using normalized scores

<table>
<thead>
<tr>
<th></th>
<th>Experimental Group 1 (B1) knowledge level</th>
<th>Pre-test</th>
<th>Immediate Post-test</th>
<th>Delayed Post-test</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>(minimizing luck factor)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>11.69</td>
<td>14.35</td>
<td>14.35</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>11.69</td>
<td>17.01</td>
<td>15.68</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>6.37</td>
<td>6.37</td>
<td>6.37</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>18.34</td>
<td>18.34</td>
<td>15.68</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>11.69</td>
<td>19.67</td>
<td>17.01</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>11.69</td>
<td>17.01</td>
<td>17.01</td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>10.36</td>
<td>10.36</td>
<td>10.36</td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>10.36</td>
<td>13.02</td>
<td>10.36</td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>9.03</td>
<td>9.03</td>
<td>9.03</td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>13.02</td>
<td>13.02</td>
<td>13.02</td>
<td></td>
</tr>
<tr>
<td>11</td>
<td>11.69</td>
<td>15.68</td>
<td>15.68</td>
<td></td>
</tr>
<tr>
<td>12</td>
<td>6.37</td>
<td>6.37</td>
<td>6.37</td>
<td></td>
</tr>
<tr>
<td>13</td>
<td>14.35</td>
<td>15.68</td>
<td>15.68</td>
<td></td>
</tr>
<tr>
<td>14</td>
<td>11.69</td>
<td>13.02</td>
<td>13.02</td>
<td></td>
</tr>
<tr>
<td>15</td>
<td>17.01</td>
<td>17.01</td>
<td>17.01</td>
<td></td>
</tr>
<tr>
<td>16</td>
<td>3.71</td>
<td>6.37</td>
<td>6.37</td>
<td></td>
</tr>
<tr>
<td>17</td>
<td>6.37</td>
<td>17.01</td>
<td>14.35</td>
<td></td>
</tr>
<tr>
<td>18</td>
<td>9.03</td>
<td>9.03</td>
<td>7.7</td>
<td></td>
</tr>
<tr>
<td>19</td>
<td>10.36</td>
<td>11.69</td>
<td>11.69</td>
<td></td>
</tr>
<tr>
<td>20</td>
<td>15.68</td>
<td>15.68</td>
<td>15.68</td>
<td></td>
</tr>
<tr>
<td>21</td>
<td>14.35</td>
<td>14.35</td>
<td>14.35</td>
<td></td>
</tr>
<tr>
<td>22</td>
<td>14.35</td>
<td>14.35</td>
<td>14.35</td>
<td></td>
</tr>
<tr>
<td>23</td>
<td>10.36</td>
<td>14.35</td>
<td>14.35</td>
<td></td>
</tr>
<tr>
<td>24</td>
<td>5.04</td>
<td>6.37</td>
<td>6.37</td>
<td></td>
</tr>
<tr>
<td>Sum of points</td>
<td>264.6</td>
<td>315.14</td>
<td>301.84</td>
<td></td>
</tr>
<tr>
<td>--------------</td>
<td>-------</td>
<td>--------</td>
<td>--------</td>
<td></td>
</tr>
<tr>
<td>mean</td>
<td>11.03</td>
<td>13.13</td>
<td>12.58</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>Experimental Group 2 (B2) knowledge level</th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Pre-test</td>
<td>Immediate Post-test</td>
<td>Post-delayed Post-test</td>
<td>(minimizing luck factor)</td>
</tr>
<tr>
<td>1</td>
<td>14.35</td>
<td>21</td>
<td>18.34</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>19.67</td>
<td>21</td>
<td>15.68</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>17.01</td>
<td>19.67</td>
<td>17.01</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>15.68</td>
<td>15.68</td>
<td>14.35</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>3.71</td>
<td>3.71</td>
<td>3.71</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>13.02</td>
<td>14.35</td>
<td>14.35</td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>3.71</td>
<td>7.7</td>
<td>7.7</td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>13.02</td>
<td>17.01</td>
<td>15.68</td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>13.02</td>
<td>14.35</td>
<td>10.36</td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>5.04</td>
<td>5.04</td>
<td>5.04</td>
<td></td>
</tr>
<tr>
<td>11</td>
<td>13.02</td>
<td>13.02</td>
<td>13.02</td>
<td></td>
</tr>
<tr>
<td>12</td>
<td>5.04</td>
<td>14.35</td>
<td>7.7</td>
<td></td>
</tr>
<tr>
<td>13</td>
<td>3.71</td>
<td>15.68</td>
<td>10.36</td>
<td></td>
</tr>
<tr>
<td>14</td>
<td>9.03</td>
<td>9.03</td>
<td>9.03</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Pre-test</td>
<td>Immediate Post-test</td>
<td>Delayed Post-test</td>
<td></td>
</tr>
<tr>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td></td>
</tr>
<tr>
<td>15</td>
<td>9,03</td>
<td>17,01</td>
<td>7,7</td>
<td></td>
</tr>
<tr>
<td>16</td>
<td>15,68</td>
<td>15,68</td>
<td>15,68</td>
<td></td>
</tr>
<tr>
<td>17</td>
<td>6,37</td>
<td>13,02</td>
<td>10,36</td>
<td></td>
</tr>
<tr>
<td>18</td>
<td>9,03</td>
<td>11,69</td>
<td>11,69</td>
<td></td>
</tr>
<tr>
<td>19</td>
<td>11,69</td>
<td>11,69</td>
<td>11,69</td>
<td></td>
</tr>
<tr>
<td>20</td>
<td>3,71</td>
<td>3,71</td>
<td>3,71</td>
<td></td>
</tr>
</tbody>
</table>

Control Group (B3) knowledge level (minimizing luck factor)

<table>
<thead>
<tr>
<th></th>
<th>Pre-test</th>
<th>Immediate Post-test</th>
<th>Delayed Post-test</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>11,69</td>
<td>11,69</td>
<td>11,69</td>
</tr>
<tr>
<td>2</td>
<td>13,02</td>
<td>13,02</td>
<td>11,69</td>
</tr>
<tr>
<td>3</td>
<td>11,69</td>
<td>11,69</td>
<td>11,69</td>
</tr>
<tr>
<td>4</td>
<td>1,05</td>
<td>1,05</td>
<td>1,05</td>
</tr>
<tr>
<td>5</td>
<td>10,36</td>
<td>14,35</td>
<td>13,02</td>
</tr>
<tr>
<td>6</td>
<td>13,02</td>
<td>17,01</td>
<td>13,02</td>
</tr>
<tr>
<td>7</td>
<td>14,35</td>
<td>17,01</td>
<td>13,02</td>
</tr>
<tr>
<td>8</td>
<td>6,37</td>
<td>6,37</td>
<td>6,37</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>---</td>
<td>-----</td>
<td>-----</td>
<td>-----</td>
</tr>
<tr>
<td>9</td>
<td>11,69</td>
<td>14,35</td>
<td>11,69</td>
</tr>
<tr>
<td>10</td>
<td>17,01</td>
<td>17,01</td>
<td>17,01</td>
</tr>
<tr>
<td>11</td>
<td>10,36</td>
<td>10,36</td>
<td>10,36</td>
</tr>
<tr>
<td>12</td>
<td>1,05</td>
<td>1,05</td>
<td>1,05</td>
</tr>
<tr>
<td>13</td>
<td>15,68</td>
<td>15,68</td>
<td>14,35</td>
</tr>
<tr>
<td>14</td>
<td>7,7</td>
<td>10,36</td>
<td>9,03</td>
</tr>
<tr>
<td>15</td>
<td>6,37</td>
<td>6,37</td>
<td>6,37</td>
</tr>
<tr>
<td>16</td>
<td>10,36</td>
<td>10,36</td>
<td>10,36</td>
</tr>
<tr>
<td>17</td>
<td>10,36</td>
<td>13,02</td>
<td>13,02</td>
</tr>
<tr>
<td>18</td>
<td>14,35</td>
<td>15,68</td>
<td>14,35</td>
</tr>
<tr>
<td>19</td>
<td>14,35</td>
<td>17,01</td>
<td>17,01</td>
</tr>
<tr>
<td>20</td>
<td>15,68</td>
<td>18,34</td>
<td>18,34</td>
</tr>
<tr>
<td></td>
<td>216,5</td>
<td>241,8</td>
<td>224,5</td>
</tr>
<tr>
<td></td>
<td>10,83</td>
<td>12,09</td>
<td>11,22</td>
</tr>
</tbody>
</table>
E.3  Distribution of original and normalized scores for all groups’ pre, immediate – post and delayed –post tests

Distribution of original and normalized scores for the experimental group 1 pre-test

Distribution of original and normalized scores for the experimental group 1 delayed post-test
Distribution of original and normalized scores for the experimental group 2 pre-test

<table>
<thead>
<tr>
<th>students/degree</th>
<th>Orig.score</th>
<th>Norm.score</th>
</tr>
</thead>
<tbody>
<tr>
<td>0--3</td>
<td>0%</td>
<td>0%</td>
</tr>
<tr>
<td>3--6</td>
<td>0%</td>
<td>25%</td>
</tr>
<tr>
<td>6--9</td>
<td>13%</td>
<td>4%</td>
</tr>
<tr>
<td>9--12</td>
<td>21%</td>
<td>17%</td>
</tr>
<tr>
<td>12--15</td>
<td>29%</td>
<td>21%</td>
</tr>
<tr>
<td>15--18</td>
<td>17%</td>
<td>13%</td>
</tr>
<tr>
<td>18--21</td>
<td>4%</td>
<td>4%</td>
</tr>
</tbody>
</table>

$s = 3,26$
$s = 4,94$

Avg 13,70

Distribution of original and normalized scores for the experimental group 2 post-test

<table>
<thead>
<tr>
<th>students/degree</th>
<th>Orig.score</th>
<th>Norm.score</th>
</tr>
</thead>
<tbody>
<tr>
<td>0--3</td>
<td>0%</td>
<td>0%</td>
</tr>
<tr>
<td>3--6</td>
<td>0%</td>
<td>13%</td>
</tr>
<tr>
<td>6--9</td>
<td>13%</td>
<td>4%</td>
</tr>
<tr>
<td>9--12</td>
<td>8%</td>
<td>13%</td>
</tr>
<tr>
<td>12--15</td>
<td>17%</td>
<td>21%</td>
</tr>
<tr>
<td>15--18</td>
<td>33%</td>
<td>21%</td>
</tr>
<tr>
<td>18--21</td>
<td>13%</td>
<td>13%</td>
</tr>
</tbody>
</table>

$s = 3,81$
$s = 5,07$

Avg 15,15

Avg 13,22
Distribution of original and normalized scores for the experimental group 2 delayed post-test

Distribution of original and normalized scores for the control group pre-test
Distribution of original and normalized scores for the control group post-test

Distribution of original and normalized scores for the control group delayed post-test