

PEOPLE'S DEMOCRATIC REPUBLIC OF ALGERIA
Ministry of Higher Education and Scientific Research
Ferhat Abbas University, Setif
Faculty of Letters and Languages
Department of English Language and Literature

**THE RELATIONSHIP BETWEEN LEARNERS' LEXICAL
COVERAGE AND THE READABILITY LEVELS
OF THE ALGERIAN ENGLISH TEXTBOOKS**

By

Saad TORKI

**Thesis submitted in candidature for the degree of
doctorate "ès-sciences" in Applied Linguistics**

Supervisor: Prof. H. SAADI

Co-Supervisor: Prof. M. MEHRACH

Board of Examiners:

Chairman: Prof. S. KESKES

Prof. University of Setif

Supervisor: Prof. H. SAADI

Prof. University of Constantine

Co-Supervisor: Prof. M. MEHRACH

Prof. University of Tetouan, Morocco

Member: Dr. H. HAMADA

M.C. University of Constantine

Member: Dr. S. LARABA

M.C. University of Constantine

Member: Dr. A. NEMOUCHI

M.C. University of Constantine

2012

PEOPLE'S DEMOCRATIC REPUBLIC OF ALGERIA
Ministry of Higher Education and Scientific Research
Ferhat Abbas University, Setif
Faculty of Letters and Languages
Department of English Language and Literature

**THE RELATIONSHIP BETWEEN LEARNERS' LEXICAL
COVERAGE AND THE READABILITY LEVELS
OF THE ALGERIAN ENGLISH TEXTBOOKS**

By

Saad TORKI

**Thesis submitted in candidature for the degree of
doctorate "ès-sciences" in Applied Linguistics**

Supervisor: **Prof. H. SAADI** Co-Supervisor: **Prof. M. MEHRACH**

Board of Examiners:

Chairman:	Prof. S. KESKES	Prof. University of Setif
Supervisor:	Prof. H. SAADI	Prof. University of Constantine
Co-Supervisor:	Prof. M. MEHRACH	Prof. University of Tetouan, Morocco
Member:	Dr. H. HAMADA	M.C. University of Constantine
Member:	Dr. S. LARABA	M.C. University of Constantine
Member:	Dr. A. NEMOUCHI	M.C. University of Constantine

DEDICATION

To
My mother

To

*The memory of my father whose last words as he lay dying were:
"My son, never quit learning".*

*The memory of my father and mother-in-law who helped me to
heed my father's advice.*

To my sisters,

To Afef - Zineb,

Mohamed-Abdallah,

Takj -Eddine,

and their Mum

To My brothers and sisters-in-law

ACKNOWLEDGEMENTS

First of all, I would like to express my deepest gratitude to Allah for His innumerable graces on me, one of which is the completion of this work.

Second, my most profound feelings and unspeakable, ineffable, and immense gratitude go to my mother who, though widowed at an early age, sacrificed her youth, life and energy and went out to work for my education.

Third, the journey that produced this study owes much to many people. Some of them can be named, others cannot, but all of them have my deepest appreciation and deserve to be acknowledged for their support in this effort.

I wish to convey my gratitude to my supervisor Professor. Hacene Saadi, Mentouri University, Constantine. I consider myself very fortunate and privileged to have conducted this research under his supervision. His comments and criticisms have played a vital part in the development of this work. I have a very special debt to him for the understanding and reassurance he gave me on many occasions I was about to give up. Without his unfailing support, kindness and ceaseless patience, this thesis would have never been written. A similar role was held by Professor Mohamed Mehrach, University of Tetouan, Morocco. I would like him to find here my most sincere thanks for his patient mentoring, guidance, support, assistance and hospitality.

Special thanks are due to the members of the Board of Examiners who have kindly accepted to examine the present thesis.

A bundle of thanks is also due to Mohammed Berrahal and Rafik Mosbah for their invaluable help and moral support.

It would be remiss of me not to express my gratitude to Abdallah Khababa and Mabrouk Korichi for their contribution in the development of a software.

Last but not the least, my deepest feelings go to my wife who constantly prodded me into working on this research, took on her all the responsibilities of the home in order to set me free for research and who endured my bad temper when things went wrong. I want to convey to her - publicly - my gratitude and my apologies for having stolen time that was hers and for the so many times I have offended her.

Needless to say, I am solely responsible for all the flaws and mistakes to be found herein.

Abstract

Key words: readability – lexical coverage –reading comprehension – EFL textbooks

In Algerian schools, textbooks provide the major if not the only written lexical input for students in classrooms. This study examined the seven English as a Foreign Language textbooks in use in these schools in order to investigate their lexical coverage and readability. The research project was designed to i) compile a textbook corpus ii) compare lexical coverage of the textbooks, and iii) assess readability. The main concern was to determine whether learners' lexical coverage was **at** the textbook readability level (Independent Reading Level), above it (Instructional Reading Level), or below it (Frustrational Reading Level).

Furthermore, the list of all lexical items occurring in the seven textbooks was compared to West's *General Service List*, and Coxhead's *Academic Word List* to assess whether textbooks provide sufficient, useful and appropriate vocabulary items.

Another purpose was to provide English teachers and educationalists in general with a means (computer software) for comparing the vocabulary levels of reading materials and textbooks destined to Algerian learners of English in order to determine what the readability and vocabulary levels are, and what additional vocabulary is required for students to reach the 95% rate of comprehension.

The methodology adopted to explore the lexical coverage was characterised by a multi-instrument computer-based approach involving computer software. The sets of textbooks were processed to generate textbook word lists.

Results have shown that all the EFL textbooks in use have low lexical coverage and readability, putting them at the frustrational level. Moreover, except for the first three textbooks there was a total discrepancy in terms of lexical coverage between the other four textbooks as the rate of common vocabulary across the textbooks was very low.

Comparison of the lexical coverage of the seven textbooks to standard vocabulary lists have revealed that Algerian students are not learning sufficient, useful and appropriate vocabulary, as Algerian learners are exposed to a low proportion of high frequency words. The study ends with implications and recommendations as to how remedy to the problem.

List of abbreviations

1 A.M: Première Année Moyenne (first year of the middle school).

1 A.S: Première Année Secondaire (first year of the secondary school).

2 A.M: Deuxième Année Moyenne (second year of the middle school).

2 A.S: Deuxième Année Secondaire (second year of the secondary school).

3 A.M: Troisième Année Moyenne (third year of the middle school).

3 A.S: Troisième Année Secondaire (third year of the secondary school).

4 A.M: Quatrième Année Moyenne (fourth year of the middle school).

AWL: The General Service List

BNC HFWL: British National Corpus High Frequency Word List

COBUILD: Collins Birmingham University International Language Database

E.F.L: English as a Foreign Language.

E.S.L: English as a Second Language.

F.L: Foreign Language.

GSL: General Service List

L1: First language

L2: Second language

LIH: Linguistic Interdependence Hypothesis

UWL: The University Word List

List of Tables

Table 1: Middle school textbooks	224
Table 2: Secondary school textbooks	235
Table 3: Sample output from FREQUENCY	249
Table 4: Labelling System	253
Table 5: Corpus compilation process	254
Table 6: Examples of types of words Excluded from the headword list	256
Table 7: Answering the research questions.....	257
Table 8: The distribution of word families in the seven textbooks	261
Table 9: Overall results: Middle school textbooks comparison	263
Table 10: New lexical items in middle school textbooks, Books 2, 3, and 4.	265
Table 11: Lexical coverage in middle school textbooks.....	268
Table 12: Distribution of headwords in the secondary school textbooks	273
Table 13: Overall results of secondary school textbooks comparison	275
Table 14: New lexical items in secondary school textbooks	277
Table 15: Lexical coverage in secondary school textbooks	279
Table 16: Lexical coverage in all textbooks	281
Table 17: Rank order of textbooks by lexical coverage	285
Table 18: Results yielded by RANGE	289
Table 19: Results from RANGE and Lextutor software	293

List of Figures

Figure	Page
Figure 1: File structure in <i>Spotlight on English Book 1</i>	227
Figure 2: File structure in <i>Spotlight on English Book 2</i>	231
Figure 3: File structure in <i>Spotlight on English Book 3</i>	233
Figure 4: File structure in <i>At the Crossroads</i>	239
Figure 5: Number of word families and word types in the BNC high-frequency word lists	248
Figure 6: Snapshot of The Compleat Lexical Tutor	250
Figure 7 Snapshot of the Lextutor results	248

List of Histograms

Histogram	Page
Histogram 1: Distribution of word families in the seven textbooks/levels	261
Histogram 2: Overall results of middle school textbooks comparison	264
Histogram 3: New lexical items in middle school textbooks, Books 2, 3, and 4	265
Histogram 4: Lexical coverage in middle school textbooks	268
Histogram 5: Distribution of headwords in the secondary school textbooks	274
Histogram 6: Overall results of secondary school textbooks comparison	276
Histogram 7: New lexical items in secondary school textbooks: Books 5, 6, and 7 expressed in percentage of words	278
Histogram 8: Lexical coverage in secondary school textbooks	280
Histogram 9: Lexical coverage in all textbooks	282

List of Charts

	Chart	Page
Chart 1:	Distribution of word families in the seven textbooks/levels	261
Chart 2:	Overall results of middle school textbooks comparison	264

List of Graphs

	Graph	Page
Graph 1:	Lexical coverage in middle school textbooks	269
Graph 2:	Lexical coverage in all textbooks	283

List of Appendices

Appendix	page
Appendix I: Academic Word List	351
Appendix II: The General Service List.....	366
Appendix III: List of function words ignored in the analysis.....	372
Appendix IV: List of words in the Stop-list	374
Appendix V: List of all headwords in the seven textbooks	429
Appendix VI: List of words from the First -1000 word list of most frequent words occurring all Algerian textbooks	447
Appendix VII: List of words from Second -1000 word list of most frequent words occurring all Algerian textbooks.....	451
Appendix VIII: Words from the Academic Word List occurring in the seven Algerian textbooks	454

TABLE OF CONTENTS

CHAPTER 1

INTRODUCTION AND STATEMENT OF AIMS

Rationale:.....	1
Statement of the Problem:	3
Research questions:	13
Importance of the study:.....	13
Theoretical framework:	17
Hypotheses:	21
Methodology:	22
Limitations of the study:.....	23
Structure of the research:.....	23

CHAPTER 2

READING THEORIES AND MODELS

Introduction	26
2.1. What is reading?.....	26
2.2. Historical perspective on reading.....	30
2.3. Schema theory	34
2.3.1. The definition of schema	34
2.3.2. Types of schema	36

2.3.2.1. Linguistic schemata.....	36
2.3.2.2. Formal schemata.....	37
2.3.2.3. Content schemata	38
2.4. Reading models	38
2.4.1. What is a reading model?.....	38
2.4.2. The bottom-up reading model.....	39
2.4.2.1. What is a bottom-up reading model?	39
2.4.2.2. Critique of the bottom-up reading model	40
2.4.3. The top-down reading model.....	43
2.4.3.1. What is a top-down reading model?.....	43
2.4.3.2. Critique of the top-down reading model	46
2.4.4. The interactive models.....	47
2.4.4.1. What is an interactive reading model?	47
2.4.4.2. Transactional Socio-Psycholinguistic Model of Reading	49
2.4.4.3. Rumelhart’s interactive reading model	52
Conclusion.....	54

CHAPTER 3

READABILITY

Introduction	56
3.1. Measuring text difficulty	57
3.1.1. Subject assessment.....	57

3.1.2. Question and answer techniques.....	58
3.1.3. Graphs and charts.....	58
3.1.4. Cloze Procedure.....	58
3.1.5. Readability formulas.....	59
3.2. What is readability?.....	60
3.3. Readability: historical overview.....	63
3.4. Factors influencing readability.....	69
3.5. Classic readability formulas.....	70
3.5.1. The Flesch Reading Ease formula.....	72
3.5.2. Flesch-Kincaid Grade Level Formula.....	72
3.5.3. The Dale-Chall Formula.....	73
3.5.4. The Gunning Fog Formula.....	74
3.5.5. The Readability Graph.....	75
3.5.6. Coleman Liau Formula.....	75
3.5.7. Automated Readability Index Formula.....	76
3.5.8. The Bormuth Formula.....	76
3.5.9. The SMOG Formula.....	77
3.6. Computational approaches.....	78
3.6.1. The Lexile Formula.....	78
3.6.2. The Strathclyde Readability Measure.....	80
3.7. Limitations of Readability Formulas.....	81

3.8. Validating Formulas for EFL	84
3.8.1. Hamsik's Study (1984).....	84
3.8.2. Brown's Study (1998)	85
3.8.3. The Miyazaki Study.....	86
3.9. Beyond Readability	87
Conclusion.....	88

CHAPTER 4

SECOND/FOREIGN LANGUAGE READING COMPREHENSION

Introduction	90
4.1. What is reading comprehension?.....	91
4.2. Purposes for reading: why we read	93
4.3. Types of reading.....	94
4.3.1. Skimming.....	95
4.3.2. Scanning.....	96
4.3.3. Extensive reading.....	97
4.3.4. Intensive reading.....	97
4.3.5. Careful reading	98
4.3.6. Browsing.....	98
4.3.7. Reading for general comprehension	98
4.4. Levels of Comprehension.....	99
4.5. Components of reading ability: skills involved.....	104

4.5.1. Automatic Word Recognition Skills.....	107
4.5.2. Vocabulary and structural (grammar) knowledge	107
4.5.3. Knowledge of discourse structure.....	108
4. 5.4. Content/World Background Knowledge	109
4.5.5. Synthesis and Evaluation Skills.....	109
4.5.6. Metacognitive Knowledge and Skills	109
4.6. Relationship between L1 and L2 reading comprehension	110
4.6.1. Reading problem or language problem.....	112
4.6.2. The Linguistic Threshold Hypothesis.....	114
4.6.3. The Linguistic Interdependence Hypothesis.....	117
4.7. Factors affecting foreign second/foreign language comprehension	124
4.7.1. Reader variables.....	126
4.7.1.1. Cognitive abilities	126
4.7.1.2. Affect	127
4.7.1.3. Culture/Background knowledge	127
4.7.2. Text variables.....	130
4.7.2.1. The Effect of Typographical features on Reading Comprehension	131
4.7.2.2. The Effect of Text Cohesion on Reading Comprehension	132
4.7.2.3. The Effect of Text Coherence on Reading Comprehension	133
4.7.2.4. The Effect of Text Type and Genre on Reading Comprehension	134
4.7.2.5. The Effect of Text Organization on Reading Comprehension	136

4.7.2.6. The Effect of Text length on Reading Comprehension	137
4.7.3. Writer variable.....	139
4.7.4. Context variable.....	140
Conclusion.....	141

CHAPTER 5

VOCABULARY AND WORD KNOWLEDGE

Introduction	143
5.1. What is a word?	144
5.1.1. Definition by orthography	145
5.1.2. Definition by phonology	146
5.1.3. Definition by meaning	147
5.2. Key terms and concepts.....	149
5.3. What is a lexical item/lexeme?.....	155
5.3.1. Citation forms	157
5.3.2. Grammatical word-forms.....	157
5.3.3. Inflection and derivation.....	160
5.3.4. Multi-part and discontinuous words	162
5.3.5. Content words and grammatical words	163
5.3.6. Clitics.....	164
5.3.7. Short forms	165
5.3.7.1. Abbreviations.....	165

5.3.7.2. Logograms	166
5.3.7.3. Contractions.....	166
5.3.7.4. Acronyms and Initialisms	167
5.3.7. 5. Clipped forms	168
5.3.8. Types of words	170
5.4. What is Vocabulary/lexicon?	172
5.4.1. Types of Vocabulary.....	173
5.4.1.1. General vocabulary	173
5.4.1.2. Active and Passive Vocabulary	173
5.4.1.3. Recognition and Reproduction Vocabulary.....	174
5.4.1.4. Writing, Reading, and Speaking Vocabulary	175
5.4.1.5. Academic and Non-Academic Vocabulary	175
5.4.1.6. Form and Content Vocabulary.....	175
5.4.1.7. Concrete and Abstract Vocabulary	177
5.4.1.8. Basic Vocabulary.....	178
5.4.1.9. Graded Vocabulary	179
5.4.1.10. Recall Vocabulary	179
5.4.1.11. Dialect Vocabulary	180
5.4.1.12. Archaic Vocabulary.....	180
5.4.1.13. Technical Vocabulary	181
5.5. What constitutes word knowledge?.....	182

Conclusion..... 186

CHAPTER 6

LEXICAL COVERAGE AND READING COMPREHENSION

Introduction 188

6.1. Reading comprehension and vocabulary knowledge 188

 6.1.1. Empirical Evidence..... 190

 6.1.2. Theoretical Perspectives 191

 6.1.2.1. Instrumentalist Hypothesis 192

 6.1.2.2. Knowledge Hypothesis 192

 6.1.2.3. Aptitude Hypothesis 194

 6.1.2.4. Access Hypothesis 194

6.2. Lexical Coverage and Reading Comprehension 195

 6.2.1. Key terms and concepts 195

 6.2.1.1. Word frequency/list 195

 6.2.1.2. Sight vocabulary 196

 6.2.1.3. Type/token 197

 6.2.1.4. Lexical density..... 198

 6.2.1.5. Word family..... 198

6.3. What is text lexical coverage? 199

6.4. How has text coverage been assessed?..... 200

6.5. Which words should second/foreign language learners know? 201

6.6. How many words should second/foreign language learners know?	204
6.7. Repetition and Vocabulary Learning.....	209
6.8. What vocabulary does a language learner need?.....	211
6.8.1. The Teachers Word Book	212
6.8.2. The American Heritage Word Frequency Book	212
6.8.3. The General Service List (West, 1953)	212
6.8.4. The University Word List (UWL)	215
6.8.5. The Academic Word List.....	215
6.8.6. Core vocabulary.....	216
Conclusion.....	218

CHAPTER 7

METHODS AND PROCEDURES

Introduction	220
7.1. Corpora and textbook corpus research	220
7.2. Description of the Textbooks:	222
7.2.2. Middle school textbooks.....	222
7.2.2.1. Spotlight on English Book One (first year).....	223
7.2.2.1.1. General layout	224
7.2.2.1.2. File Structure Description	225
7.2.2.2. Spotlight on English Book Two (second year)	227
7.2.2.2.1. General layout	227

7.2.2.2.2. File Structure Description	227
7.2.2.3. Spotlight on English Book Three (third year, revised edition)	230
7.2.2.3.1. General layout	231
7.2.2.3.2. File Structure Description	231
7.2.2.4. On the Move (fourth year).....	233
7.2.2.4.1. General layout	233
7.2.2.4.2. File Structure Description	233
7.2.3. Secondary school textbooks.....	234
7.2.3.1. At the Crossroads (first year)	235
7.2.3.1.2. General layout	235
7.2.3.1.3. Unit Structure Description.....	236
7.2.3.2. Getting Through (second year).....	237
7.2.3.2.1. General layout	237
7.2.3.2.2. Unit Structure Description.....	239
7.2.3.3. New Prospects (third year)	240
7.2.3.3.1. General layout	240
7.2.3.3.2. Unit Structure Description.....	240
7.3. Method of analysis	241
7.3.1. Instruments.....	241
7.3.1.1. Range and Frequency programs for Windows based PCs	242
7.3.1.2. The Compleat Lexical Tutor	248

7.3.1.3. TextCompare software	249
7.3.1.4. TextMaster.....	250
7.3.2. Procedure	250
7.3.2.1. Preparing textbooks for analysis	250
7.3.2.1.1. Creating Word Lists from English Textbooks	250
7.3.2.1.1.1. Step 1: Scanning the books	251
7.3.2.1.1.2. Step 2: Cleansing the untreated text	251
7.3.2.1.1.3. Step 3: Generating word lists from textbooks	253
7.3.2.1.1.4. Step 4: Refining the word lists	253
7.4. How to answer the research questions.....	255
Conclusion.....	257

CHAPTER 8

DATA ANALYSIS AND RESULTS

Introduction	258
8.1. Lexical coverage and readability in middle school textbooks.....	259
8.1.1. Data analysis	259
8.1.2. Interpretation.....	269
8.2. Lexical coverage and readability in secondary school textbooks	271
8.2.1. Data analysis	271
8.2.2. Interpretation.....	279
8.3. Putting it all together: Lexical coverage and readability in middle and secondary	

school textbooks	280
8.4. Do Algerian EFL textbooks provide sufficient, useful and appropriate vocabulary items?.....	286
8.4.1. Results.....	288
8.4.2. Interpretation.....	293
Conclusion.....	294

CHAPTER 9

FINDINGS, IMPLICATIONS, LIMITATIONS, AND RECOMMENDATIONS

Introduction	295
9.1. The overall picture of the thesis	295
9.1. Conclusion.....	301
9.2. Pedagogical implications.....	302
9.3. Limitations of the study.....	307
9.4. Recommendations for Future Research.....	308
Concluding remark	309
BIBLIOGRAPHY	1
Relevant websites	37
APPENDICES	38
Academic Word List.....	39
Sublist 3 of Academic Word List.....	41
Sublist 4 of the Academic Word List	42

Sublist 5 of the Academic Word List	44
Sublist 6 of the Academic Word List	46
Sublist 7 of the Academic Word List	48
Sublist 8 of the Academic Word List	50
Sublist 9 of the Academic Word List	51
Sublist 10 of the Academic Word List	53
APPENDIX II.....	54
APPENDIX III	60
APPENDIX IV	61
APPENDIX V List of all headwords in the seven textbooks.....	61
APPENDIX VI.....	61
APPENDIX VII.....	61
APPENDIX VIII	61

CHAPTER 1

INTRODUCTION AND STATEMENT OF AIMS

The purpose of this introductory chapter is to present the rationale for the study and its aims, review some of the important studies upon which the theoretical framework of the present study is based, and state the problem to be addressed as well as the questions to be answered and the hypotheses considered. In addition, it spells out the importance of the study. The terms used are defined and the limitations of the study are set. Finally, it introduces the structure of the thesis.

Rationale:

It is a byword that reading is one of the basic ways of acquiring information in our society and, as Grabe (2002, p.2) noted, it is “one of the most important language skills in academic settings”. Another scholar, Alderson, (2000) contends that:

... reading is an essential skill for English as a Second/Foreign Language (ESL/EFL¹) students; ... and the most important skill to master. With strengthened reading skills, ESL/EFL readers will make greater progress and attain greater development in all academic areas.

(p. 20)

¹ ‘Second language’ is defined throughout the whole study after Stern (1983) as referring to “the chronology of language learning; a second language being any language acquired after the native” (p. 12). Hence, no distinction is made throughout this study between English as a Second Language (ESL) and English as a Foreign Language (EFL). The two terms are used interchangeably.

It is also almost a truism that the information age we are living in requires that our students be proficient readers in English. The importance of this language is not to be argued. However, this importance is well expressed by Zabrocky and Moore (1999) who stated that:

Adolescents entering the adult world in the 21st century will read and write more than at any other time in human history They will need literacy to cope with the flood of information they will find everywhere they turn In a complex and sometimes even dangerous world, their ability to read will be crucial. (p. 99)

However, reading English is a complex process for native speakers, let alone for non-native speakers. Taylor and Hiebert (1994) believe that a substantial number of native students have difficulties with reading. With regard to non-native speakers (ESL/EFL learners), Alderson (1984) reported that reading in a language which is not the learner's first language is a source of considerable difficulty. Furthermore, Shokrpour (2005) contended that: "reading difficulty has been and continues to be one of the most important aspects of reading comprehension" (para.7). Having reading comprehension difficulties generally means having trouble with one or more of the foundation skills that make up the complex act of reading. These difficulties can be attributed to many factors. Attempting to address the difficulty which students encounter, Haynes (1989) suggests that to help second language readers comprehend English text quickly and accurately, language teachers need a good understanding of

where reading difficulties are apt to occur and of what can be done in the short and long run to help students overcome these difficulties. Another reading specialist, Grabe (2002) stated that: “If teachers and curriculum developers are to help students make significant progress in reading instruction, they need to understand how reading works. Only in this way can we make informed decisions to guide effective reading instruction” (p. 5).

Statement of the Problem:

One of the problems I have encountered repeatedly over the years I have been teaching English is how to tell whether a particular piece of writing is likely to be comprehensible to a particular group of readers. This problem is more acute when it comes to choosing a reading passage for the design of examinations. In fact, a daily concern for frontline teachers and a recurring question is: is the level of vocabulary in this reading passage appropriate for my learners? Three possible solutions to the problem could be suggested in this particular context.

A first solution is to guess, to make an estimate of the comprehensibility of the piece of writing with the skills developed from personal experience and the feedback from readers about similar texts. In the absence of reliable means, EFL teachers in Algeria rely inescapably on this educated guess for their decisions on the comprehensibility of a particular piece of writing for their learners. No matter how accurate one's memory may be, there are always occasions when one cannot tell exactly whether a certain word or expression has been taught or when it was taught, or in which

lesson it was first introduced. This problem is even worse in secondary schools and universities as there is a wide gap between the middle school, the secondary school, and the university. Secondary school teachers and university lecturers are generally unaware of the contents of textbooks used in the school their students come from, especially with the introduction of new textbooks at both the middle and secondary levels².

A second solution could be administrating a pilot comprehension test to find out how well the intended groups of readers may understand the reading materials. In such a case a pilot comprehension test should be designed to find out how well the intended groups of readers may understand the reading materials. Obviously such a test which would be a comprehensibility test is time consuming, expensive and requires expertise, i.e. an adequate training in test construction and administration or at least knowing how to test comprehension. This makes it out of the reach of teachers.

A third solution to the problem adopted in some countries such as the USA, England, and Canada, is to refer to the ranking or grading index of the reading material. Unfortunately, such a grading is not available in published reading material in Algeria.

In the Algerian situation, the textbook continues to be the most essential and, in the majority of cases, the only aid in the hands of the teacher and the learner through which the given curriculum is transacted. In most classrooms textbooks are the major resources for both teachers and learners. Textbooks give teachers accessible and

² Textbooks publishing dates are as follows:

Middle school: 1st year 2003-2004; 2nd year, 2004-2005; 3rd year, 2005-2006; 4th year, 2006-2007

Secondary school: 1st year, 2004-2005; 2nd year, 2005-2006; 3rd year 2007-2008/2009

plentiful teaching ideas, contents and materials. For students, they serve as the major written exposure. More importantly, textbooks provide students lexical input for vocabulary acquisition.

This situation consequently places a heavy responsibility on the quality of the textbook for ensuring effective teaching-learning interactions and outcomes. It is notorious in the field of teaching and learning languages in educational settings that for any instructional material to be effective, it has to be necessarily planned, keeping in view the requirements of the learners particularly with reference to their age, their educational and cultural background, their immediate environment as well as their proficiency in the language which is the object of instruction. The textbooks are supposed to comply with these criteria. The selection of an appropriate reading passage is critical. If the passage chosen is inappropriate for whatever reason, the chances that learners understand that particular passage are substantially jeopardized. The problem raised here is then whether the textbooks in use for teaching and learning English in the Algerian schools are matched to learners' level and whether teachers are equipped with tools that help them gauge their textbooks difficulty level.

Several authors have discussed factors to consider in assessing the difficulty of texts. Textbooks and text difficulty have then been the topic of a substantial body of research and writing (Britton and Black, 1985; Liesveld, 1988; Muth, 1989; Britton and Guelgoez, 1991, Sawyer, 1991; Chall et. al, 1996; Chambliss and Calfee, 1998; Alexander and Jetton, 2000; Goldman and Rakestraw, 2000; Linderholm, et al., 2000;

Graves and Graves, 2003). Reviewing the literature that dealt with the topic, Graves et al. (op. cit.) identified a set of ten factors that affect a learners' comprehension. These ten factors can be divided into two groups. The first group comprises six factors inherent to the text itself: vocabulary, sentence structure, length, elaboration, coherence and unity, and text structure (or organization, i.e., narratives or exposition). However, the authors are quick to point out that since reading is an interactive process that involves both the reader and the text, no text factors are fully independent of the reader. The second group comprises four factors that involve both the reader and the text: familiarity of content and background knowledge required, audience appropriateness, quality and verve of the writing, interestingness.

The level of reader comprehension of the text is, as research tends to confirm, determined by how well the reader variables interact with the text variables. Comprehension is then seen as a function of the difference between reader ability and text readability. Studies which approached the issue have relied on readability formulas. Traditionally, studies on readability have focused on linguistic and psycholinguistic factors to explain text difficulties. Early readability studies (Chall and Dale, 1948; Flesch, 1943) investigated observable text characteristics (e.g., number of words in a sentence, number of syllables in a word, number of prepositions, and vocabulary frequencies). Studies conducted in the last decades have continued to delve further into factors affecting readability (Gross and Sadowski, 1985; Fry, 1989; Alliende, 1987, 1990; Greenfield, 2003; Litz, 2005; Campbell and Weir, 2006; Crossley, 2006). More

recently, researchers have made use of computer science. These studies have tried to explain text difficulties by measuring texts readability and the ability of readers by attempting to place the two constructs on the same scale. Examples of such studies are: The Lexile Framework (Stenner, 2004), The Strathclyde Complexity Measure (Campbell and Weir, 2006), and Corpus analysis studies. The Lexile Framework for Reading, has been defined by its designers as a scientific approach to reading and text measurement, based on two well-established predictors of how difficult a text is to comprehend: semantic difficulty (word frequency) and syntactic complexity (sentence length).

The Strathclyde Complexity Measure was developed with technology-based testing in mind to account for the shortcomings of other traditional means. It makes use of Latent Semantic Analysis (LSA) which is a theory and method for extracting and representing the contextual-usage meaning of words by statistical computations applied to a large corpus of texts (Landauer, Foltz and Laham, 1998). Crossley (2006) working within a Corpus Analysis framework examined second language reading texts empirically using a computational tool known as Coh-Metrix which is a tool that quantifiably measures different linguistic features of text, as a means of providing answers to how texts can be better assessed for their readability.

Broadly speaking, factors affecting reading comprehension, as listed by Alderson (2000), can be classified into two general categories: reader variables, and text variables. The first category of variables includes factors internal to readers such as

reader's background knowledge, reader's skills and abilities, and reader's motivation and attitude. The second category of variables includes factors such as text content, text type and genre, text organization, text typographical features, and text readability which are internal to texts rather than to readers.

Investing such a broad area of research is beyond any single study. This study aims at investigating reading comprehension difficulties of Algerian EFL students with reference to text variables, more particularly to lexis or vocabulary³. Based on personal experience as an EFL teacher for over 30 years and on the review of the literature, we believe that there is a strong need for investigating and discussing in a specific and thorough way the difficulties and problems encountered by Algerian EFL students in reading English. Consequently, this study purports to identify the source of comprehension difficulties in the process of reading English texts in Algerian official textbooks currently in use in Algerian schools. Printed textbooks and other instructional materials have been and, in spite of the advent of the electronic media, will probably continue to occupy an extremely important place in the teaching-learning process in the schools in Algeria.

Studies of factors contributing to text difficulty have shown that vocabulary knowledge constitutes the main factor related to the difficulty of a text (Alderson and Urquhart, 1984). As McCarthy (1990: viii) noted, “no matter how well the students learn grammar, no matter how successfully the sounds of L2 are mastered, without

³ No distinction is made between the terms “lexis” and “vocabulary”. The two terms are used interchangeably

words to express a wide range of meanings, communication in an L2 just cannot happen in any meaningful way”. Vocabulary is regarded as the variable which can more accurately predict reading comprehension. Furthermore, the ability to recognize vocabulary both in and outside of context (also known as ‘sight vocabulary’) has been stressed as of vital importance behind the reading process. Numerous studies have shown a strong correlation between vocabulary and comprehension (Davis, 1944, 1972; Harrison, 1980; Stahl and Fairbanks, 1986; Alderson, 2000; Coxhead, 2000; Read, 2000; Nation, 2001, Ghahraki and Sharifian, 2005; Moore, 2005, Laufer, 1986, 1987b, 1988, 1989a, 1992a, 1997, 2009; 2010; Laufer and Ravenhorst-Kalovski, 2010; Laufer and Hill, 2000).

Reviewing the literature related to successful reading comprehension for foreign language learners, Laufer (2010) concluded that regarding the relationship between vocabulary and reading most researchers agree that vocabulary is a good predictor of reading, if not the best (Bernhardt and Kamil, 1995; Laufer, 1992b; Nation 2001, 2006; Qian, 1999, 2002; Ulijn and Strother, 1990). Nation and Coady (1988) also stated that vocabulary difficulty has consistently been found to be the most significant predictor of overall text readability. Vocabulary knowledge is so strongly correlated to reading comprehension that some researchers use the two almost synonymously (Carver, 1994). Crossley and McNamara (2008) added that speaking and reading skills are also greatly enhanced by improvements in one’s vocabulary and that lexical growth correlates strongly with academic achievement. Beglar (2009) reminds us that “vocabulary

acquisition is a crucial, and in some senses, the central component in successful foreign language acquisition and, as educators we know that vocabulary is the heart of a language” (p. 7). Learners depend on vocabulary as their first resource (Huckin and Bloch, 1993) and a rich vocabulary makes the skills of listening, speaking, reading, and writing easier to perform (Nation, 1994). Therefore, how many words you need to know in order to do certain things is important in second language acquisition (Miura, 2005; Nation, 2006). Researchers have suggested that among the text-based components, vocabulary is the most important and even the most crucial factor in reading comprehension (Laufer, 1989a; Laufer and Sim, 1985; Nation, 1990). For example, Ulijn and Strother (1990) suggest that reading depends chiefly on understanding the meaning of the text’s words. Cooper (1984) described vocabulary as being the key ingredient to successful reading while other researchers argue that “no text comprehension is possible, either in one’s native language or in a foreign language, without understanding the text’s vocabulary” (Laufer 1997a, p. 20). They maintain that when the percentage of unknown vocabulary in a given text increases, the possibility of comprehending the text decreases (Hirsh and Nation, 1992; Hu and Nation, 2000; Laufer, 1989b, 1990, 1997). Laufer (1989b) was more specific when she revealed the importance of having sufficient vocabulary for reading comprehension, claiming that a reader whose vocabulary is insufficient to cover at least 95% of the words in a passage will not be guaranteed comprehension. Readers themselves consider vocabulary knowledge to be the main obstacle to second language reading comprehension. Yorio

(1971) surveyed second language students, who reported that vocabulary was their main problem in reading comprehension. Still, other researchers such as Sternberg (1987) assert that one can predict a readers' ability to comprehend text based on their vocabulary knowledge. Sternberg states that the level of vocabulary knowledge of a reader may indeed determine their level of comprehension. The general conclusion is that understanding the words in a text makes it easier to understand the text as a whole. The plethora of works referred to above converge on one point that is: text difficulty is a function of the percentage of known or unknown words in a text.

Thus, the primary purpose of the study is to assess the readability of the materials used to teach Algerian learners to read in English through lexical coverage. The research was conducted to determine if the readability level of textbooks is within the range of the students' lexical coverage.

More specifically it seeks to determine whether textbooks used to teach English in Algerian middle and secondary schools are **at**, **above**, or **below** the students' lexical coverage. Worded differently, it seeks to find out if students' vocabulary coverage ensures reading comprehension of the reading passages in their textbooks.

Lexical/text coverage refers to "the percentage of running words in the text known by the readers" (Nation, 2006, p. 61). The ultimate objective is to find out if the readability level of Algerian textbooks is within the range of the student's lexical coverage. Worded differently it seeks to find out if the students' vocabulary coverage ensures reading comprehension of the reading passages of their textbooks. It is my

contention that through careful selection of the vocabulary included in textbooks, we can improve our student's level of English, increase their comprehension and best prepare them for the university.

The assumption made behind lexical coverage is that there is a lexical knowledge threshold which marks the boundary between having and not having sufficient vocabulary knowledge for adequate reading comprehension. This constitutes the amount of unknown vocabulary which can be tolerated in a text before it interferes with comprehension. Some researchers regard one unknown word in every twenty words, roughly in every two lines of a text, as the necessary level beneath which readers are not expected to read an authentic text successfully (Laufer, 1989b; Read, 2000; Schmitt and McCarthy, 1997). Laufer's research tends to confirm that for a text to be suitable for a reader (at the reader's "instructional level"), that reader should know 95% of the words in the text. That is, a lack of familiarity with more than 5% of the running words in a text (one unknown word in less than 20 words) can make of reading a simple passage a daunting activity.

Hence, this project will rely on student's lexical coverage to determine the readability level of textbooks, knowing that readability depends primarily on vocabulary knowledge and that it is a function of lexical coverage. In fact, do textbooks provide useful and appropriate vocabulary items for students in Algeria? What words do the textbooks include and present to the learners? After the seven-year English learning, have Algerian students learned adequate vocabulary to further their studies at

the university level? These are underlying questions which will undoubtedly be of interest to frontline teachers, English syllabus designers, and education policymakers. They will certainly provide to these people pedagogical implications to decide what measures could be done to foster vocabulary learning. These questions could give insights for teachers to understand the effectiveness of textbooks as lexical.

Research questions:

The ultimate objective is to answer the following main research question:

*Are textbooks used by Algerian English as a Foreign Language students **at, above, or below** students' lexical coverage?*

However, in the course of answering this main question there is a set of other questions that this research endeavour asks in order to predict the textbook readability based on lexical coverage. These are:

Research question 1: What is the lexical coverage of Algerian EFL textbooks?

Research question 2: What is the readability level of Algerian EFL textbooks?

Research question 3: Do those textbooks provide sufficient, useful and appropriate vocabulary items?

Importance of the study:

From the point of view of the quality of a textbook, its level of readability is a critical factor which could help or hamper the understanding or learning by the learner

depending on the extent of match/mismatch between the linguistic competences demanded by the text and that available with the learner. To facilitate this understanding, the readability of the texts used must be ensured in the textbook. The idea is to have students read materials that are not so difficult that they cause frustration but challenging enough to build vocabulary and comprehension skills (Appelt, 2006). If the readability of the texts is higher than the reader's lexical coverage, it may be frustrating. If, on the other hand, readability is far below the student's lexical coverage, the text may be felt too easy and become boring and uninteresting.

Recently, a certain number of English textbooks have been issued by the Algerian educational authorities (see dates above). Studies specific to the Algerian context would have facilitated informed decisions regarding appropriateness and sequencing of content to match the linguistic competence of the learners. The dearth of such studies specifically in the area of readability in Algeria, however, reinforces the need and priority for such studies, since these could contribute significantly to the improvement of the quality of instructional material being produced. It should be mentioned that to the best of my knowledge, no study has been conducted in the area of readability of English textbooks in Algeria relying on text coverage as yet. Therefore, there is an imperative need for in-depth empirical studies of the existing textbooks in the above context. Such studies would be of some help to many people as the outcome of the study will be corpora to be used as basis for comparison using computer software. The

data collected, the software, and the results of the study could be further exploited by a certain number of people among whom:

Teachers: Teachers could be provided with a tool that would help them find the right measure to decide whether a particular reading passage is suitable for a particular group of readers. They would be able to say whether a reading comprehension text, a test, or a lecture is likely to be understood and hence decide on the level of difficulty and take informed decisions to choose with confidence materials that will improve students' reading skills and take the guesswork out of matching readers with appropriate texts. Moreover, such studies can be very useful in teacher development and professional growth by helping "teachers move beyond impressionistic assessments and ... acquire useful, accurate, systematic, and contextual insights into textbook material" (Litz, 2005, p. 8). As far as teachers are concerned, one additional reason for the need of such a study is the fact that it would provide teachers of English with a sense of familiarity with their textbook content, thus assisting in identifying the particular strengths and weaknesses of the reading material in use. This would ultimately assist them with making optimum use of the textbook strong points, and recognizing the shortcomings of reading material, if any. Finally, results will serve as a gauge to Algerian students reading ability.

Course book designers and writers: Course book designers and writers will find the outcomes of this project useful for curriculum development as it will determine the vocabulary content of each textbook. Armed with documented evidence, syllabus

designers can take informed decisions as to the selection of instructional material for future instructional material. Ultimately, this can be used at least for the vocabulary gradation of textbooks. They will be able to say to what extent the vocabulary of a textbook of a certain level matches the vocabulary of another textbook (how many words in a given syllabus have been covered by the preceding one?). They will also be able to know the extent to which the vocabulary of a unit or file within a textbook matches that of another unit or file (how many words in a given unit or file have been covered by the preceding unit or file?). The research outcome of this study will help them find out which high frequency words (words which occur most frequently in English) learners don't know and determine what students need to learn then include them in the syllabi. Finally, educational authorities could make use of the results and corpora of the present study to gauge the lexical level of students at any stage of their curriculum, set goals for vocabulary teaching and learning, and use them for textbook evaluation within the limits of the present study.

Authors and publishers of graded readers: Authors of graded readers destined to Algerian students of English could exploit the results and data of the present research to match their material with levels of their intended audience.

Researchers: Further research can be carried out relying on the results and corpora of the present study (see the section "suggestions for further research" in this study).

To sum up, studies related to vocabulary and to the readability level of the books and the reading levels of students are badly needed in Algeria. Improvement in teaching reading will be fruitless unless these are taken into consideration. It is hoped that this will open up the field of readability⁴ and corpus analysis in Algeria, to demonstrate the need to make use of readability studies and corpus analysis in order to match students with textbooks, and above all, to provide a body of data for future research and syllabus design.

Theoretical framework:

This study draws on literature related to the investigation of text readability, text difficulty and corpus linguistics. Estimating the readability of text by examining its linguistic characteristic is a long-standing empirical tradition. Such a tradition holds that learning vocabulary is the fundamental step to learn a foreign language and, as already mentioned, that vocabulary knowledge is the single best predictor of reading comprehension.

The concept of readability has been given great attention in the past fifty years. The notion that textbook reading difficulty needs to be matched with student reading ability has been emphasized by readability investigators (Shokrpour, 2005). Text difficulty has been a concern of educational researchers and practitioners during this era (Chall and Conrad, 1991) and many have used different methods to assess the difficulty

⁴ However two studies are worth mentioning here: Lakehal-Ayat – Benmati, K. (2008) and Belouahem, R. (2008). These two studies tackled readability but from a perspective quite different from the present one.

of the text (Chall and Dale, 1995). In fact, one of the most important aspects of textbook development has been considered to be texts of appropriate difficulty by educational publishers. Matching the difficulty of textbooks and readers' reading ability has been taken into account by publishers, writers, editors and teachers in order to use the text successfully (Chall and Conrad, op. cit.; Harrisi-Sharples, 1983; Day, 1994; DuBay, 2004; Crossley et al., 2007).

Readability formulas measure certain features of the text that can be put into mathematical calculations. The way they work is that the reader can be questioned or tested on something they read and then the material they read can be tested with the formula(s). The reader's success in understanding the material as measured on an exam can be correlated to the readability score of the text itself. As already noted, controlling the readability of reading instructional material has led to the emergence of different readability formulas. Many of such formulas have been widely used for several decades within first language since Dale and Chall's (1948) seminal article and have been considered as the most reliable and valid (Klare, 1963; 1984). The two factors considered as predictors of difficulty and common to all formulae are vocabulary difficulty and syntactic difficulty. In fact, readability research of English shows that: "for most purposes, a simple two-factor formula that is based on word familiarity or length, and sentence length or complexity, is sufficient because these two variables are most predictive of comprehension difficulty" (Willows et al., 1981 p. 181). These authors also assert that: "In measures of readability of a text, vocabulary difficulty has

consequently been found to be the most significant predictor of overall readability” (idem. p. 97).

Vocabulary difficulty has been measured according to the familiarity of the vocabulary to the reader. If a word is familiar to the reader, its vocabulary difficulty is low, and vice versa. Vocabulary familiarity is commonly measured by word frequency. Words frequency means the frequency with which a given word occurs in a sample of the target language. Word frequency and vocabulary difficulty are, as Lin (2002) noted, “... inverse in variation. The higher the word frequency, the smaller will be the vocabulary difficulty, and vice versa” (p. 153).

Measures of syntactic difficulty rely on sentence length which is often considered as an indicator of syntactic complexity: the longer a sentence is, the more syntactically complex it is and therefore more likely to be difficult to understand. According to Fry (1989, p. 95): “... the two inputs of sentence length and word difficulty accurately predict how easily a given passage will be understood by the average reader”.

As per the validity of formulas in second language contexts, several studies have tested them. To name but a few, Hamsik (1984) concluded that “readability formulas and graphs do measure readability of ESL students and that they can be used to select for them material appropriate to their reading level”. Nilagupta (1977) came to the same conclusion. More recently, Lin (2002), Greenfield (2003), Crossley (2006) and Weir and Ritchie (2006) used already existing formula or developed their own formula to assess text difficulty and readability in EFL contexts.

Researchers within corpus linguistics have endeavoured to find out which words and word families are necessary to know in order to reach 95% of lexical knowledge that enables reading comprehension. The most frequent 2000 words of English account for about 80% of most texts. The conclusion was that learning the 1,000 most frequent words and word families in the English language allows a reader to comprehend around 74% of words in an academic text (and the percentage is higher for newspapers, fiction, and conversational English). Learning the second most frequent 1,000 words and word families (known as the “2k list”), covers an average of 5% more (in academic texts). So knowing the first 2,000 most frequent word families in the language brings us up to 80% lexical familiarity with academic texts (Coxhead, 1998; Nation 1990).

This principle has been adopted by some dictionaries such as *Longman Dictionary of Contemporary English* in which the defining vocabulary amounts to the same figure. This is noted in *Longman Dictionary of Language Teaching and Applied Linguistics* (2002) as follows:

Defining vocabularies are used to write definitions in dictionaries for children and for people studying foreign languages. They are based on research into word frequency. In the Longman Dictionary of Contemporary English, all definitions are written using a 2000 word defining vocabulary, so that anyone who knows the meaning of those 2000 words will be able to understand all the definitions in the dictionary. (p.147)

ESL/EFL research posits that a text with no more than approximately one unknown word in every twenty words (5% unknown) would be considered easy. This is known as the independent reading level and the comprehension rate would be 95%.

A text with one unknown word in every ten words (10% unknown words) would be at the instructional reading level, that is challenging but manageable text for the reader and the comprehension rate would be 90%.

Finally, if the reader is unfamiliar with more than one unknown word in every ten words the text is considered difficult and the comprehension rate would be less than 90%. The text is said to be at the frustration reading level (Laufer, 1997).

Hypotheses:

The hypotheses to be proved or disproved would be as follows:

H1: The EFL textbooks used by Algerian students are **above** the students' lexical coverage. If so, then the textbook readability is **low** (reading material is **difficult**).

H2: The EFL textbooks used by Algerian students are **at** the students' lexical coverage. If so, then the textbook readability is **medium** (reading material is of **medium difficulty**).

H3: The EFL textbooks used by Algerian students are **below** the students' lexical coverage. If so, then the textbook readability is **high** (reading material is **easy**).

We can determine whether Algerian student are learning sufficient, useful and appropriate vocabulary items by using a computer based tool to compare the lexis in the textbook to the existing lists of most frequent words. Cobb (2000), Laufer (2010) and many others explained that if a student knows the first 2000 most frequently used English words plus the 570 words of the Academic Word List (AWL, Appendix 1), then he/she knows about 90% of the words he/she will meet in any academic text. Furthermore, ninety percent of the dictionary's concepts can be expressed in 850 words (Cobb, 2000).

Methodology:

The study purports to analyse the seven English as a Foreign Language textbooks in use in the Algerian schools. In order to answer the research questions, the present study will first look at the quantity of vocabulary distribution in all the textbooks. For the measurement of vocabulary levels of textbooks a full list of the words occurring in each textbook is needed to form a basis of comparison. For such a task a computer method was adopted involving the use of software tools. Hence, the contents of the seven textbooks were converted into machine-readable format text files, so that they could be “read” by the computer programme used to analyze them. The methodology adopted was divided into two main steps: the first step consisted in corpus compilation and the second step consisted in the study of the corpus with software. More details are provided in Chapter Seven.

Limitations of the study:

There are two limitations that need to be acknowledged and addressed regarding the present study.

1. Generalizations generated from the study will be restricted to EFL textbooks in use in Algerian schools.
2. The readability findings will be limited to the texts contained in the textbooks used in this study.

Structure of the research:

The study consists of an introduction and nine chapters. The introduction provides an overview of the study and presents the problem which the study tries to clarify, the rationale, the methodology to be followed, the aim of the study and its organisation.

Chapter Two reviews the reading theories and models. It examines the various definitions provided for the term ‘reading’, sketches a historical overview of reading and the reading models since the early decades of the last century. Special emphasis is laid upon interactive reading models and the schema theory as they pertain directly to the present project.

Chapter Three is devoted to the literature on readability and the assessment of text difficulty. The bulk of this chapter is devoted to the landmarks studies on readability and the controversy regarding the formulas. The history of various ways whereby text difficulty and readability have been assessed is surveyed. The literature on the variables affecting readability and the methods of readability measurement are reviewed along

with criticism directed at the use of readability formula. It also discusses the validity of readability formula in EFL contexts.

Chapter Four considers reading comprehension in a second/foreign language. Its primary objective is to help the reader gain insights into the nature of reading comprehension. It aims at shedding light on the factors which influence and constrain reading comprehension bearing in mind that reading is an interactive process. It considers the relationship between reading comprehension in a first language and a second/foreign language, highlighting differences and similarities as well as the influence of first language on second/foreign language. The factors that affect L2/FL reading comprehension are then reviewed.

Chapter five discusses what is meant by *word*, *word knowledge*, and *vocabulary*. It aims at clarifying what is considered as a *word* as this is of paramount importance to this project. It presents the debates around the definition of this term as well as what constitutes word knowledge, what is meant by the term vocabulary and the different types of vocabulary.

Chapter six provides an overview of research concerning vocabulary learning and comprehension on the one hand, and vocabulary learning and lexical text coverage, or simply lexical coverage, on the other. It examines the issues of how some aspects of vocabulary might affect the stability of text coverage. The relationship between vocabulary and comprehension is discussed in terms of both theoretical and empirical basis for the connection between understanding the meaning of individual words and

understanding text. After defining a few useful terms, it moves on to the central issue in this work, lexical coverage. A brief historical overview of the evolution towards recognition of the importance of lexical competence within second/foreign language learning is briefly sketched. The bulk of this chapter is devoted to the review of the literature on lexical coverage.

Chapter seven exposes the method and procedures developed in this study. Chapter eight analyses the data collected by means of the various instruments used. Chapter nine is a concluding chapter. It opens with an overall picture of the dissertation then moves on to expose the findings. Pedagogical implications of these findings are considered next. The limitations of the study are discussed before ending by formulating recommendations for future research.

CHAPTER 2

READING THEORIES AND MODELS

Introduction

As already mentioned above, reading is a complex process in first language, let alone in second/foreign language. It includes intricate processes and skills. Understanding reading in a second/foreign language requires the examination of the psychological processes involved in first language reading. Consequently, this chapter opens with an examination of the various definitions provided for the term 'reading'. Following this is a historical perspective on reading and the reading models since the early decades of the last century. Some of these models continue to influence research to this day. This is done in order to situate reading within the broader fields of linguistics and psychology. Special emphasis is laid upon interactive reading models and the schema theory as they pertain directly to the present project.

2.1. What is reading?

Defining reading is not an easy task. Rather, it is an enormous task and there is a plethora of definitions provided in the literature bearing witness of the stark disagreement of scholars on the definition of this term. Most of people think of reading as a simple, passive process that involves reading words in a linear way and getting their meaning at the same time.

The American Heritage Dictionary defines reading as: "The act or practice of rendering aloud written or printed matter". Webster's Seventh New Collegiate

Dictionary (1971) gives as many as seven principal definitions of the word read, such as "to receive or take in the sense of (as letters or symbols) by scanning;" "to understand the meaning of (written or printed matter);" and "to attribute a meaning or interpretation to (something read)" (the brackets in the quotations are original). However, such definitions do not seem to offer a satisfactory answer to the question "what is reading?" Terms such as "take in", "sense", "understand", "meaning" and "interpretation" themselves are controversial and need to be defined to obtain a clearer definition and understand what is really meant by reading.

Over the past few decades, many researchers have investigated the process of reading and have attempted to gain insights into the process of reading. Urquhart & Weir (1988, p.22) define reading as follows: "Reading is the process of receiving and interpreting information encoded in language form via the medium of print". Anderson (1999, p. 1) states: "Reading is an active, fluent process which involves the reader and the reading materials in building meaning. Meaning does not reside on the printed page, nor is it only in the reader". Still, Grabe and Stoller (2002), conclude that it is possible to present a single sentence definition of reading. They state: 'Reading is the ability to draw meaning from the printed page and interpret this information appropriately' (p. 6). Another prominent researcher into reading Goodman (1967) defines reading in the following terms:

Reading is a psycholinguistic guessing game. It involves an interaction between thought and language. Efficient reading does not result from precise perception and identification of all elements, but

from skill in selecting the fewest, most productive cues necessary to produce guesses which are right the first time. The ability to anticipate that which has not been seen, of course, is vital in reading.
(p. 260)

These definitions share in common the view that in reading decoding print and reading comprehension are inseparable. That is, reading comprehension is the obverse of reading. However, these definitions are insufficient to help us understand completely the true nature of reading because they do not convey that a reader has several possible purposes for reading and each purpose emphasizes a somewhat different combination of skills and strategies. They also do not emphasize the many criteria that define the nature of reading abilities, as they do not explain how reading is carried out as a cognitive process (Lin, 2002).

Hence, the literature review reveals that a clearly stated, empirically supported, and theoretically unassailable definition of reading is still not available. Bernhardt (1991) summed up the situation: "if the question 'what is reading?' could be answered in the pages of this book, most educational researchers could retire or select alternative careers" (p. 5). It is so because reading includes intricate processes and skills which differ according to tasks, purposes and readers' abilities. Without knowing these elements independently, the concept of reading remains unclear.

A common view in the research literature on reading is that reading is basically divided into two components: decoding (word recognition i.e., identifying the words in print) and comprehension (constructing an understanding from them). There ensues a

coordination of identifying words and making meaning so that reading is automatic and accurate which is known as fluency.

It appears then, as Alderson (2000, p. 3) points out here that it is widely accepted to draw a distinction between the process of reading and the result of that process, the product. The process is what we mean by “reading” per se: the interaction between a reader and the text. The process, he further argues, is likely to be dynamic, variable and different for the same reader on the same text at a different time or with a different purpose in reading. It is even more likely then, that the process will be different for different readers on different texts at different times and with different purposes.

Alderson describes this process as follows:

During that process it would appear that many things are happening. The reader is looking at print, deciphering the symbols on the page, deciding what they “mean” and how they relate to each other. The reader is presumably also “thinking” about what he or she is reading: what it means to him or her, how it relates to other things he or she has read, to things he knows, to what he expects to come next in texts like this. He or she is presumably thinking about how, entertaining, boring or crazy the text is. He may also be consciously reflecting on the difficulties or ease he is experiencing when reading, and on ways of overcoming the difficulties or of continuing the enjoyment. He may be completely unconscious of how he is reading and he may be fully absorbed in “reading”. (p. 3)

In order to get a more profound understanding of reading in a second/foreign language it is necessary to look at the psychological processes involved in reading from L1 and insights gained from research on reading in a second/foreign language . Such a task calls for a brief overview of the history of reading and the reading models that have been posited in the professional literature since the early decades of the last century. Thus, the following is an attempt to review briefly the literature on research in reading in L1 and L2.

2.2. Historical perspective on reading

L1 reading research, according to Samuels and Kamil (1984, p. 22) is “just a little more than a hundred years old”. Its beginning was marked Emile Javal’s publication of his paper on eye movement in 1879 and serious attempts at building explicit models describing the entire reading process have a history of about half a century. Javal observed that a reader's eyes do not sweep smoothly across print but make a series of short pauses, or saccades, at different places until reaching the end of a line, when they move to the beginning of the next in a smooth, unbroken fashion (Huey, 1968 cited in Paulson and Goodman, 1999). Before Javal, it was assumed that the eye glided unceasingly across text. Perhaps the first concrete insight into the reading process made possible by eye-movement research was provided in 1891 by Landolt who observed subjects' eye movements while they read different types and genres of text, and discovered that "reading of a foreign language required more pauses, as did also the reading of detached words, numbers, and lists of proper names (Huey, 1968, p. 19 cited in Paulson and Goodman 1999). Landolt thus provided the first evidence that the eyes

do not proceed on a regular, predetermined path, but vary depending on the type of reading being done. Study of their movements therefore provides a window to the cognitive processes of perception and comprehension that take place during reading.

Subsequent research focused on the perceptual span. The perceptual span is the extent into the periphery within which the visual information are acquired and used during fixations in reading. It was found that readers have very limited perceptual spans. For example, on average, the perceptual span of English readers is about four character spaces to the left of fixation and 14 to the right (McConkie & Rayner, 1976). Underwood and McConkie (1985) further demonstrated that letters are distinguished only within eight character positions to the right. Beyond that point letters become indistinguishable and only word boundary and word shape information are obtained. Normal skilled readers have been found to make no more than 4-5 saccades in a second and thus 4-5 fixations per second. It was also found that not every word in a text is fixated and that not only do different readers read differently, but individual readers read differently in different circumstances (Judd & Buswell, 1922, cited in Fisher and Shebilske, 1985)

During the early decades of the previous century, approaches to reading were dominated by the Grammar Translation Method, a method which required enormous amounts of memorization and translation. Following traditional instructional methods for classical languages such as Latin and Ancient Greek, L2/FL learners were regularly required to learn grammar rules, do translation exercises, read texts in the original, and respond to text questions orally and in writing.

When the Grammar-Translation Method was replaced with the audio-lingual approach in the fifties, the importance of reading decreased as the speaking and listening skills were given more prominence. During the sixties, reading was viewed as a mere support skill for grammar and vocabulary acquisition. Ever since, reading has remained of diminished third place in importance, closely followed by writing.

In the first sixty to seventy years of the twentieth century research in reading was under the overwhelming influence of behaviourism and positivism. Influenced by these two movements, the study of reading was focused mainly on directly observable events external to individual readers (Ellis, 1994; Samuels & Kamil (1984). During this period, the emphasis in reading research was to find out, describe and provide models for how visual stimuli such as printed words became associated with word-recognition responses. The dominance of behaviourism with its rejection of mentalism in psychology, as well as its emphasis on the observability of data, refrained most attempts at constructing theories or models about events that were not directly observable, such as reading processes. Under the influence of behaviourism, reading was viewed as a process of decoding linguistic orthographic symbols into mental semantic codes (Ellis, *op. cit.*; Pearson & Stephens, 1994). In this type of approach, major attention was focused on the forms and structures of language in which the text was coded. Accordingly, printed texts were structurally analyzed into their componential parts: letters, at the first level; words, at the second level; sentence, at the third level; and text, at the final fourth level (Brown, 1998). The act of reading was then thought to be a linear process of decoding, beginning from the first level of letters and progressing

serially up through the levels of words and sentences and concluding at the final level of text. Reading was considered as "translating graphic symbols (letters) on a printed page into an oral code (sounds corresponding to those letters)" (Pearson & Stephens, 1994, p. 23). It followed that "... reading comprehension or understanding of the text was likewise conceptualized as being able to remember the text" (Brown, 1997, p. 36).

Since the mid-1960s, with the developments in cognitive psychology, much of the research in reading has moved away from visual signal perception and started to investigate what happens in the recesses of the mind that allow humans to make sense of the visual information captured during fixations (Klein, 1988). The sixties were soon followed by an era during which researchers advocated a much stronger emphasis on reading as part of a meaning-making process (Goodman, 1967; Smith 1973, 1978). This period witnessed a burst in model-building activity.

After the demise of the audio-lingual approach in the late 1970s, reading re-emerged. The increasing emphasis on reading led to a psycholinguistic model or theory of reading, the theoretical perspectives of which were taken and expanded upon by first language (L1) and second/foreign (L2) researchers alike during the 1980s (Bernhardt, 1984, 1985, 1986; Coady, 1979; Goodman, 1988, 1992, 1996; Goodman and Goodman, 1980; Smith, 1978, 1985, 1988; Swaffar et al., 1991).

It was only during the 1980s that much attention was focused on reading in a second or foreign language. This era witnessed a boom in literature on reading in a second or foreign language (e.g. Alderson and Urquhart, 1984; Bernhardt, 1991; Grellet, 1981; Ulijn, 1977; Ulijn and Kempen, 1976; Swaffar et al., 1991). The

impetus of such a boom was to understand better the process of reading. Since the mid-eighties, there has been a significant shift towards describing the reading process either in terms of skills and knowledge areas within a cognitive process. We find in the literature many theoretical models that have been developed to explain the process. However, the following considers briefly models which have been most influential in the development of methods for the teaching of reading and the conception of textbooks and which are of particular relevance to this project. These are the bottom-up view of the reading process (Gough's, 1972), the top-down model of the process (Goodman 1971, 1970, 1975, 1988), the interactive model (Rumelhart, 1977), an alternative model to the aforementioned ones. These models have been greatly influenced by what is known as schema theory.

2.3. Schema theory

Schema theory is a framework which encompasses a range of theoretical models. It is considered here because it has had profound implications for the process of reading and a great impact on understanding reading. Before proceeding any further, the concept of schema (plural schemata) must be defined.

2.3.1. The definition of schema

As far back as 1932, Bartlett saw memory as constructive and mental representation was built from current discourse and background knowledge. Schema was an active feature organizing the pieces to develop memory. Bartlett (1932) describes how knowledge is constructed and represented in memory. This theory explains how and where information is stored and the process by which readers

combine their own background knowledge with the information in a text to comprehend that text, in short, how we make connections between ideas.

Widdowson, (1983) describes schema as cognitive constructs which allow for the organization of information in long-term memory. Cook (1989, p. 69) states, "the mind, stimulated by key words or phrases in the text or by the context, activates a knowledge schema". Widdowson & Cook both emphasize the cognitive characteristics of schema which allow us to relate incoming information to already known information. This covers the knowledge of the world, from everyday knowledge to very specialized knowledge, knowledge of language structures, and knowledge of texts and forms they take in terms of genre, and organization.

According to Carrell (1984a), the role of background knowledge in language comprehension has been formalized as schema theory. Carrell states that schema theory holds that any text, either spoken or written, does not carry meaning by itself; rather, a text only provides directions for listeners or readers as to how they should retrieve or construct meaning from their own, previously acquired knowledge. Such knowledge is called the reader's background knowledge; the previously acquired knowledge structures are called schemata. The process of interpretation is guided by the principle that every input is mapped against some existing schema and that all aspects of that schema must be compatible with the input information. This principle results in two basic modes of information processing, called bottom-up and top-down processing (these models are discussed below).

Schema theory is then based on the belief that "every act of comprehension involves one's knowledge of the world as well" (Anderson & Pearson in Carrell et al. 1984, p. 73). Thus, readers develop a coherent interpretation of text through the interactive process of "combining textual information with the information a reader brings to a text" (Grabe 1988, p. 56). Readers' mental stores are termed 'schemata' (after Bartlett). Schemata are then mental structures stored in memory and called upon during the process of making sense of a text. These constructions, which make comprehension while reading, are formed by three types of schemata.

2.3.2. Types of schema

Generally, there are three major types of schemata, namely, linguistic schemata, formal schemata and content schemata, which are closely related to reading comprehension. These issues are discussed in more details in the section on reading comprehension.

2.3.2.1. Linguistic schemata

Linguistic schemata refer to readers' existing language proficiency in vocabulary, grammar and idioms. They are the foundation of other schemata. As widely acknowledged, linguistic knowledge plays an essential part in text comprehension. Without linguistic schemata, it is impossible for the reader to decode and comprehend a text. Therefore, the more linguistic schemata a reader has in his mind, the faster the reader acquires information and the better understanding the reader may get.

2.3.2.2. Formal schemata

Formal schemata are the organizational forms and rhetorical structures of written texts. They include knowledge of different text types and genres, and also include the knowledge that different types of texts use text organization, language structures, vocabulary, grammar and level of formality differently. Formal schemata are described as abstract, encoded, internalized, coherent patterns of meta-linguistic, discourse and textual organization that guide expectation in our attempts to understand a meaning piece of language. Readers use their schematic representations of the text such as fictions, poems, essays, newspaper articles, academic articles in magazines and journals to help comprehend the information in the text. Studies show that the knowledge of what type and genre the text is can facilitate reading comprehension for readers because the type of the text will offer detailed evidence of the content of the text. Nonetheless, compared with the linguistic and content schemata, the formal schemata offer less power in the reading process (Carrell, 1984b). Schooling and culture play the largest role in providing one with a knowledge base of formal schemata For example; -talking about formal schemata readers know that a text describing process is organized in a different way than one of an argumentative nature. So, when they read a descriptive text they activate a schema comprised of specific signalling devices such as the words ‘firstly, secondly, thirdly’, use of particular tenses such as simple present, and use of passive voice.

2.3.2.3. Content schemata

Content schemata are related to factual information about topics of various types. They refer to the background knowledge of the content area of a text, or the topic a text is about. They include topic familiarity, cultural knowledge and previous experience with a field. Content schemata deal with the knowledge relative to the content domain of the text, which is the key to the understanding of texts. Since one language is not only the simple combination of vocabulary, sentence structure and grammar but also the bearer of different levels of the language's culture. To some extent, content schemata can make up for the lack of language schemata, and thus help learners understand texts by predicting, choosing information and removing ambiguities. Many studies show that readers' content schemata influence their reading comprehension more greatly than formal schemata. On the whole, the familiarity of the topic has a direct influence on readers' comprehension. The more the reader knows about the topic, the more easily and quickly he gets the information of the text. Therefore, if one wants to be an efficient reader, he needs to try to know the knowledge about more fields and topics. Learners with more prior knowledge can better comprehend and remember more the text.

2.4. Reading models

2.4.1. What is a reading model?

In the last few decades reading researchers have been studying the link between the reading process (what goes on in the brain) and how to teach reading. Depending on their interpretation of the reading process, they have developed a model of reading.

The term model has been defined as a systematic set of guesses or predictions about a hidden process, which are then subject to testing through experimental studies. A reading model then provides a description of the reading process. It is a visual representation of a theory which describes what goes in the eyes and the mind when readers are attempting to understand a text. Singer and Ruddell (1985, p. 88) define it as "... a graphic attempt to depict how an individual perceives a word, processes a clause, and comprehends a text."

2.4.2. The bottom-up reading model

2.4.2.1. What is a bottom-up reading model?

The bottom-up model of reading is primarily concerned with the recognition of individual letters, phonemes and words. This model believes that the reading process begins with individual recognition of letter and phonemes. According to Gough et al. (1996), this model functions as follows: first, the graphemic information enters through the visual system and is transformed at the first level from a letter character to a sound, that is from a graphemic representation to a phonemic representation; second, the phonemic representation is converted, at level two, into a word. Meaning of the whole text is a process of building understanding of individual letters to the word level, then to the sentential level, and finally the text level. This process is also referred to as "data-driven". In this view learners are perceived as being almost passive decoders of visual stimuli.

Gough et al. posit that reading is a strictly serial process: letter-by-letter visual analysis, leading to positive recognition of every word through phonemic encoding and

lexical, syntactic and semantic rules are applied to the phonemic output which itself has been decoded from print. In this view, bottom-up models is seen to operate on the principle that the written text is hierarchically organized (i.e., on the grapho-phonetic, phonemic, syllabic, morphemic, word, and sentence levels) and that the reader first processes the smallest linguistic unit, gradually compiling the smaller units to decipher and comprehend the higher units (e.g., sentence syntax). The bottom-up approach to reading comprehension is well represented in LaBerge and Samuels' (1974) Automatic Information Processing Model. In the model, reading is viewed as a process of decoding, from bottom to top, from part to whole, from surface to deep and from external to internal. The bottom-up approach to reading comprehension dominated the study of reading comprehension up till the 1960s.

2.4.2.2. Critique of the bottom-up reading model

Since the inception of the bottom-up approach in the 1960s, a number of problems have been identified because some occurrences that are found to take place during reading were not accounted for in this model (Carrell, Devine, & Eskey, 1988). Researchers have come to the conclusion that reading comprehension is not as text-driven as it was viewed in the bottom-up model. Instead of a uniform interpretation, readers have been found to read different meanings into the same text. Researchers in the field of reading research have also noticed that, although visual information derived from the orthographical symbols is important for reading comprehension, the non-visual information already acquired by the reader played an even more important role in reading comprehension (Goodman, 1967, 1970, 1971; Smith, 1985).

According to Eskey (1973), the bottom-up reading model is inadequate because it underestimates the contribution of the reader who makes predictions and processes information. It fails to recognize that students utilize their expectations about the text, based on their knowledge of language and how it works. Stanovich, (1980) argued that an important shortcoming of this model is the fact that it is difficult to account for the role of prior knowledge of text topic as a facilitating variable in word recognition and comprehension. Other researchers (e.g., Coady, 1997; Eskey, 1988; Goodman, 1988; Haynes, 1993) have argued that bottom-up models do not account for observed reading phenomena. For example, researchers conducting miscue analysis studies (analysis of the mistakes readers make in oral reading) concluded that readers do not passively take in the information from the text, but rather are actively involved in predicting meaning based on both cues from the text (inferencing) and their background knowledge.

Rumelhart (1977) sees an essential flaw in the bottom-up model because processing is seen as proceeding only in one direction which implies that no higher level information ever modifies or changes lower level analysis. The following illustration cited by Rumelhart refutes the idea that readers rely only on "bottom-up" processing. In this case readers are able to identify the word "read" correctly only by employing higher level semantic and syntactic processing.

I read that story yesterday.

Today, I am going to read a new story.

From this example, it is evident that to pronounce the word *read* correctly in each case, the reader draws upon world knowledge and meaning as well as grammatical

sense to facilitate word identification. Hence it is obvious that cognitive processing at higher levels influences lower or surface level processing. Reading thus seems to be an interactive, rather than a linear process based only on letter perception.

Other support for this position (Rumelhart, 1977) that reading is not entirely "bottom-up" has been demonstrated in a number of psychological studies in which: (1) subjects are cognizant of more letters when a word is presented than when a string of unrelated words is presented (Huey, 1968); (2) letter strings are perceived as the original word even when a letter is deleted or replaced by one or two letters (Pillsbury, 1897, as cited in Rumelhart, 1977); and (3) a letter is perceived more accurately when it is part of a word than when it is among a set of unrelated letters (Reicher, 1967). Not only that, but McClelland and Johnston (as cited in Rumelhart, 1977), demonstrate that letters are more accurately perceived when they are embedded in orthographically regular rather than orthographically irregular strings, suggesting that the perception of letters depends upon the surrounding letters. Rather than being perceived singly then, letters are often perceived in clusters.

Other relevant research reported by Rumelhart (1977) indicating how semantic processing directs lower level processing includes a series of experiments in which the time required to identify such word pairs as bread-butter, doctor-nurse as opposed to such pairs as bread - doctor, nurse - butter is shorter when the words are semantically related. This adds further support to the thesis that word perception is influenced by meaning. As Stanovich (1980) contends, serial stage models of reading run into

difficulty because no mechanism is afforded whereby high-level processes can influence lower levels.

Such dissatisfaction with the bottom-up model and a growing interest in the study of the constructive role of the reader in reading have led to the development of a different type of approach to reading comprehension, which has come to be known as the top-down approach. The rise of this approach in reading research indicates a significant shift of the focus of the study in reading research from data-driven to concept-driven or reader-driven (Carrell, 1991; Eskey, 1988).

2.4.3. The top-down reading model

2.4.3.1. What is a top-down reading model?

One of the earliest models of reading that questioned the bottom-up models was proposed by Goodman (1965, 1966, 1967/1976). It was called Psycholinguistic Model of Reading. In this model, readers are said to rely on their existing syntactic and semantic knowledge structures to minimize their reliance on the graphic display and the grapho-phonemic knowledge. This model of reading is still largely influential. Goodman refers to reading as a ‘psycholinguistic guessing game’ (Goodman, *op. cit.*). It is based on a consideration of schema theory which says that comprehension depends on the activation of schemata. This issue is discussed below.

Smith (1985), unlike Goodman, did not propose a model of reading. But his description of the linguistic and cognitive processes in reading has called attention to a similar issue that Goodman tries to deal with in his model. Smith noticed that redundancy was inherent at all levels of language and that readers enjoyed remarkable

flexibility in creating a meaning for the text at hand. His findings have lent top-down models considerable support. According to Smith (1985), reading comprehension is dependent on two kinds of information: (1) visual information, which is perceived from the print and (2) non-visual information, which comprises readers' understanding of the relevant language, their familiarity with the subject matter and their general ability in reading. The non-visual information is in readers' heads and is "behind the eyes" as Goodman (1970) puts it. The relationship between visual information and non-visual information in terms of their contribution to reading comprehension is mutually inverse: the more non-visual information has been acquired, the less visual information is needed; and the less non-visual information is available, the more visual information is required (Smith, 1985). To put in another way, it seems as if the comprehension of a text requires a definite amount of information -- the more one gets through non-visual information, the less one depends on visual information.

According to Smith and Goodman, reading can be considered as a matter of asking questions about what is being read and getting the answers to those questions through reading (Goodman, 1970). It is in this sense that Smith (1985) claims that the aspect of reading that all have in common is that questions are asked of the text. Comprehension occurs when answers to these questions are found. Letter identification, word identification and the comprehension of meaning are independent consequences of asking different kinds of questions of text. Comprehension need not require word identification, which in turn need not require letter identification. (p.124)

The top-down reading model, also known as inside-out model, concept-driven model, whole-to-part model, is almost the reverse of the bottom-up model. In this model the flow of information is considered to proceed from the top downward so that the process of word identification is dependent upon meaning first. The processing sequence proceeds from prediction about meaning to attention to progressively smaller units. According to theorists in this camp, what readers bring to the text separately in terms of both their prior knowledge of the topic and their knowledge about language, assists them in predicting what the upcoming words will be. Bottom-up processing models (e.g., LaBerge and Samuels, 1974) take a cognitive perspective. They are data-driven, emphasize lower-level processes such as letter and word recognition, and most importantly, emphasize textual decoding due to the primary priority placed upon the text as input. Therefore, the top-down reading model suggests that processing of a text begins in the mind of the readers with meaning-driven processes, or an assumption about the meaning of a text. From this perspective, readers identify letters and words only to confirm their assumptions about the meaning of the text. The proponents generally agree that comprehension is the basis for decoding skills, not a singular result, and meaning is brought to print, not derived from print. It makes the reader an active participant in the reading process as opposed to the bottom-up process where reading is seen as a passive linguistic decoding process.

However, later Goodman (1989) admitted that this model is interactive,.

Goodman wrote referring to this model:

... it is one which uses print as input and has meaning as output. But the reader provides input too, and the reader, interacting with text, is selective in using just as little of the cues from text as necessary to construct meaning. (p. 76)

He also stated: "...the goal of reading is constructing meaning in response to text. It requires interactive use of grapho-phonetic, syntactic, and semantic cues to construct meaning". (p. 77)

2.4.3.2. Critique of the top-down reading model

The main critique that has been addressed towards the top-down reading model is that it emphasizes higher level skills as the prediction of meaning by means of context clues or background knowledge at the expense of lower skills like the rapid and accurate identification of lexical and grammatical forms (Eskey, 1988). It also tends to deemphasize the perceptual and decoding dimensions of that process. Therefore, this model is good for the skilful, fluent reader for whom perception and decoding have become automatic, not for the less proficient, developing reader. In this regard Weber (1984) argues that a top-down model of reading is essentially a model of the fluent reader and does not account for all the needs of students who are acquiring reading skills. Concerning second or foreign language learners, in many cases and for many texts, the reader has little knowledge of the topic and cannot generate predictions. Even if a skilled reader can generate predictions, this would take much longer than it would to recognize the words (Stanovich, 1980).

2.4.4. The interactive models

2.4.4.1. What is an interactive reading model?

An interactive reading model is a reading model that recognizes the interaction of bottom-up and top-down processes simultaneously throughout the reading process. It attempts to take into account the strong points of the bottom-up and top-down models, and tries to avoid the criticisms raised against each, making it one of the most prominent approaches to the theory of reading today. According to Rosenblatt (1994), the bottom-up and the top-down approaches are both right in the sense that readers do perceive in a bottom-up manner the letters and words in print and that they do use in a top-down manner their prior linguistic knowledge and the knowledge of the world to get meaning from the print.

Interactive approaches to reading (Rumelhart, 1977, 1985) acknowledge the simultaneous interaction of both lower-level processing skills (identification and decoding) and higher-level reasoning and comprehension skills (inferencing and interpretation). As Grabe (1991, p. 383) asserts, "reading involves an array of lower-level rapid, automatic identification skills and an array of higher-level comprehension / interpretation skills." The concept of interaction is based on the assumption that there is a complex cognitive psychological relationship between reader and text, and a simultaneous activation of readers' multiple component skills and their background world knowledge as they attempt to (re)construct the information present or implied in the text.

Two examples of interactive processing models are those theorized by Stanovich (1980) and Swaffar et al. (1991). Stanovich claimed that the development of reading fluency needs to be viewed as an "interactive-compensatory" model of individual differences wherein readers compensate for deficiencies at the word level (lower level) by relying more on context (higher level). Similarly, Swaffar et al. assessed readers based on their affective factors such as motivation and different learning styles, their background world knowledge, and their linguistic knowledge. Their integrated approach to language learning is founded on the belief that readers engage in reading for meaning.

One approach that proved to be popular and which was greatly influenced by K. Goodman's (1967) and Coady's (1979) psycholinguistic model, was the Constructivist Model advocated by Bernhardt (1986). This model includes the following six elements: prior knowledge, word recognition, phonemic/graphemic features, metacognition, syntactic feature recognition, and intratextual perceptions (i.e., "how the reader perceives and then reconciles each part of the text with the preceding and succeeding discourse context" Bernhardt, 1991, p. 122). The interaction of these factors would result in comprehension. This model contributed to reshaping researchers' view on reading which came to be regarded as involving readers, not just the reading text (Bernhardt, 1986, 1991b). The basic tenet of such an approach is that the reading process comprises the interaction of reader and text and that comprehension is subject to the influences of a wide range of factors (see next chapter). Another important contribution of this model is the keen observation that comprehension of a reading passage may be hampered when that passage contains unfamiliar cultural referents.

In sum, it can be argued quite convincingly that positing a solely bottom-up or top-down processing model for reading will fail to capture the complex interactive nature of the reading process. Alternatively, an interactive approach to reading appears to offer a better explanation of the cognitive processes believed to be at work here. Yet notwithstanding this model's improved explanatory power, it is unclear how the interactive, or indeed any of these processing models, can be translated into effective, simple-to-use teaching practices with long-lasting results.

It is also important to note here that many of these approaches overlap, thus making absolute distinctions and comparisons difficult to detect. Nevertheless, these approaches have underscored current research efforts; some of them have considerably influenced current thinking in the teaching of reading, while others continue to shape our understanding of reading and reading comprehension (Liontas, 2002). One of the most prominent theories on reading to date and which has had important pedagogical implications is the transactional theory of reading, writing, and written texts originally advanced by Goodman in 1967. Originally, Goodman referred to his model as a psycholinguistic one to reflect its language-processing essence. This model was refined throughout the following three decades and came to be known as the Transactional Socio-Psycholinguistic Model of Reading. This model is sketched below.

2.4.4.2. Transactional Socio-Psycholinguistic Model of Reading

According to the transactional theory, reading is not a personal event but rather a social event, an act of communication between the reader and the writer, or an act of

non-simultaneous communication with the reader on the one side and the behaviour residue of the writer on the other. Rosenblatt (1994) explains that every reading act is an event, or a transaction involving a particular reader and a particular pattern of signs, a text, and occurring at a particular time in a particular context. Instead of two fixed entities acting on one another, the reader and the text are two aspects of a total dynamic situation. The "meaning" does not reside ready-made "in" the text or "in" the reader but happens or comes into being during the transaction between reader and text. Thus, within the transactional theory of the reading process the meaning does not reside solely in the text, nor does it reside only in the reader. Rather, when reader and text interact -- i.e., when a reader reads a text -- the meaning that the reader constructs from the text is a third entity, not exactly what the author created nor what the reader could have created without a text. Therefore, each reading -- whether by the same reader or by a different one -- involves a different transaction and, thus, a somewhat different meaning.

The transactional theory provides a useful perspective for reading comprehension. It questions the nature of the meaning of a text, which has long seemed static, unquestionable and fossilized in black and white in the bottom-up models of reading. According to the transactional theory (Weaver, 1994), the text is simply a set of signs or marks, which are potentially capable of being interpreted as verbal symbols, and which do not possess a meaning that can be imposed on all readers. A text is viewed merely as one participating party in the communicative act of reading and the meaning of the text has to be constructed or negotiated through the

transaction between the reader and the text. It appears that, unlike the top-down approach which considers meaning as something that the reader brings to the text, transactional theory of reading attaches special emphasis to the process of negotiation for the meaning between the reader and the text.

According to Goodman, reading is *making sense* of the text. The text is not viewed as controlling a passive reader; instead, the reader is seen as an active user of language. As a reader reads, the text must be sampled and interpreted through the reader's unique personal background knowledge and experience. The active involvement of the reader with print and meaning, not words. The reader is rather seen as central to the act of reading: the reader's reaction to particular pieces of print on a page that may or may not have connections to larger contexts is of secondary importance.

Furthermore, in making sense (i.e., in constructing meaning), the reader constructs his or her own text parallel to the printed text. The reader behaves in such a manner because a text is never a complete representation of the writer's meaning. Since much needs to be inferred during the reading process, the reader can only comprehend that which s/he brings to the "transaction" of reading. This explains the construction of a parallel text by the reader using the smallest amount of available text information and one's own existing linguistic and conceptual knowledge to build meaning. Goodman (1992, p. 19) argues, "construction of the text is a necessary concomitant of the construction of meaning." Therefore, effective reading involves making sense of print, *not* accurate word identification. Viewed under such a perspective, it becomes obvious that *meaning is in both the reader and the writer*, and not inherently in the text itself.

Goodman (1992) investigated the meaning-making processes of readers while they read out loud whole stories that they had not seen before. During their reading, it was discovered that readers made miscues, not errors as would be argued by those who believe that reading is a sequential word identification process. In other words, Goodman found that his readers produced unexpected responses to the text. As a result, the miscues he discovered became for him "windows on the reading process" (Goodman, 1992, p. 3). If readers use cues in the text to construct meaning, he consequently hypothesized, reading cannot possibly be a passive process. Instead, reading must be a receptive language process where readers are active users of language.

In sum, transactional theory highlights the interaction that must be involved between the reader and the writer via the text in reading. For an interactive approach to reading comprehension, therefore, it is necessary to include not only the interaction between bottom-up and top-down processing and the interaction among different sources of prior knowledge but also the interaction between the reader and the writer via the text (Grabe, 1988).

2.4.4.3. Rumelhart's interactive reading model

In Rumelhart's view reading is at once a perceptual and a cognitive process and a skilled reader must be able to make use of sensory, syntactic, semantic, and pragmatic information to accomplish the task. These various sources of information appear to interact in many complex ways during the process of reading (Rumelhart, 1977/1995).

According to Rumelhart the interpretation of the meaning of what is read depends very much on the contexts in which the text is encountered and the understanding of information at one level of analysis can often depend on the understanding of information at a higher level. In his view, the interpretation of the meaning of the text may be said to be the product of the simultaneous interactions among all knowledge sources applied in the course of reading. In this model, the perception of words depends on the syntactic and semantic environment in which the words are encountered. Similarly, the perception of syntax relies on the semantic context in which the string of words appears. The same is also true for the interpretation of the meaning of the text, for it is dependent on the general context in which the reader encounters the text. Syntactic context effects are evidenced by oral substitution errors where an incorrect word is substituted for a correct one. Semantic context effects can be seen in our interpretations of words with two meanings, such as "bank". It follows, therefore, that reading comprehension may be perceived as a procedure of producing the most probable interpretation of the message encoded by the writer in the text, an interactive use of the sensory, syntactic, semantic, and pragmatic information available to construct the meaning of the text.

2.4.4.4. Critique of the interactive reading models

The main criticism that has been directed to interactive models of reading is that other psychological and socio-cultural contextual factors, such as attitude, interest, motivation, physical environment, time, place, etc., have all been left out, although they have been found to have some effects on reading comprehension (Chall & Dale, 1995;

Johnson and Pearson, 1984; Ruddell et al, 1994) and their pertinent role in L1 and L2 reading has long been acknowledged to have an effect on metacognition and text comprehension (Kern, 1988, 1989, 1992, 1994).

Conclusion

This section reviewed the different definitions of ‘reading’ available and which are generally divided into two major types: a) those that equate reading with interpretation of experience generally, and b) those that restrict the definition to the interpretation of the graphic symbols. Then it considered briefly the history of the reading process before moving on to the consideration of schema theory, as part of a psycholinguistic model of reading which holds that efficient comprehension requires the ability to relate textual material to one’s own knowledge. Some reading models were reviewed. The bottom-up model describes the reader as arriving at meaning by moving from letters to words to phrases and sentences and arriving at meaning. The top-down model describes the reader as deriving meaning primarily from predictions about the text and background knowledge. The interactive model posits that both processes work together.

In sum, the literature review reveals that reading is a complex phenomenon. A complete picture of reading can only be captured by considering the purpose for reading (why we read), the components of reading ability (what skills are involved), and the key processes involved in reading comprehension (how we read) (Grabe, 2002). Such issues

will be discussed in the next chapter on reading comprehension within the theoretical framework of interactive approaches and schema theory presented above.

CHAPTER 3

READABILITY

Introduction

It is almost a truism among educators and reading specialists that one of the basic tenets of reading pedagogy is to match the ability level of the students as closely as possible to the difficulty level of the text. It is also axiomatic that a too difficult text leads to frustration and a too easy one leads to student boredom. For this reason researchers have endeavoured to find ways of accommodating readers' abilities and reading material level of difficulty. In such a quest, they attempted to identify in measurable terms the factors that affect success in reading and understanding a text, i.e. what makes some texts easier to read than others and tried to develop valid means which serve such a purpose. In other words, they attempted to find out what makes a text more readable than another and hence develop tools that assess it.

The purpose of this chapter is then to review the literature on the assessment of text difficulty. The bulk of this chapter is devoted to the landmarks studies on readability and the controversy regarding the formulas. Starting from a consideration of how text difficulty has been measured, it moves on to how the concept of readability has been defined with a mention of a historical review of the various ways whereby text difficulty and readability have been assessed. The literature on the variables affecting readability and the methods of readability measurement are reviewed along with criticism directed at the use of readability formula. Finally it discusses the validity of readability formula in EFL contexts.

3.1. Measuring text difficulty

A survey of the literature shows that there have been many attempts to assess the level of difficulty of reading material, using a variety of techniques. The most commonly used techniques may be arranged into five categories that follow:

Subject assessment;

Objective question and answer techniques;

Graphs and charts;

Cloze Procedure;

- Readability formulas.

3.1.1. Subject assessment

In the absence of convenient quantitative methods, assessments of the level of difficulty of reading material have involved the use of a subjective judgment. However, Chall (1958) reported that while teachers were consistent in their rankings of the difficulty of books, their matching of books and students were inaccurate and inadequate.

Smith and Dechant (1961) have summarized pertinent research and have noted that teachers and others concerned with students' reading have not been too successful in their judgments of the difficulty level in books. In 1971 Moyle reported two studies involving grading of books by committees of experienced teachers. The results of these studies show a pattern typical of assessment by panels - i.e. the grading done by the committees is much more consistent than by individuals (Gilliland, 1972). Studies

reported by Chall (1958) demonstrated the unreliability and inconsistency of even expert judgment.

3.1.2. Question and answer techniques

This procedure, which is essentially measuring comprehension of content, while being more impartial and controlled than subjective estimates, has several limitations which restrict its utility. One is that multiple choice questions are affected by the ability of the reader to make an inspired guess based upon an imperfect understanding of content. Thus, they provide a somewhat inadequate measure of the level of difficulty.

3.1.3. Graphs and charts

These have been used in an attempt to facilitate the assessment of the level of difficulty. Tables offer the advantage of requiring little or no calculation. They are easier for preparing and evaluating data than formulas. Pry (1972) has developed a table or graph for use in estimating difficulty levels of reading materials.

3.1.4. Cloze Procedure

This procedure was first published by Taylor in 1953. Instead of using a specific formula, the Cloze Test involves individuals filling in missing words within a text. The word cloze is related to the concept of closure, the human tendency to complete a partly finished pattern. The theory originated in Gestalt psychology and assumes that in figuring out the missing word, the mind goes through a process of sampling, predicting, testing, and confirming the appropriate word choice. The theory holds that the higher the individual's reading ability, the greater the success of predicting the missing words. If we know the reading ability of the population taking the Cloze Test, which we could

if they were a representative sample of a given group, then the test can then be used to determine the reading difficulty of the text used. In the original Cloze, the deleted words from the text are taken at regular intervals, and every fifth or tenth word is standard. The reliability of the test is increased the more missing words there are, with a minimum of at least 50 is recommended. Given this, to obtain a reliable outcome from the test, a text of at least 250 words would be required. The accuracy of the test would also increase if different versions of text were used and tested by different groups of readers. Each version would have a different sequence of deleted words so that all words within the text were deleted within the complement of versions.

Like any test assessing reading difficulty, the problem arises over what is considered a successful completion of the text: inserting 50% of missing words, 75% or 100%. Today cloze procedures are considered more suitable to assess readers' abilities than to measure the difficulty level of text. The cloze technique does not predict whether the materials are comprehensible; it is an actual try-out of the material. It tells you whether a particular audience group can comprehend the writing well enough to complete the cloze test (Rabin, 1988).

3.1.5. Readability formulas

These are methods for measuring text difficulty. They are mathematical equations based upon an analysis of easily identifiable aspects of the text mainly vocabulary difficulty and sentence complexity. Before considering them, the term “readability” must be defined and the factors influencing text difficulty must be discussed.

3.2. What is readability?

Readability has been defined in various ways and it seems that there is no consensus on the exact definition of the concept of readability. Consequently, various definitions have been proposed. Right from the onset it is possible to say that simply defined, readability refers to the ease with which a reader is able to understand a text. It is often confused with legibility, which concerns typeface and layout. For the purpose of this study, it is necessary to mention a few definitions that have been put forth by the leading scholars.

The Dictionary of Psychological Terms defines readability in the following terms: “Readability is the quality of a written or printed communication that makes it easy for any given class of persons to understand its meaning, or that induces them to continue reading.”

Klare (1963, p. 27) defines readability as the “the ease of understanding or comprehension due to style of writing”. This definition focuses on writing style, in contrast to factors like format, features of organization and content (DuBay, 2004). In contrast, McLaughlin (1968) stresses the importance of the reader taking into account the importance of specific reader characteristics, such as reading skill, motivation, relevant knowledge, and how these interact with the text. He defines readability as “the degree to which a given class of people find certain reading materials compelling and, necessarily, comprehensible” (p.185). On the other hand, Burton (1991) focuses on the interaction between the reader and the text: “Readability is concerned with the degree to which the reader can share meaning with the writer” (p.1).

Maybe the most comprehensive definition of readability is that of Chall and Dale (1948,) for whom readability is:

“The sum total (including interactions) of all the elements within a given piece of printed material that affect the success which a group of readers have with it. The success is the extent to which they understand it, read it at optimum speed and find it interesting”. (p.11)

The definitions above differ in the type of variables they mention as intervening in success in reading. However, a common denominator to the above definitions is understanding or comprehension on the part of the reader. It can then be safely inferred that the term readability refers to all the factors that affect success in reading and understanding a text, or as Wimmer and Ominick (2005) put it, it is the “sum total of the entire elements and their interactions that affect the success of a piece of printed material” (p. 331).

It should be pointed out that in the literature the concept of readability has often been used as though it describes the same phenomenon as understandability or comprehension (Adelberg and Razek 1984, Klare 1984). Jones (1997) disagrees with this point of view and argues that readability refers to qualities of texts while understandability depends on the reader’s background, prior knowledge, interest, and other factors. Readability has also been viewed as a larger concept that includes comprehension (Harrison 1980). In fact, it appears from the literature review that readability has two common meanings, one applying to document design, the other to

language. Readability as it is applied to document design is closely associated with legibility. Legibility refers to the arrangement, style, and appearance of printed words. It is concerned with typeface and layout, i.e. typography.

It pertains to the visual and perceptual aspects and so refers to the quality of handwritten or printed material that is clear enough to be read. It concerns the attribute of characters that makes it possible for each one to be identifiable from others. This depends on such features as form of characters and the amount of space between characters (Sanders and McCormick, 1993). Thus, when the shapes of different letters within one typeface can be discriminated fairly quickly, that typeface is said to be highly legible. A distinction needs to be made between *character legibility* and *legibility of printed text*. Character legibility is the ease with which a person can identify an individual character as a particular letter. Legibility of text refers to the ease with which groups of characters are correctly identified as a word, with the result that the reader perceives meaningful sentences. In this sense, it refers to how easily letters and words can be recognized (Meade & Smith, 1991).

High legibility is very important for books. If a text is not very legible, this will reduce the reader's reading speed and give him a feeling of uneasiness resulting in an increase of the mental effort needed to identify the letters correctly. Legibility is affected by level of illumination, background contrast and reader fatigue.

Readability as it is applied to language is concerned with the comprehensibility or understandability of a piece of written text and which Selzer (1983, p. 73) defines as: "...the efficiency with which a text can be comprehended by a reader, as measured by

reading time, amount recalled, questions answered, or some other quantifiable measure of a reader's ability to process a text..." For Donald (1993, p. 48): "Readability means understandability. The more readable a document is, the more easily it can be understood."

In this study the terms readability, understandability, and comprehensibility are used as synonyms and are therefore used interchangeably. It is this type of readability as it pertains to language, not as it pertains to document design that is considered here, and this is the variety of readability that the readability formulas purport to measure.

3.3. Readability: historical overview

According to Klare (1963, 1974, 1984), concern about text readability can be dated back as far as eleven thousand years ago. Modern interest in readability began in the late 19th century. One of the early landmarks of readability studies were L. A. Sherman's statistical analysis of literature, the vocabulary-frequency lists, and the classic readability formulas.

In 1880, Sherman, compared average sentence length in older prose with more popular modern one. He noticed a progressive shortening of sentences over time. He decided to look at this statistically and began by counting. He found that sentence length averages shortened over time: 50 words per sentence for Pre-Elizabethan times, 45 words per sentence for Elizabethan times (1558– 1603), 29 words per sentence Victorian times (1837-1901), and 23 words per sentence in Sherman's time. In our time, the average is down to 20 words per sentence. He concluded that shorter

sentences and concrete terms increase readability (DuBay, 2004). Sherman's findings indicated that the length of a sentence affects readability, that a given writer is consistent in the average length of his sentences, and that sentence complexity decreases as sentences get shorter (Sherman, 1893). These same three factors have played an important part in the development of readability formulas since almost every readability formula uses sentence length as a measure of difficulty.

The 1920's witnessed a re-emergence of interest in America. Kitson (1921) is considered as its pioneer. He used the number of syllables in a word and the number of words in a sentence as indices of the relative difficulty of newspapers and magazines. Lively and Pressley (1923) used a word frequency index based on Thorndike's *Teacher's Word Book* (1921) to estimate vocabulary difficulty. The early readability study and the original reasons for development of readability tests rose from the need of American junior high school science teachers to provide them with books which let them teach scientific facts and methods rather than getting entangled in teaching the science vocabulary necessary to understand the texts. Thus, the first readability studies were conducted by asking students, librarians, and teachers what they believed made texts readable. Research broadened during the 1930's and throughout the 1940's, deriving formulas that could accurately predict readability using the least number of factors. The literature review reveals that over the past fifty years, the concept of readability has been revitalized and the notion that textbook reading difficulty needs to be matched with student reading ability has been emphasized by readability investigators.

According to DuBay (2004), history of concern about readability is distinguished by two clear-cut periods. The first era, known for the classic or early readability studies, started in the late 19th century and concluded in the 1940's. Such a period culminated with the publication of the Flesch and Dale-Chall formulas which gained popularity among publishers, educators, and teachers who were concerned with finding practical methods to match texts to the skills of readers, both students and adults. The second era of new readability studies began in the 1950's and was influenced by the contributions of linguistics and cognitive psychology. Researchers explored how the reader's interest, motivation, and prior knowledge affect readability. These studies in turn stimulated the creation of new and more accurate formulas (DuBay, 2004).

Another milestone in the history of readability studies was Thorndike's *Teacher's Word Book* (1921) which was the first extensive listing of words in English by frequency. It provided teachers with an objective means for measuring the difficulty of words and texts. It laid the foundation for almost all the research on readability that would follow as it suggested, for the first time, a means to measure difficult words through mathematical formulas (Chall, 1988; Klare, 1963). Thorndike's work was based on the observation that the more frequent a word is used, the more familiar it is and the easier to use.

Thorndike's book was followed by another landmark work by G. K. Zipf with the publication of *Human Behavior and the Principle of Least Effort* (1949), in which he enunciated a mathematical relationship between the hard and easy words, called Zipf's Curve. Another researcher, Johnson (1946, cited in DuBay 2004) found that twenty-five

percent of the 67,200 words used in the 24 life stories written by university freshmen consisted of these ten words: *the, I, and, to, was, my, in, of, a, and it*. The first 100 most frequent words made up almost half of all written material. The first 300 words made up about 65 percent of it (Fry et al, 1993). It was shown that the first words learned are the simplest and shortest. These first, easy words were also the words used most frequently. This notion of energy saving became one of the bases of research on the frequency of words. Reviewing research on word frequency, Klare (1968) concluded:

Not only do humans tend to use some words much more often than others, but they also recognize more frequent words more rapidly than less frequent, prefer them, and understand and learn them more readily. It is not surprising, therefore, that this variable has such a central role in the measurement of readability. (p. 65)

Early researchers studied surface characteristics of written texts to determine the extent readers could comprehend texts. They considered any measurable elements of writing such as the number of personal pronouns in the text, the average number of syllables in words or number of words in sentences in the text. Then they compared the data with certain predetermined standards. One such standard was tabulating the average grade level of students who could correctly answer a certain percentage of questions from the text. Researchers judged the characteristics with the most accurate standards as indicators of readability, which they developed into readability formulas. There ensued the design of mathematical equations which correlate these elements.

Designers of early formulas began with 289 elements of content, style of expression and presentation, format and organization, and subsequently reduced them down to the factors which could be counted most reliably such as number of personal pronouns, average number of words in a sentence, percentage of different words, and number of prepositional phrases. However, ultimately, these factors were reduced to two: sentence length and word frequency because these two factors consistently rated the highest on regression equations to predict comprehensibility (Zakaluk & Samuels, 1997).

Beginning in the 1950's and with new insights from linguistics and cognitive psychology, new dimension such as reader's interest, motivation, and prior knowledge became to be explored as variables affecting reading comprehension and readability. Factors investigated were described as "semantic" if they concern the word used and "syntactic" if they concern the length or structure of sentences.

This era was characterized by a certain number of features (DuBay, 2004). The number of researchers and studies that took interest in readability grew considerably (Klare 1952, 1963, 1974-75, 1984; Chall, 1958; Chall and Dale, 1995; Flesch, 1955, 1964, 1979). There were a growing number of researchers and studies that took interest in readability and attempted to improve readability formulas. There was also the introduction of the cloze test (see above). This was considered as a new tool to test the properties of texts and readers with more accuracy and detail. A certain number of studies went beyond surface text variables by looking at manner in which these reader variables affect readability. More recently, and with the advent of computers and the

generalization of their use, computational approaches to the study of readability emerged. Such issues are discussed with more details in the next sections.

There are at least 200 readability measures for texts, all of which primarily measure two aspects of the sentence: word difficulty, usually through word length or syllable count, and sentence difficulty, typically through sentence length (DuBay, 2004). The same researcher noted: “A vocabulary test on the meaning of words is the strongest predictor of verbal and abstract intellectual development. The knowledge of words has always been a strong measure of a reader’s development, reading comprehension, and verbal intelligence” (p.12).

Chall and Dale (1995, p. 84) wrote, “It is no accident that vocabulary is also a strong predictor of text difficulty.” It is generally admitted that the first words we learn are the simplest and shortest. These first, easy words are also the words we use most frequently. Johnson (1946) showed that twenty-five percent of the 67,200 words used in the 24 life stories written by university freshmen consisted of these ten words: *the, I, and, to, was, my, in, of, a, and it*. Fry et al, (1993) established that the first 100 most frequent words make up almost half of all written material. The first 300 words make up about 65 percent of it.

It is not the purpose of this study to review all the readability measures as it is beyond its scope. However, to acquaint the reader with how text difficulty has been measured the most popular readability formulas will be reviewed. But first what factors affect text difficulty, and consequently readability?

3.4. Factors influencing readability

It is mentioned above in the introduction to this chapter that researcher have attempted to identify in measurable terms the factors that affect success in reading and understanding a text, i.e., what makes some texts easier to read than others and tried to develop valid means which serve such a purpose.

The first real readability formula was that of Lively and Pressey (1923) which was based purely on vocabulary difficulty. Vogel and Washburne (1928) counted number of different words in a sample, the number of prepositions, the total number of words not on the Thorndike 10,000 most frequent word list, and the number of clauses in 75 sentences. Gray and Leary (1935) investigated more than 200 style elements and the relationships between them. By combining variables that were highly predictive but not related to each other they created a readability formula with five variables. These variables have been narrowed down to only a few which have high correlations with the others and so can be used to represent them. They grouped the variables under a) content, b) style, c) format, and d) general factors of organization, although their readability formulae took into account only variables listed under style. Chall (1958) included the reader and concluded that only four types of elements seem to be significant for a readability criterion, namely vocabulary load, sentence structure, idea density and human interest. This is completely in accordance with her previous definition of readability (see under 4.2 above).

Later the predictor variables were further narrowed to just two main text factors: vocabulary difficulty and grammatical difficulty. Depending on the formula, vocabulary

difficulty may be represented as word familiarity, average word length in syllables, proportion of long words, average word length in either characters or syllables, or proportion of monosyllable words. Grammatical difficulty is typically measured by the average number of words or syllables per sentence, based on a strong association of sentence length with, for example, the incidence of compound-sentence and embedded-clause constructions, which are much harder to count (Bormuth, 1969, 1971; Chall & Dale, 1995). Dale and Chall (1948) also used two variables (average sentence length and word familiarity) in their formula. Many other formulae of varying complexity followed, notably those of Fry (1964, 1977), Bormuth (1966), Coleman & Liao (1975). Some researchers proposed other factors to insert in formulae or different methods of establishing readability such as (to name only two) counting abstract words (Flesch, 1943, Cohen, 1975) and propositional density, and inferences, (Kintsch, 1974).

It is generally accepted that readability indices and factors influencing text difficulty are of two sorts: Some of them are directly related to the text, others are directly related to the reader. Text-related factors can be subdivided into physical factors and linguistic factors. The discussion of all the variables which affect readability is beyond the scope of this study. The factors of interest, that is the variables which affect reading comprehension, will be discussed in the next chapter.

3.5. Classic readability formulas

It appears from the above that readability formulas are predictive devices that offer estimates of the difficulty of a text. The purposes for which readability formulae

are designed might vary, but they tend to have one objective in common, that is, to predict the difficulty of the text for the intended readers without their actual participation in the course of readability evaluation. The prediction of the text's difficulty is expressed as a grade level or a score on some set scale (Greenfield, 2004). Text difficulty is measured according to the familiarity of the vocabulary to the reader. If a word is familiar to the reader, its vocabulary difficulty is low, and vice versa. The common measurement of vocabulary familiarity is word frequency. By word frequency, is meant the frequency with which a given word occurs in a sample of the target language. Word frequency and vocabulary difficulty are inverse in variation. The higher the word frequency, the smaller will the vocabulary difficulty be, and vice versa (Lin, 2002, p.170)

According to Hayden and Prichard (2008), in the United States of America, many government agencies now require that documents, such as loan agreements, rental agreements, and property purchase contracts meet specific readability levels. Over thirty states have some form of plain language, or readability component in their insurance laws, and a number of states specify that insurance policies must be written at a minimum Flesch Reading Ease score (see below) to be judged “sufficiently readable”.

Out of the vast array of readability formulas, a few have gained popularity over the others and are exposed below. The first mentioned are milestones from the early period, the classic readability formulas. The last two are examples of more recent readability studies involving computational approaches and which were developed thanks to the advent of computers.

3.5.1. The Flesch Reading Ease formula

This formula was developed in 1948 and is one of the most widely used readability formulas in use today. It is a standard used by many government agencies and is the U.S. Department of Defense standard. It is considered suitable for all kinds of text. This formula uses average sentence length and number of syllables. It multiplies the average number of words in the sentence by 1.015 and the total syllable count by .846. The sentence length and syllable count are then added and subtracted from 206.835 to arrive at a readability score (DuBay, 2004, p. 20).

The formula reads as follows:

$$\text{Flesch Reading Ease Score} = 206.835 - 84.6 \times \text{ASW} - 1.015 \times \text{ASL}$$

ASW = average number of syllables per word

ASL = average sentence length

It measures reading from 100 (for easy to read) to 0 (for very difficult to read). The higher the score, the easier the document is to read. It is based upon a 50% comprehension rate. A zero score indicates text has more than 37 words on the average in each sentence and the average word is more than 2 syllables. This formula also has been incorporated into most word processing programmes including MS Word.

3.5.2. Flesch-Kincaid Grade Level Formula

The Flesch-Kincaid Grade Level formula, developed for the U. S. Navy in 1974, is a modification of the Flesch Reading Ease formula. The Flesch-Kincaid formula is a U.S. Department of Defense standard and is used by the U.S. Navy. This formula is

also widely used elsewhere and is considered to be well suited for business publications and journals.

$$\text{Flesch Kincaid Grade Level} = (0.39 \times \text{ASL}) + (11.8 \times \text{ASW}) - 15.59$$

ASL = average sentence length

ASW = average number of syllables per word

This formula gives a result in terms of US school grade levels, which will not be directly comparable with Algerian school years.

3.5.3. The Dale-Chall Formula

This is another very influential formula, published by Edgar Dale and Jeanne Chall in 1948. It was designed in such a way as to correct certain flaws of the Flesch Reading formula. The formula uses two variables, average sentence length and a percentage of difficult words. The idea behind this formula is that readers typically find it easier to read, process and recall a passage if the words are familiar. The Formula uses a count of “hard” words. It calculates the grade level of a text sample based on sentence length and the number of “hard” words. These “hard” words are words that do not appear on a specially designed list of common words familiar to most students. It is based on the Dale list of 3,000 familiar words. In addition to the percentage of words found on the Dale list, the formula uses average number of words per sentence (Chall and Dale 1995).

The formula is as follows:

$$\text{Raw Score} = 0.1579 \text{ PDW} + 0.0496 \text{ ASL} + 3.6365$$

Raw Score = Reading Grade of a reader who can answer one-half of the test questions on the passage.

The first measure (PDW = Percentage of Difficult Words) is the percentage of words in the passage not found on the Dale Word List. The second measure (ASL = Average Sentence Length) is the average number of words per sentence, which is calculated by dividing the total number of words in the passage by the total number of sentences in the passage.

The main idea behind the Dale-Chall readability formula is that, should a text be written with familiar words, it becomes easier to read and, therefore, expressed ideas are easier to understand and remember.

3.5.4. The Gunning Fog Formula

The Fog Formula is another commonly used readability measure, published by Robert Gunning in 1952. It became popular due to its ease of use. The formula is based on two variables, average sentence length and the number of words with more than two syllables per 100 words (DuBay, 2004, p. 24). The output is the Fog Index expressed in a grade level score. It estimates the number of years of education that a reader needs to understand the material. This formula calculates a grade level based on the average sentence length and the number of complex words, where “complex” is defined as words containing three or more syllables. The Gunning Fog Formula targets 90% comprehension. It is widely used for evaluating business publications and journals.

The Fog Index is given by the following equation:

$$GL = 0.4(ASL + \text{hard words})$$

Where:

GL = grade level.

ASL = as in the previous formula.

Hard words = number of words with more than two syllables per 100 words.

3.5.5. The Readability Graph

Fry developed a "Readability Graph" in 1965 for predicting readability. He used the common formula variables of syllables per 100 words and words per sentence. The user marks the counts of the variables on a graph and then reads the readability grade score directly from it.

3.5.6. Coleman Liau Formula

The Coleman-Liau readability formula (1975) calculates grade level based on average sentence length and average number of characters per word instead of syllables per word. Like the preceding formulas as well as many of the others, its output approximates the U.S. grade level thought necessary to comprehend the text. Coleman-Liau relies on characters instead of syllables per word. This formula is considered appropriate for text written for American 4th grade to college level readers.

The Coleman-Liau Index is calculated with the following formula:

$$CLI = 0.0588L - 0.296S - 15.8$$

Where:

L is the average number of letters per 100 words and S is the average number of sentences per 100 words.

3.5.7. Automated Readability Index Formula

The Automated Readability Index (ARI) formula was developed by Smith and Senter in 1967 to assess U.S. Air Force materials. Like Coleman Liau formula, this test calculates the U.S. grade level of a text sample based on average sentence length and average number of characters per word, and hence is easier to automate. The Automated Readability Index formula is considered most suitable for technical documents and manuals.

$$\text{ARI Grade Level} = (4.71 \times \text{ACW}) + (0.5 \times \text{ASL}) - 21.43$$

ACW = average number characters per word ASL = average sentence length

3.5.8. The Bormuth Formula

This readability measure is considered to be one of the most accurate. It was developed by John Bormuth in 1969 and was later adapted by the U.S. College Entrance Examination Board in 1981. It uses three variables, average sentence length in words, average word length in characters and the number of words on the original Dale-Chall list of 3000 words. Bormuth used 330 passages which were about 100 words each and which ranged in difficulty from first grade to college and covered a wide range of subject matter. The cloze procedure of deleting every fifth word was used as the criterion of difficulty (Klare, 1980). The findings of Bormuth about the reliability of the classic variables were confirmed by MacGinitie and Tretiak (1971) who concluded that average sentence length is the best predictor of syntactic difficulty (DuBAy,2004).

3.5.9. The SMOG Formula

Published by G.H. McLaughlin in 1969, the SMOG formula (Simple Measure Of Gobbledygook) is one of the simplest formulas to use and was based on the idea that semantic and syntactic difficulty predictors should be multiplied, instead of added. The formula is calculated based on sentence length and number of complex words, where “complex” is defined as words containing three or more syllables. This formula returns grade level scores. It assesses the educational level needed to understand material at a 100% comprehension level

The SMOG Formula is:

$$GL = 3 + PC$$

Where:

GL = grade level

PC = polysyllable count (number of words with more than 2 syllables in a 30 sentence sample)

In fact, the advent of computer word processing has made the formulas more accessible than ever. In older versions of Microsoft Word, it was possible to get a readability report on a Word document using the Flesch, Flesch-Kincaid, and Bormuth formulas built into that application. With Microsoft Word 1997/2007, the number of formulas has been reduced to include only the Flesch and Flesch-Kincaid.

3.6. Computational approaches

Beginning in the 1980's, the advent of computers has made it easier for researcher to apply compute and apply the formulae discussed above. Such an ease of calculation boosted computational linguistics. Aspects of a text such as word frequency, syntactic complexity, and many other indices were difficult to calculate a few decades years ago were computed effortlessly. Researchers were able to explore other variables with other approaches. Examples of such approaches are The Lexile Framework for Reading and the Strathclyde Readability Measure

3.6.1. The Lexile Formula

A more recent application of traditional readability formulas is known as the Lexile Framework and which has been defined by its designers (Wright and Stenner, 1998) as a scientific approach to reading and text measurement. Stenner (2004, p.2) wrote:

The Lexile Framework for Reading purports to measure in a common unit, called Lexiles, the traits of reader ability and text readability.

Based on these measures, reading comprehension can be calculated based on the gap between reader ability and text readability. When reader ability far exceeds text readability, then comprehension should approach unity. Conversely, when text measure far exceeds reader measure, then the probability of little or no comprehension should approach unity.

The Lexile Formula is based on two components. The first is a measure of sentence length, which by hypothesis indicates the level of syntactic complexity. It is based on what the proponents of the formula call the syntactic axiom: the shorter the sentence, the easier the passage is to read. Worded differently, the length of a sentence is a good indicator of how hard it is to read. Longer sentences take longer to read and require more concentration to understand. The longer a sentence is, the more likely it is to contain multiple phrases and clauses, which will require the reader to comprehend a number of ideas as well as the relationship between them.

The second component is a measure of semantic complexity, based on word familiarity in terms of inverse frequency of appearance in a corpus of text works. Again, the more frequently a word is encountered the more chances there are for the reader to know its meaning.

Lexiles evaluate the semantic difficulty of words by their frequency in standard written text. The authors have developed a process that assigns a value to the reading capacity of a person. The central idea is that, when a person is reading with 75% comprehension, they are at optimal reading capacity. The process therefore assesses a person's level of reading comprehension, and then calculates what they called the Lexile value of texts they can read with 75% comprehension. This is the measure of their reading capacity.

The Lexile Framework for Reading is made up of Lexile reader measures and Lexile text measures, both of which are put on the Lexile scale. A Lexile measure is defined as the numeric representation of an individual's reading ability or a text's

readability (or difficulty), followed by an “L” (Lexile). There are two types of Lexile measures: *Lexile reader measure* and *Lexile text measure*. A Lexile reader measure typically is obtained when an individual completes a reading comprehension test. Once a field study has been performed to link Lexile Framework with the test, the individual’s reading score can be reported as a Lexile measure. A Lexile text measure is obtained by evaluating the readability of a piece of text, such as a book or an article. The Lexile Analyzer, a software programme specially designed to evaluate reading demand, analyzes the text’s semantic (word frequency) and syntactic (sentence length) characteristics and assigns it a Lexile measure. .

The Lexile scale runs from below 0L (zero Lexile) to above 2000L. Scores 0L and below are reported as beginning reader and scores above 2000L are reported as high Level. By way of example, the novel *Gone with the Wind* by Margaret Mitchell scores 1100L, *A Tale of Two Cities* by Charles Dickens scores 990, and American university textbooks 1250L-1450L.

3.6.2. The Strathclyde Readability Measure

The Strathclyde Complexity Measure (Weir and Ritchie, 2006) has been developed within the computer science department at the University of Strathclyde. This method has been developed with technology-based testing in mind and to account for the shortcomings of other traditional means. The complexity of words is derived using a frequency list. This provides an indication of word commonality, which is regarded by its designers as indicative of the word’s likely perceived difficulty. To account for the complexity of a text, corpus analysis was used to identify words which

appear less often than others from a frequency list created from the target text: the less a word appears in the text the more complex it is. The frequency list used contains approximately one million words and is based on the British National Corpus. The output of this approach is similar to other approaches that is, a numerical score is generated between 0 and 100. The score position on the scale directly indicates the relative complexity of the considered text. However this approach does not make use of educational grade level as some of the other approaches do.

It should be pointed out that the usefulness and validity of such formulae were called into question (Irwin and Davis, 1980; Davison & Kantor, 1982).

3.7. Limitations of Readability Formulas

For a time during the 1980s, readability formulas came under attack when viewed from the viewpoint of psycholinguistic theories of reading (Bruce & Rubin, 1988; Rubin, 1985; Bruce, Rubin, & Starr, 1981; Smith, 1988). Nevertheless, the formulas have survived and are still widely used on account of their consistently high predictive validity. That is to say, they have been found empirically to do a good job of discriminating text difficulty even though it is not obvious why they should or how they could (Chall & Dale, 1995; Fry, 1989).

According to Connaster (1999, p. 274): “Readability formulas fail to predict text difficulty’ and he cites an experiment to demonstrate that ‘text difficulty’ is ‘a perception of the reader and therefore cannot be objectively calculated by counting syllables, word length, sentence length, and other text characteristics”

The field of readability research has witnessed several, quite heated debates about the shortcomings of readability formulas. Some of the main limitations of readability formulas that have been identified in the relevant literature are the following:

- They cannot measure conceptual difficulty: No formula takes into account the content of the document being evaluated.
- They do not take into account word order: Readability scores remain the same even if the text is scrambled. For example, the phrase “the baby is scared by the dog” will have the exact same score with the phrase “the dog is scared by the baby”, or even “baby by the dog is the scared”. It is clear that the second phrase means the opposite and the third is incomprehensible but readability formulas are unable to detect that.
- Readability formulas do not yield similar results for the same text because they use different variables and different criterion scores (Bruce & Rubin, 1988).
- They do not discriminate between readers; they assume that they are all alike.

Readability formulas make no distinctions based on reader’s characteristics. They do not take into account the reader factor (Kintsch & Vipond, 1979). That is, “they take no account of differing purposes, maturity and ability of readers” (Redish, 2000, p. 134). Readability formulas can’t measure the context, prior knowledge, interest level, difficulty of concepts, or coherence of text.

Zakulek and Samuels (1958) are quite critical of readability measures, concluding they “are inadequate because they consider only one source of information - that contained on the printed page. They say:

The same text materials may be very easy for one reader yet extraordinarily difficult to another. For example, prior knowledge will greatly influence how well a reader can understand text dealing with a particular topic. While a coastal engineer might easily read a technical report on the topic of tidal erosion, most of us would find the same text incomprehensible. (p. 32)

The more serious criticism of the underlying tenets of readability formulas is that “the practice of using simpler words and shorter sentences is unlikely to enhance readability ... because it does not take into account the numerous other text variables that are known to affect comprehension” (Singh, 2000: 216). The formula cannot measure the circumstances in which the reader will be using the text or form - both the psychological and the physical situations.

As discussed above, research on schemata has shown prior knowledge and background knowledge have been shown to be equally important factors in comprehension and retention of information. The ease of reading that the reader experiences is also directly influenced by the writer's use of physical, syntactic, semantic and contextual cues which cannot be measured by these tests. Such clues include the legibility factors mentioned above. Bruce and Rubin (1988) summarized the

situation by stating that readability formulas have limitations because formulas do not measure all the factors that influence the comprehensibility of a text.

3.8. Validating Formulas for EFL

Readability formulas were originally designed to predict difficulty in first language reading material. Historically, they were based on first language reading data. Some researched endeavoured to find out whether these formulas are also valid for EFL/ESL readers. The most salient ones are Hamsik (1984), Brown (1998), and Greenfield (2003). These studies addressed this issue with divergent results. Brown (1998) found that classic formulas were not very accurate predictors of EFL difficulty, while Hamsik (1984) and Greenfield (1999) found that they predicted for EFL about as well as they did for native English readers.

3.8.1. Hamsik's Study (1984)

Hamsik (1984), investigated the ESL validity of the Flesch, Dale-Chall, Fry, and Lorge readability measures. Hamsik gave cloze tests on 18 academic passages to 40 Intensive English Center students at an American university. The students are described as being from the Middle East, South America, and "the Far East." Hamsik found significant positive correlations of .775 to .819 between the rank orders of difficulty of the passages as indicated by the cloze scores and as predicted by each of the four readability measures. On the strength of this evidence, Hamsik concluded that "the four readability formulas and graphs...do measure readability for ESL students and that they can be used to select material appropriate to the reading level of ESL students" (p. iv).

Hamsik's small heterogeneous sample of ESL readers did not permit discriminating any effect of first language background. With this in mind, Hamsik included among her recommendations one that "future studies of this sort should take account of L1 background" (p. 55). She also suggested that it might be possible to develop a readability index for ESL students that would be more accurate than existing formulas.

3.8.2. Brown's Study (1998)

Further investigation of ESL/EFL formula validity was Brown's 1998 article. In an earlier study of cloze item difficulty, Brown (1992) had administered cloze tests to nearly 2300 Japanese EFL university students. For the new study he reanalyzed the data for difficulty at the passage level and compared the observed mean cloze scores on the passages with scores predicted by six readability measures: the Flesch, Flesch-Kincaid, Fry Graph, and Gunning-Fog. Brown found Pearson correlations ranging only from .48 to .55, leading him to conclude, "...first language readability indices are not very highly related to the EFL difficulty" (p. 27).

To address this need, Brown developed a new formula using his observed EFL scores as the criterion, scaled to yield an EFL Difficulty Index ranging from 1 to 92. Multiple regression analysis found the best fit or most accurate prediction to be made using four text variables: syllables per sentence, passage frequency (how many times the deleted item appears elsewhere in the text), percentage of long words (seven or more letters), and the percentage of function words. The resulting formula, which he called the *EFL Difficulty Estimate*, yielded a stronger correlation with the observed EFL

scores than did the classic formulas in his tests. Brown speculated, "EFL/ESL readability might best be estimated separately for students from different language backgrounds" (p. 30). In other words, Brown suggests, we need to replace the classic readability formulas with new formulas specific to different language groups. His formula was offered as one that might be used with Japanese EFL. Brown's formula is as follows:

Brown EFL Difficulty Estimate

$$\begin{aligned} \text{EFL Difficulty} &= 38.7469 + (.7823 \times \text{Syllables per Sentence}) \\ &+ (-126.1770 \times \text{Passage Frequency}) + (1.2878 \times \% \text{ Long Words}) \\ &+ (.7596 \times \% \text{ Function Words}) \end{aligned}$$

3.8.3. The Miyazaki Study

The Miyazaki study (Greenfield, 1999) also involved Japanese university students and checked the Flesch Reading Ease and Flesch-Kincaid formulas along with the Coleman-Liau, New Dale-Chall, and Bormuth formulas. The EFL participants in this study were 200 Japanese students enrolled in a college in Japan. Careful randomized testing procedures were followed, based on Bormuth (1971). Fifth-word deletion cloze tests were constructed on 31 of the 32 Bormuth academic passages. Pearson correlations between observed EFL mean cloze scores and scores predicted by the formulas are .691 for the New Dale-Chall formula, .765 for Coleman-Liau, .845 for Flesch Reading Ease, .847 for Flesch-Kincaid, and .861 for Bormuth.

The Miyazaki formula is as follows:

Miyazaki EFL Readability Index

EFL Difficulty = $164.935 - (18.792 \times \text{Letters per Word})$

$- (1.916 \times \text{Words per Sentence})$

($R = .862$, adjusted $R^2 = .723$, $SE = 10.558$, $N = 31$, $p < .0001$)

This formula delivers a reading ease score on a nominal 100-point scale, 100 being easiest.

3.9. Beyond Readability

More recent studies conducted in the psychometric tradition have incorporated both readers' ability and characteristics of texts (e.g., Carver, 1977; Stenner, 1997). Carver (1977) maintained that the prediction of reading comprehension is enabled by the intersection of the ability level of the readers and the characteristics of text. However, in traditional readability studies, ability levels were often scaled using standardized tests, and these measures were not scaled in the same metric as text difficulty (Carver, 1977). In Carver's (1977) *National Reading Standards*, each grade ability score on the test (G_a) was calibrated to reflect a 0.50 probability that an individual can read and understand, or comprehend, the passages at the same grade of difficulty (G_d) according to the Rauding scale (reading and understanding scale). The Rauding scale measures the grade difficulty of reading and understanding. Grade 5 ability means that the average accuracy is likely to be 75 percent of grade 5 materials. A choice of a 75 percent target comprehension rate is obtained through empirical evidence (Squires, Huitt, & Segars, 1983; Crawford, King, & Brophy, 1975). The theoretical assumption of comprehension in using the Rauding scale is that the rate of reading is constant and the accuracy of comprehension during reading can be predicted

from a measure of material difficulty and individual ability. However, the Rauding theory was criticized because it is very mechanical, serial, and not comprehensive. In this sense “the theoretical assumption does not support every day reading phenomena such as skimming and studying” (Pearson, 1984).

Conclusion

This chapter provided an account of how measuring text difficulty has been approached through readability studies. Modern research into readability began in 1921 when Thorndike published a list of English words used more frequently in texts. Assumptions were made that the more frequently a word was used, the more familiar readers became with it and the easier it was to read. During the 1920's, research activities concentrated on looking for word factors that could be used to predict readability. Research broadened during the 1930's and throughout the 1940's, deriving formulas that could accurately predict readability using the least number of factors. More recent formulas have used computer software and corpus analysis. A wide range of measurement criteria have been employed to assess readability. Every conceivable linguistic factor has been included in the scores of different formulas, and some formulas include a dozen or more factors. However, the conclusion after almost a century of research was that the addition of more factors does little to more accurately predict readability and renders the formulas much more difficult to use. Put another way, “...counting more things does not make the formula any more predictive of reading ease but takes a lot more effort” (Stephens, 2000).

Text difficulties have been measured by semantic and syntactic factors. The most frequently used factors are word complexity and sentence length. Most readability formulas are based on two principal measures: (1) the number of polysyllabic words (more than 3 syllables), and (2) the average number of words per sentence. The underlying assumption of readability formulas employing these parameters is that monosyllabic words and short sentences are easier to read than complex sentences with a high proportion of polysyllabic words. Formulas range from those that are simple enough to be calculated by hand (e.g. SMOG, Fry) to more complex equations calculated by readability software or through the grammar analysis tools available in most modern word processor software packages (e.g. Flesch Reading Ease and Flesh-Kincaid formulas in MS Word and WordPerfect). The main criticism addressed to readability formulas is that word complexity and sentence length are rather simple “surface features” exclude any consideration of content, grammar, or organization, and reader variables. The effect of such variables on reading comprehension shall be considered in the next chapter.

It appears from the literature review of research on text difficulty and readability that vocabulary comes at the forefront of the factors which make a text easy or difficult. For this reason many researchers have turned to the investigation of the role of vocabulary comprehension. Such is the main topic of Chapter five.

CHAPTER 4

SECOND/FOREIGN LANGUAGE READING COMPREHENSION

Introduction

Right from the onset, it should be cautioned that research on reading comprehension in a second/foreign language has grown remarkably in the past few decades and is one of the most researched issues. Therefore, it is not easy to synthesize the array of research literature in L2/FL reading comprehension.

In the previous chapter, it was concluded that a complete picture of reading can only be captured by considering the purpose for reading (why we read), the components of reading ability (what skills are involved), and the key processes involved in reading comprehension (how we read) (Grabe, 2002). It was also mentioned that theories and models of reading have changed, from seeing reading as primarily receptive processes from text to reader to interactive processes between the reader and the text.

The goal of this chapter is to help the reader gain insights into the nature of reading comprehension. It aims at shedding light on the factors which influence and constrain reading comprehension bearing in mind that reading is an interactive process. It will review reading comprehension research in second/foreign language comprehension. A definition of reading comprehension will first be delimited as applicable to this project. Second, the relationship between reading comprehension in a first language and a second/foreign language will be examined, highlighting differences

and similarities as well as the influence of first language on second/foreign language. Finally, the factors that affect L2/FL reading comprehension are considered.

4.1. What is reading comprehension?

As already mentioned, reading comprehension is the obverse of reading as the purpose of any reading is to comprehend what is read. Moreover, just as for the definition of reading, there exists a plethora of definitions for reading comprehension. However, most of them seem to converge that reading comprehension is the process of simultaneously extracting and constructing meaning through interaction and involvement with written language (Snow, 2002, p. 11). Within information-processing theory, reading comprehension is currently defined as the process of relating new or incoming information to information already stored in memory (Anderson & Pearson in Carrell et al. 1984, among many others). The implication of such a definition is that reading comprehension is viewed as a constructive process, that is, a process that combines individual units to form new configurations. In fact, researchers in various fields such as computer science, artificial intelligence, cognitive psychology, sociolinguistics, and English as a second/foreign language have generally acknowledged that understanding is not a process of breaking complex units of language into simpler ones but, rather, a process of taking multiple units and building them into representations (Bernhart, 1986).

The literature reveals that there are two theories that describe reading comprehension: the first one views reading comprehension as a whole skill that cannot

be broken down into smaller sub-skills and depends on the information and knowledge in the reader's head that he/she uses while reading, i.e., it is the continuous interaction between the reader and material read. This theory emphasizes the use of teaching strategies that encourage extensive reading by the students.

The second theory views comprehension as a cognitive activity in which readers process different types of information to acquire knowledge about the topic presented in the text (Richek, List and Lerner, 1983). They decode printed symbols, derive word meaning, identify inter- and intra-sentence syntactic relationships, cohesive ties, and organizational structure; make inferences, and use their background knowledge about the topic to help them connect the new information in the text with what they already know about the topic. Lack of information in any linguistic part may result in comprehension difficulties. This theory views comprehension as an aggregate of sub-skills (Richek, List & Lerner, 1983).

In this study reading comprehension is defined (after Moore, 2005 p. 2) as 'the process of simultaneously extracting and constructing meaning through interaction and involvement with written language'. Within the scope of such a definition, reading comprehension entails the three elements mentioned above: the reader, the text, and the activity in which comprehension is a part. Any discussion of reading comprehension should include these elements each of which will be discussed separately with some details. Consideration of reader includes all abilities, knowledge, and experiences that a person brings to the act of reading. Text is broadly construed to include any printed text

excluding electronic text as this project investigates reading material contained mainly in textbooks. Activity includes the purposes, processes, and consequences associated with the act of reading. These three elements are inseparable from the sociocultural context within which they exist and by which all are influenced. However, a clear picture of L2 comprehension cannot be captured without considering similarities and differences between reading in L1 and SL/FL, as Grabe (2002) put it:

Useful extended definitions of reading can be developed at two levels:

1) purposes for reading (why we read), and 2) components of reading ability (what skills are involved). A yet more complete picture is created by considering key processes involved in reading comprehension (how we read). (p. 2)

The last issue (how we read) has been discussed in the previous chapter.

4.2. Purposes for reading: why we read

In everyday life we read for a variety of purposes in a variety of circumstances. Grabe (2002) provides a list of them: Scanning, skimming, reading for general understanding, reading to learn, reading to integrate information, and reading to evaluate critically, reading as search process, expeditious reading, reading to write, and reading while writing. In academic settings, students' purpose, as (Jordan, 1997:143) points out, can include:

- To obtain information (facts, data, etc)

- To understand ideas, or theories, etc
- To discover author's viewpoints
- To seek evidence for their own viewpoint (and to quote) all of which may be needed for writing essays, reports, etc.'

As we read for different purposes Grabe (2002) noted, we often vary the ways that we use the cognitive processes and knowledge resources central to reading. At the same time, the actual processes and resources for reading themselves do not generally vary just how they are used in combination. So we can still talk about reading as a single ability, while also recognizing levels of variability in response to differing purposes and tasks. To understand this consistency across purposes, a definition of reading must include a description of the component skills comprising reading abilities. But, to help us identify the skills and sub-skills involved in reading, it is also useful to try to distinguish between different levels of understanding of a text. It is necessary then to consider different levels of understanding then move on the skills involved in reading comprehension. Still, since the purpose for reading is closely related to the type of reading, a word must be said about the types of reading.

4.3. Types of reading

There exists in the literature good number of types of reading most of which are irrelevant to the present project as they do not apply to the Algerian context. The following is a quick overview of the most popular types of reading.

4.3.1. Skimming

Skimming refers to the way of reading in which readers quickly run their eyes across a whole text to quickly gather the most important information, or 'gist'. It is not essential to understand each word when skimming. Examples of skimming are: going over a newspaper quickly in order to get the general news of the day; going over a magazine quickly to discover which articles you would like to read in more detail; going over a business or travel brochures quickly to get informed; and going over an article to see if it may be of interest in your research. In all these cases it involves reading rapidly for the main points.

Skimming is a high speed reading process and involves visually searching the sentences of a page for clues to meaning. It is conducted at a higher rate 700 words per minute (wpm) plus than normal reading for comprehension (around 200-230 wpm) (sometimes known as rauding), and results in lower comprehension rates, especially with information-rich reading material. In classroom settings, skimming on its own should not be used when complete comprehension of the text is the objective. It is mainly used when researching and getting an overall idea of the text.

There are many strategies that can be used when skimming. Some people read the first and last paragraphs using headings, summarizes and other organizers as they move down the page or screen. One might read the title, subtitles, subheading, and

illustrations. Others read the first sentence of each paragraph. This technique is useful when one is seeking specific information rather than reading for comprehension.

Skimming works well to find dates, names, and places. It might be used to review graphs, tables, and charts

4.3.2. Scanning

Scanning or searching reading is a common reading activity when readers extract necessary pieces of information from a text without reading through the whole text. It refers to the way of reading in which readers run their eyes quickly over the text looking for the specific piece of information they need. Scanning is used on schedules, meeting plans, etc. in order to find the specific details required. Examples of Scanning are: the 'What's on TV' section of the newspaper; a train/airplane schedule and a conference guide. Other examples are looking up a word in a dictionary, searching through a telephone directory, searching for key words or ideas, and reading a timetable or advertisements for getting information

In short, scanning is reading rapidly to find a specific piece of information when one knows what they are looking for, so they are concentrating on finding a particular answer. Scanning involves moving eyes quickly down the page seeking specific words and phrases. The spread of the Internet may well accelerate the need for this type of reading. When scanning, some strategies can be useful such as author's use of organizers such as numbers, letters, steps, or the words *first*, *second*, or *next*. Bold faced letters, italics, or in a different font size, style, or colour can be useful.

4.3.3. Extensive reading

Extensive reading can be defined as reading a large quantity of text, where reading confidence and reading fluency are prioritized. Extensive reading is used to obtain a general understanding of a subject and includes reading longer texts for pleasure with emphasis on overall meaning. Examples of extensive reading are: a novel one reads before going to bed or magazine articles. However, in reality, extensive reading is not promoted in many schools in Algeria.

4.3.4. Intensive reading

Intensive reading is used on shorter texts in order to extract specific information. It includes very close accurate reading for detailed information. Intensive reading is used to grasp the details of a specific situation. In this case, it is important that the reader understands each word, number or fact as it may be the case for a contract. Intensive reading is the approach most often adopted in the reading classes in Algeria. This type of reading, which seeks to grasp the whole message, including both arguments and supporting details, encourages careful, literal processing of text.

Anderson (1999), while highlighting the superiority of extensive reading, still emphasizes the importance of intensive reading through which readers can develop strategies and skills which they can transfer to extensive reading contexts. Nutall (1996) also states the importance of teaching how meaning is produced through intensive reading which is intended to train readers to cope with the texts.

4.3.5. Careful reading

Urquhart and Weir (1998) point out that careful reading is associated with reading to learn. The reader attempts to handle detailed information in the text, to master content including details, evaluate material, outline, summarize, paraphrase, analyze, solve problems, memorize, evaluate literary value or read poetry. Thus, reading rate seems to be rather slower than other types of reading because in this type of reading, readers often require rereading and inferencing to connect information with background knowledge.

4.3.6. Browsing

Browsing is the sort of reading where readers do not have any particular goals for reading and parts of a text may be skipped fairly randomly and there is little need to integrate the information. We often browse magazines or newspapers just for fun. In the classroom, as there are limited resources written in English, students have few opportunities to make use of browsing.

4.3.7. Reading for general comprehension

According to Grabe and Stoller (2002, p.14) this is the most basic purpose of reading though it is actually more complex than commonly assumed. Because reading for general comprehension ‘requires rapid and automatic processing of words, strong skills in inferring a general meaning representation of main ideas, and efficient coordination of many processes under very limited time constraints’.

Other authors define reading as the act of simultaneously ‘reading the lines, reading between the lines, and reading beyond the lines’ (Alderson, 2000, p. 9).

Reading the lines refers to the act of decoding the words in order to construct the author's basic message; reading between the lines, refers to the act of making inferences and understanding the author's implied message; and finally, reading beyond the lines involves the judging of the significance of the author's message and applying it to other areas of background and knowledge.

4.4. Levels of Comprehension

First of all, it should be mentioned that no distinction is made throughout this project between the terms *understand* and *comprehend*. *Webster's New World College Dictionary* (2004) clarifies this:

Understand and *comprehend* are used interchangeably to imply clear perception of the meaning of something, but, more precisely, *understand* stresses the full awareness or knowledge arrived at, and *comprehend*, the process of grasping something mentally (a foreigner may comprehend the words in an American idiom without understanding at all what is meant). (p. 1586)

Levels of understanding (or comprehension) refer to the different depths of comprehension and different analysis of what is meant. More precisely, the term 'levels of comprehension' refers to the thinking processes that are stimulated in order to arrive at answers to reading comprehension questions. Bloom's Taxonomy of Comprehension (1956) has been used by educators for many years as a way of ensuring that all levels of

thinking occur when printed text is read. Each level builds upon the previous one and ensures that a comprehensive thought process occurs. Specifically, the six levels of Bloom's Taxonomy of Comprehension are: knowledge, comprehension, application, analysis, synthesis and evaluation.

Knowledge level: It is the foundational level of comprehension that emphasizes a literal interpretation. The *who, what, where, when, and how* are identified.

Comprehension Level: It is the basic level of understanding. It involves the ability to know what is being communicated in order to make use of the information.

Application Level: It is the ability to use a learned skill in a new situation. That is, transferring an understanding of the events to other areas.

Analysis Level: It is the ability to break content into components in order to identify parts, see relationships among them, and recognize organizational principles. Comparison and contrast can be utilized to better understand the relationships between the different elements involved in the incidents.

Synthesis Level: It is the ability to combine existing elements in order to create something original. Comprehension at this level includes putting the events into a broader context.

Evaluation Level: It is the ability to make a judgment about the value of something by using a standard.

To facilitate the teaching of reading comprehension, Barrett (1982, cited in Al-Jarf, 2007) proposed taxonomy of reading comprehension levels. Worthy of note are six categories of comprehension outcomes:

Literal comprehension: Literal comprehension refers to an understanding of the straightforward meaning of the text i.e. the literal recognition, recall or verification of facts, vocabulary, dates, times, locations, comparisons, cause-effect relationships and character traits. These constitute the basic or surface meaning of the text. Literal comprehension entailing recognition and recall of ideas and information explicitly stated in the reading selection. Questions that require literal comprehension include items that can be answered directly from the text at the sentence level. They can be questions that require either copying or rephrasing of information from the text. At this level the reader or student can attempt to answer the question: What did the author say?

Some specific reading skills at the literal level of comprehension are: identifying specific information or noting details, sequencing ideas when explicitly signals are given, and following instructions. These skills, specially the first two, are scanning skills

Reorganization: Reorganization deals with the organizing of ideas and information explicitly stated. It is based on a literal understanding of the text; students must use information from various parts of the text and combine them for additional understanding. For example, we might read at the beginning of a text that a person was born in a certain year and then later at the end of the text that she died in certain year. In

order to answer this question, ‘How old was the person when he/she died?’ the student has to put together two pieces of information that are from different parts of the text.

Inferential comprehension: Inferential comprehension (also known as interpretive level) involves more than a literal understanding. Making inferences involves more than a literal understanding. This involves students combining their literal understanding of the text with their own knowledge and intuitions. Answers are based on material that is in the text but not explicitly stated. Inferential Comprehension requires the use of ideas and information, explicitly stated, as the basis for making intelligent ideas and using this to make intelligent guesses and/or hypotheses. At this level the reader or student can attempt to answer the question: What was meant by what was said. Some reading skills at the interpretative level of comprehension are: identifying the main idea, inferring character traits, forming conclusions, anticipating or predicting an action and drawing implications.

Inferences can be categorized as implications, conclusions, generalizations or predictions.

Implication is any inference or expectancy that maybe logically implied or understood, but not directly stated, form the author’s arguments in a text or utterance.

Conclusion draws together factual evidence into a statement about the nature of a phenomenon.

Generalization is a statement about the behaviour of a large population based on the observable behaviour of a similar but smaller sample group.

Prediction is a statement about future behaviour or action.

Evaluation: (also known as the critical level), refers to judgments of reality or fantasy, fact or opinion, adequacy or validity, appropriateness, worth, desirability and acceptability. It also refers to judging the language and effect of the text in the light of appropriate criteria. The principal focus of this level of comprehension is the aspect of evaluation. Critical level of comprehension is the judgment of validity or worth of what is read or heard, based on sound criteria of standards developed through previous experiences. Critical evaluation occurs only after our students have understood the ideas and information that the writer has presented. At this level, students can be tested on the following skills: the ability to differentiate between facts and opinions; the ability to recognize persuasive statements; and the ability to judge the accuracy of the information given in the text.

Integration: Integration refers to an understanding of the explicit meaning of a text but also requires accessing information from various parts of the text in order to answer a given question. Questions that require integration include items that require thinking about how ideas or information in the passage relate to each other as well as to an understanding of its main idea and supporting details.

Appreciation: Appreciation refers to emotional responses to the content, plot or theme, sensitivity to various literary genres, identification with characters and incidents, reaction to the author's use of language, and response to generated images.

Alderson (2000) refers to the literal understanding of a text, understanding of meanings not directly stated in the text, or an understanding of the main implications of a text. There is also the distinction between understanding details and the main idea of a text as well as reading the lines, reading between the lines and reading beyond the lines (the critical evaluation of a text). These distinctions refer to the product of reading and enable us to describe and evaluate the differences in understanding among readers. This is because, Alderson argues, it is believed that inferred meanings are somehow ‘deeper’ than literal meanings, that it is more difficult to reach a critical understanding of a text than it is to infer meanings and that both are more difficult than just understanding the literal meaning. This hierarchy of difficulty and value has led to the assumption that the same hierarchy applies to acquisition, but the empirical justification for these assumptions is slim and the picture is more complex (Alderson, 2000, p. 8).

In fact, it is possible to comprehend the words, but not the meaning of a sentence, and sentences but not the organization of a text. Researchers relate the former to ‘microprocesses’ or local, phrase by phrase understanding and the latter to ‘macroprocesses’ or global understanding (van Dijk, 1978 cited in Alderson, 2000, p. 9).

4.5. Components of reading ability: skills involved

Reading skills can be described roughly as ‘a cognitive ability which a person is able to use when interacting with written text (Urquhart and Weir, 1998, p. 88).

However, since there are a number of skills taxonomies, it can be difficult to grasp the

whole picture of reading skills (see, Urquhart and Weir, 1998, pp. 90-91; Brown, 2001, p. 307). In fact, despite the fact that psychologists and educators have been conducting research on various aspects of reading for more than a century (e.g., Thorndike, 1917a, b, c; Dewey, 1935; Robinson, 1963; ; Smith, 1971; Singer and Ruddell, 1976; Johnston, 1983, Alderson and Urquhart, 1984; Alderson, 1990a, b, 2000), there are still controversies surrounding the exact nature of skill, or skills, that are involved in reading comprehension, either in L1 or L2/FL.

The analysis of the studies that have been conducted appears to be complicated by the fact that usually many different dimensions are found and similar dimensions are often named differently, as well as different dimensions being named alike. In general, however, studies that have addressed the nature of reading skill(s) seem to have subscribed to one of the two views: a) reading is a unitary, holistic, and indivisible skill which cannot be split into different sub-skills (e.g., Alavi, 2002; Alderson, 1990 a, b; Andrich and Godfrey, 1979; Lunzer et al., 1979); and b) reading skill consists of various sub-skills (e.g., Bloom, 1956; Davis, 1968; Munby, 1979).

Hughes (1989) refers to ‘macro skills’ and ‘micro skills’ of reading comprehension. The distinction between these two levels of sub-skills is not made explicit, but it appears that the term ‘macro skills’ refers to understanding the general ideas in the text (e.g., information, gist, argument) while ‘micro skills’ refers to recognizing and interpreting the linguistic features of the text (e.g., referents, word meanings, discourse indicators) Ghahraki and Sharifian (2005).

However, it should be pointed out that some researchers questioned the very notion of the so-called reading comprehension skills. They advanced that such skills do not exist. In this regard, Rosenshine (1980) more states that ‘there is simply no clear evidence to support the naming of discrete skills in reading comprehension’ (p. 552). Alderson and Urquhart (1984) repeated this, and Alderson (1988) stated it. Gardner (1978) suggested that these ‘skills’ are better regarded as activities involving comprehension than as categories of abilities’ (p. 72). In other words, when reading, we engage in activities such as recalling word meanings, inferring, drawing conclusions, and so on, but these are all aspects of the act of comprehending (i.e., reading) which will be considered later.

The present research will adhere to Grabe's (1991) taxonomy shown below because he uses general categories and these skills and knowledge are obviously important to teach reading in the Algerian educational context.

- a. Automatic recognition skills
- b. Vocabulary and structural knowledge
- c. Formal discourse structure knowledge
- d. Content/ world background knowledge
- e- Synthesis and evaluation skills / strategies
- f. Metacognitive knowledge and skills monitoring.

4.5.1. Automatic Word Recognition Skills

Automatic recognition skills allow readers to identify letters and words without being consciously aware of the process. They imply that the vast majority of words in a text will be recognized speedily directly and, on the whole, without recourse to phonological decoding. Good readers are able to read rapidly because they can recognize most words automatically, and therefore process this information very efficiently. The combination of rapid and precise word recognition has proven to be an effective predictor of reading ability, particularly of young readers.

4.5.2. Vocabulary and structural (grammar) knowledge

Second/foreign language researchers broadly agree that ‘vocabulary development is a critical component of reading comprehension’ (Grabe, 1991, p.392). Laufer (1997) emphasizes that reading comprehension is strongly related to vocabulary knowledge, more strongly than to the other components of reading and points out that even if students have good metacognitive strategies or background knowledge of the world, these will not be of much help in L2 before a solid language base has been reached. Vocabulary knowledge is most frequently used as an indicator of the linguistic threshold.

For fluent L1 readers of English, a recognition vocabulary of 40,000 words are estimated to be sufficient for academic needs; but a recognition vocabulary of 5,000 to 7,000 words are estimated to be adequate for EFL/ESL academic coursework (Barnett, 1989; Grabe, 1991; Strother and Ulijn, 1987). On the basis of experimental data, Laufer

(1992c, p. 100) suggests that ‘the knowledge of 3000 word families is the lexical threshold of reading comprehension in L2 [English]’. More details are provided in Chapter 6.

Structural knowledge is also important. Grabe and Stoller (2002) highlighted the importance of grammar in L2 reading context, especially through hours of exposure to print in order to develop automaticity in using information from grammatical structures to facilitate reading. In the Algerian secondary educational context, explicit grammar instructions are essential for successful reading because of time constraints. On the other hand, teachers should provide students with more opportunities to encounter English passages to enhance the automaticity of their reading.

4.5.3. Knowledge of discourse structure

Research has shown that Knowledge of the structure of formal discourse / rhetorical organizational patterns (i.e. formal schemata) is very useful in reading comprehension. It is important for students to notice some basic text conventions in English texts. To see how ideas are typically arranged can help students to read a text more easily. According to Geva (1983) and Carell (1984a, b), there is empirical evidence that readers' background knowledge of text structure and discourse cues significantly affects their reading in a second or foreign language. Thus, explicit teaching of rhetorical organization of text facilitated EFL/ESL students' reading comprehension and assists them in understanding and remembering the text. To make students more aware of discourse structure, explicit teaching of discourse markers and

several basic discourse structures are necessary. This knowledge is useful both in skimming and scanning.

4.5.4. Content/World Background Knowledge

As discussed in the previous chapter, background knowledge includes all the experience that a reader brings to a text. Anderson (1999) points out that activation of prior knowledge facilitates comprehension. Activating the reader's knowledge of the subject matter and cultural content of the text is a significant factor in both reading comprehension and recall. Research shows that L2 learners can better recall information from text on topics familiar to them than readings of equivalent difficulty level on subjects with which they are less familiar. Moreover, readers can more easily comprehend and recall texts of which the content is based on their own culture than texts based on unfamiliar and more remote cultures.

4.5.5. Synthesis and Evaluation Skills

Fluent readers are able not only to comprehend the text, but to make judgments about the information, the author's purpose, and the usefulness of the text. They usually use strategies like predicting to assist them in anticipating text development and evaluating the author's perspective as they read.

4.5.6. Metacognitive Knowledge and Skills

Metacognitive knowledge is knowledge about how learners think and self-regulate their cognitive processes. It includes knowledge about language and ability to recognize structural and rhetorical features of the text using suitable strategies for

accomplishing particular goals. In the reading process, metacognitive skills include recognizing main ideas, adjusting reading rate for skimming, scanning, paraphrasing, and summarizing, guessing meaning from word formation rules, prefixes, and suffixes, and taking notes. The ability to effectively apply metacognitive strategies to the reading process is a key ingredient to skilled reading, especially for older readers.

4.6. Relationship between L1 and L2 reading comprehension

A crucial issue for investigators in the field of second/foreign language reading is the extent to which L1 and L2/FL reading differ. First, L2/FL readers do not have the same language resources as L1 readers at the outset of learning (Alderson, 2000; Grabe and Stoller, 2002, Urquhart and Weir, 1998). Second, L2/FL readers have much lower levels of lexical, grammatical, and discourse knowledge at beginning stages of L2/FL reading than L1 readers do when they begin to read. In addition, L2 readers have much less overall exposure to L2 print (Day and Bamford, 1998); in contrast, L1 readers are consistently exposed to native language print from a very early age. L2/FL readers also vary considerably in their own L1 reading abilities, creating an added complexity (Grabe, 2002). In addition to linguistic differences, L2/FL readers do not have the same social and cultural knowledge as L1 readers. Neither do they have the same motivation for reading.

Research of the 1970's assumed that L1 reading ability transferred to L2. This view is known as the 'Reading Universals Hypothesis' (Goodman, 1971, pp. 140-42). For a long time it was assumed that second/foreign language reading is marginal and

derivative from first language reading and that result of research on L1 reading will apply in L2 contexts. Bernhardt (1991) summed up the situation by asserting that many L1 reading researchers tended to see L2/FL studies ‘...as ‘marginal and derivative’ because they viewed L2 reading itself as ‘merely a slower, bastardized version of doing the same task in the native language’ (p. 2). Within such a view, reading in a second language was then often viewed as merely a slower version of doing the same task in the native language. This implies that second language tasks are mapping tasks – that is replacing one mode of behaviour with another. Then, from the late 1970s, Clarke (1979; 1980) and others argued that good L1 readers who were weak in L2/FL suffered a ‘short circuit’ that reverted them to poor reading strategies.

In the context of the ongoing debate, two competing hypotheses have been proposed to describe differences between L1 and L2 reading. Broadly framed within the schema theoretic model, the two views attempt to account for the performance of novice L2 readers who are already skilled in L1 reading. The first, Clarke's ‘short circuit hypothesis,’ now referred to as the Linguistic Threshold Hypothesis (LTH), states: in order to read in a second language, a level of second language linguistic ability must first be achieved. The Threshold Hypothesis presupposes then a language threshold in L2 proficiency below which transfer of reading strategies acquired in L1 over to L2 reading is ‘short-circuited’. In other words, language is the key factor in reading. The second hypothesis is referred to in the recent literature as the Linguistic Interdependence Hypothesis (LIH), stating that: reading performance in a second

language is largely shared with reading ability in the first language. The idea is that once the ability to read has been acquired, it can be transferred to a second language. Alderson (1984) presented Clarke's Short Circuit Hypothesis in the form of a question, which became well-known later in L2/FL reading research literature, that is, 'Reading in a foreign language: a reading problem or a language problem?' (p.1)

4.6.1. Reading problem or language problem

As early as 1976, Cummins (1976) hypothesized that there might be a threshold level of L2 competence which pupils must attain both in order to avoid cognitive disadvantages and allow the potentially beneficial aspects of becoming bilingual to influence their cognitive functioning. He stated: 'Those aspects of bilingualism which might accelerate cognitive growth seem unlikely to come into effect until the child has attained a certain minimum or threshold level of competence in his second language' (p.23). In his discussion of this issue, Alderson (1984) noticed Clarke's notion of 'language competence ceiling', which hampered the good L1 readers in their attempt to use effective reading behaviours in L2 reading, and related it with Cummins's (1976) 'threshold level of linguistic competence'. Alderson (1984) summarized the notion of L2 threshold level as 'foreign language readers will not be able to read as well in the foreign language as in their first language until they have reached a threshold level of competence in that foreign language' (p.19). Alderson (id.) concluded tentatively that L2/FL reading 'appears to be both a language problem and a reading problem, but with firmer evidence that it [i.e., L2 reading] is a language problem, for low levels of foreign

language competence, than a reading problem' (p.24). However, he points out that further methodological measures need to be taken; particularly providing the missing information of controlling the same individuals reading ability in L1 and L2, and the L2 or FL proficiency. Taking this missing information into account, Carrell (1991) investigates the same question starting with the premises of an equation stipulating that:

$$\text{L2 Reading} = \text{L1 Reading} + \text{L2 Language Proficiency}$$

The researcher has confirmed her equation as she writes that '... both first language reading *ability* and second language *proficiency* have significant effects on second language reading ability' (ibid., p 167, italics original). She found that for one of her investigated groups (speakers of English as an L1 who were reading in Spanish as their FL), the language proficiency accounted for a greater proportion of the variance in second language reading. Taillefer (1996, reported in Ganderton, 1998) supported the same equation advocated by Carrell (ibid.) and concluded that for L2 readers of lower language proficiency, other factors outside of the simple aforementioned equation may be significant. Most notably affective factors, such as lack of confidence in L2 or anxiety, may make them afraid of incorporating their L1 reading knowledge into the L2 reading situation. Despite these factors, the researcher points out that the threshold hypothesis and the significance of the two factors of L1 reading ability and L2 language proficiency do hold some importance in characterizing the L2 reading process.

4.6.2. The Linguistic Threshold Hypothesis

As already mentioned, the issue regarding the relations among L2 reader's L1 reading ability, L2 proficiency and L2 reading comprehension was first brought to the attention of the second language reading research community in Clarke's (1979) 'The Short Circuit Hypothesis of ESL reading' (p. 19).

In his experiments, Clarke (1980) tried to address two questions: '1) Can the psycho-linguistic perspective of reading explain the reading performance of proficient adult Spanish-speaking readers, reading in Spanish and English? 2) Do proficient L1 readers transfer their reading skills to the second language?' (p.204). In his experiment, the subjects were compared for their reading ability in their L1 and L2 using a cloze procedure and a miscue analysis. The underlying assumption was that, if reading was the same in all languages, given equivalent proficiency in L2, the superior reading skills of the good readers would provide them with an equal advantage over the poor readers in both L1 and L2. The research data collected in the experiment showed, however, that good L1 readers did not demonstrate the expected advantage over poor L1 readers when confronted with difficult questions, although as a group they were better L2 readers than the poor L1 readers. Clarke speculated on the basis of his experimental data that low level L2 readers who were good L1 readers showed a reduced ability to utilize the good reader strategy. They were not able to rely more on semantic than on syntactic cues when reading in L2, and thus demonstrated a reduced superiority over poor L1

readers when both read L2 materials. The short circuit hypothesis has since found considerable support in other studies. Laufer's study (1992c) has led her to suggest that:

If language teachers and course designers are concerned with improving the L2 reading of low proficiency learners, it is advisable not to bother about their cognitive ability and reading skills in L1. A more serious pitfall for such learners is insufficient knowledge of vocabulary. (p. 101)

According to Laufer, L2 readers cannot transfer their L1 reading skills to their L2 reading task if they have not achieved L2 proficiency as reflected in their sufficient L2 vocabulary. Ulijn and Salager-Meyer (1998) also note in summary of the relationship between 'L2 lexical threshold' and 'L1 skill transfer' that:

... helping students to develop text problem solving skills (using context to figure out interpretations, intentions, conclusions, etc.) is a good idea, but helping them to improve their word identification skill is an even more important goal in setting the stage for the successful use of such comprehension strategies and in paving the way toward independent and fluent reading. (p.83)

Several other studies (cited in Bernhardt and Kamil, 1995) have explicitly approached the question concerning the relationship between L1 and L2 reading and have provided direct evidence of the Linguistic Threshold Hypothesis. In Cziko's studies

(1978, 1980), native French speaking and French-as-an-L2 students in junior high school used both contextual and graphic clues in reading French. The students who were less competent in French tended to rely more heavily on lower-level textual clues than higher-level contextual clues. It appeared that only with increasing proficiency in the target language were subjects able to use discourse constraints, which were held to be an important ability in the reading process in general. Roller (1988) reported L2 vocabulary development in 300 Shona-speaking children between third and fifth grades and concluded that ‘transfer between languages is minimal and certainly is not occurring early and readily’ (p.315).

With more accuracy, the same area has been investigated elsewhere showing how much each of the two variables accounts for in the variance. Bernhardt and Kamil (1995) have found that L1 reading ability accounts for between 10 and 16 per cent in second language reading and language proficiency for 30 to 38 per cent. They believe that ‘while language proficiency accounts for a greater proportion of the variance, first language reading also makes a significant contribution’ (p. 25). However, these figures do not compare well with those of Lee and Schallert (1997). These looked into the threshold level in 809 Korean students who exhibited a wide range of ability both in reading Korean and English and in their English proficiency. The authors contend that language proficiency plays a more critical role in reading in L2 than does L1 reading ability. They found that L2 proficiency accounts for 56% of the variance in L2 reading, and only 30% by L1 reading. This difference, the authors claim, ‘supports the contention that L2

reading ability owes more to L2 proficiency than to L1 reading' (p. 732). This finding goes in line with Carrell's (1991) in which she claims that 'proficiency in the foreign language accounted for a greater proportion of the variance in L2 reading than did reading ability in the first language' (p. 168). Furthermore, Lee and Schallert (op. cit.) that L2 proficiency is a more significant predictor in L2 reading than L1 reading ability. Also, the high correlations between L1 and L2 reading abilities in students with high L2 proficiency and the low ones in those with low L2 proficiency suggest the existence of a threshold level of language proficiency.

The studies reported above provide enough evidence to claim the existence of a threshold level which not only paves the way for the transfer of L1 reading ability into SL/FL reading but also predicts success in reading in the target language. The existence of such a threshold is believed to enable -or else hinder- the transfer of L1 reading ability in SL/FL reading.

However, The Short Circuit Hypothesis advanced by Clarke was challenged by the Linguistic interdependence Hypothesis.

4.6.3. The Linguistic Interdependence Hypothesis

The Short Circuit Hypothesis advanced by Clarke was challenged by the hypothesis which advocates language interdependence and reading universals. According to Jolly (reported in Alderson 1984), success in L2 reading depends crucially on one's L1 reading ability rather than on one's level of the target L2, because L2 reading requires the

transference of old skills, not the learning of new ones. This view was shared by Coady (1979), who found that many L2 learners had very poor reading habits to transfer from their L1 and that, in many cases, they learn in L2 the reading skills that ‘should have been learned in first language instruction’ (p.12). Such a hypothesis about the interdependence between the readers’ L1 and L2 reading ability found a ready echo in Goodman's ‘reading universals hypothesis’ (Goodman, 1973, 1988). According to Goodman (1973), ‘the reading process will be much the same for all languages’ (p.27). According to Carrell (1991), some researchers in bilingual education also maintained that:

Reading or learning to read is accomplished only once, and that once learners have matured in their ability to read in the first language, the awareness of the reading process transfers to the second language and does not need to be relearned. (p.159)

Some other studies, however, found evidence for the Linguistic Interdependence Hypothesis. Lapkin and Swain (1977) found no difference between bilinguals’ reading ability and control native speakers, for both L1 and L2, nor did they discover any ‘glaring’ differences between native and non-native reading performance. Cummins (1980) found in his experiments that L1 and L2 were interdependent in terms of cognitive academic language proficiency and that older learners of L2 acquired L2 more rapidly than younger learners because their L1 cognitive academic language proficiency was

better developed. Pritchard (1990) also found that the subjects used the same metacognitive strategies in ESL reading as in their own Spanish reading.

It appears then that the Linguistic Threshold Hypothesis and the Linguistic Interdependence Hypothesis have appeared to be able to find supportive evidence in research. A study carried out by Carrell (1991) might well cast some light on the indeterminacy of the issue. She conducted an experiment with two different groups of subjects. One group consisted of 45 native speakers of Spanish studying English as a second language; and the other group comprised 75 native speakers of English studying Spanish as a second language. Each group read an easy and a difficult passage and completed two multiple-choice tests. Carrell (1991) found in the statistics of her studies that 'both first language reading ability and second language proficiency have significant effects on second language reading ability' (p.167) and that the two variables: L1 reading ability and L2 proficiency level could account for up to 39 per cent variance in L2 reading ability. The most interesting finding, however, was about 'the relative importance' of each of the factors for each of the two groups of subjects. For the group with Spanish as their native language and English as their second language, reading ability in L1 accounted for a greater proportion of the variance in L2 reading ability than did proficiency in the L2 language. But for the group with English as their native language and Spanish as their foreign language, proficiency in L2 accounted for a greater proportion of the variance in L2 reading ability than did reading ability in L1. One

possible explanation for the difference of the relative importance of L1 reading ability and L2 language proficiency to L2 reading ability may be that, as Carrell (1991) put it:

... proficiency level in the second language is more critical for learners at slightly lower proficiency levels (as would be true for the English L1 group in this study, who were determined to be at Levels 2, 3, and 4) when compared to learners at slightly higher levels (as would be true for the Spanish L1 group in this study, who were determined to be at Levels 3, 4, and 5). This last potential explanation, would, of course, fit with the views of the 'language threshold' researchers... (p.168)

In other words, if one's proficiency level in L2 is low, one may not be able to access L1 strategies (for whatever reasons) but if one's proficiency is high then L1 strategies are available for L2 reading comprehension. Similarly, Royer and Carlo (1991) found in their studies that reading skills in L1 do get transferred to reading in L2 as their second language proficiency develops. Bossers (1991) conducted a similar experiment to investigate the relation between L1 reading, L2 reading and L2 knowledge. In his experiment, 50 native speakers of Turkish in Dutch post-secondary education read 4 passages both in L1 and L2 and answered 16 multiple-choice questions per text. Bossers found that 'L1 reading was more implicated in the reading of skilled L2 readers than in the reading of less skilled L2 readers' (p.57), and that 'the results [of the experiment] point in the direction of a language threshold' (p.59).

The indeterminacy of L1 reading ability and L2 language proficiency as a predictor of the reader's L2 reading ability was given further support in the experiment conducted by Bernhardt and Kamil (1995). In the experiment, they used students in three levels of Spanish institutions at the United States Air Force Academy: beginning, intermediate and advanced. The subjects took three reading comprehension tests: the test in Spanish, the English translation of the Spanish reading comprehension test and a more difficult English reading comprehension test. Their findings seemed equally as inconclusive as Carrell's. 'The question framing this paper, 'Is second language reading a language problem or a reading problem?' is not unambiguously answerable: the answer is 'yes' to both sides of the question; and the answer is also 'no' to both sides of the question' (Bernhardt and Kamil, 1995, p.32). So far as L2 reading comprehension is concerned, the question that better formulates the two hypotheses seems not to be the one proposed by Alderson but the two offered by Bernhardt and Kamil (1995). They represented the two seemingly opposite hypotheses as 'How first language (L1) literate does a second language reader have to be in order to make the second language (L2) work?' and 'How much second language (L2) knowledge does a second language reader have to have to make the first language (L1) reading knowledge work?' (p.32)

Indeterminate as the two questions may appear, the presentation of the two hypotheses has presumed the existence of the thresholds of L1 and L2. It implied that L2 readers cannot transfer their L1 reading ability to their L2 reading if their L1 has not reached its threshold, nor can their L1 reading ability transfer to their L2 reading if their L2 has not

reached the threshold. The solution to the two questions, therefore, lies in the identification of the thresholds of both L1 and L2. One's L1 reading ability can transfer to L2 reading only if one's L1 and L2 have each reached their respective thresholds.

As Devine (1988) pointed out, despite the indeterminacy of the relative importance of L1 reading ability and L2 language proficiency in relation to L2 reading comprehension, there is little doubt about the importance of L2 language proficiency in L2 reading comprehension, especially with the readers of low L2 proficiency. 'The critical interaction of language proficiency and reading ability is now generally well accepted. Most researchers and classroom teachers, in fact, take as a given that second language readers must reach a certain level of second language competence before they can effectively read in the second language' (p.261). She continued to note that most L2 reading instructors and researchers would agree that L2 readers must attain a level of proficiency in L2 before there can be genuine interaction with texts in L2. According to Ulijn and Salager-Meyer (1998), it is taken as a given that 'second language readers must reach a certain level of second language competence before they can effectively read in the second language' (p.83). What seems to be of greater concern is not the existence of the threshold level of the L2 but 'to what extent does language proficiency place any limit on reading performance in a second language?' (Barnitz, 1985, p.11) Research on L2 reading comprehension has so far resulted in some basic understanding of the notion of Linguistic Threshold.

First, linguistic threshold appears to be fluid not absolute. The general findings of research show that low reading achievement in L2 is more often than not related to low L2 proficiency and those readers with low L2 proficiency are especially handicapped in their ability to utilize the knowledge learned through L1 reading, particularly the knowledge about contextual constraints and cohesive devices (Devine, 1988). The findings of research (discussed above) suggest that L2 readers will not be able to read effectively unless they have developed some L2 proficiency. Despite what the term may literally suggest, Linguistic Threshold is said not to be understood as an absolute term. It varies from reader to reader and from task to task. According to Alderson (1984), the concept of the linguistic threshold for successful L2 reading depends on the answers to the following questions:

... to what extent is it [the linguistic threshold] syntactic, semantic, conceptual, discursal? Does the level of the threshold vary for different learners, and for different tasks? Is it conceivable that good first-language readers will require a lower threshold before being in a position to utilize their good reading strategies? Will the attainment of a higher level of competence compensate for a poor first-language reader? (p.21)

Since these questions are precisely the questions that considerable research of L2 reading comprehension is still working at, no consensus has emerged on the issue. But studies have shown that the more difficult the reading task, the higher the linguistic

threshold is likely to be (Alderson, 1984), and the more one knows about the general subject of a text, the lower the linguistic threshold is likely to be (Hudson, 1988).

In summary, the linguistic threshold hypothesis and the linguistic interdependence hypothesis concern two important aspects of L2 reading comprehension. It is still not possible to decide precisely the extent to which, and the stage of L2 learning when, L2 reading is determined by L2 general proficiency. Neither is it possible to specify the extent to which the existing L1 reading ability may determine the success in L2 reading comprehension. Just as Ulijn and Salager-Meyer (1998) point out, a well balanced 'blend' of the two viewpoints is generally adopted today by most researchers and classroom teachers, acknowledging a balanced and critical interaction between L2 proficiency on the one hand, and reasoning processes and reading ability on the other, and taking as given the notion of linguistic threshold.

4.7. Factors affecting foreign second/foreign language comprehension

There is little dispute among researchers that the reader plays the central role in an act of reading. While in early approaches to language learning the reader was believed to be a passive receiver of information, he/she is now considered an active participant in a reading activity. As stressed several times throughout this study, reading has come to be considered as an interactive process between text and reader making the ultimate product the result of interaction between the reader and other variables. Based mainly on the research and arguments put forth by Goodman (1968; 1970; 1973) and Smith

(1971; 1973), it is now accepted unanimously that the reader contributes more than ‘the visual symbols on the page’ do (Grabe, 1991, p. 377). The reader’s contribution to text comprehension differs from reader to reader because readers are different in their ‘shared knowledge, language skills, strategies ... and ‘other personal characteristics’ ’ (Alderson, 2000, p. 128).

It has also been noted earlier with Hosenfeld (1979) that the level of reader comprehension of the text is determined by how well the reader variables interact with the text variables (text type, structure, syntax, and vocabulary) . It is believed then that a plethora of factors affect the reading process to a considerable extent. In Chapter 3 (see under 3.4.), it was noted that it is generally accepted that readability indices and factors influencing text difficulty are of two sorts: some of them are directly related to the text, others are directly related to the reader.

Moreover, it was also noted in the introduction to this study that broadly speaking, factors affecting reading comprehension, as listed by Alderson (2000), can be classified into two general categories: reader variables, and text variables. The first category of variables includes factors internal to readers such as reader's background knowledge, reader cognitive abilities, and reader's motivation and attitude. The second category of variables includes factors such as text content, text type and genre, text organization, text typographical features, and text readability which are internal to texts rather than to readers. Two other variables may affect reading comprehension, namely, the context, and the writer. Context variable refers to all situational elements such as the

time of reading and the place of reading, as well as the larger socio-economic context. Writer variable refers to the text-producer. Each of the aforementioned variables is discussed in some detail in the following.

4.7.1. Reader variables

4.7.1.1. Cognitive abilities

One reader characteristic that influences reading comprehension is one's cognitive abilities. Reading performance is dependent upon cognitive abilities to the extent that some scholars relate reading to thinking. Goodman (1970, p. 108), for example, notes that efficient reading results from 'an interaction between thought and language.' Others have emphasized readers' cognitive strategies and meta-cognitive knowledge in affecting reading comprehension (Baker & Brown, 1984, p. 34; Aslanian, 1985; Johnson, 1998; Schoonen et al., 1998). The point is that the reader's ability and success in relating new information to old information seem to be vital for understanding. The reader's (cognitive) ability to link all sources of information to one another in order to construct the writer-intended meaning has been recognized by Trabasso (1981), Linderholm, Everson, van den Broek, Mischinski, Crittenden, & Samuels, (2000), and Johnson (1998). The latter, for instance, assert that 'Individual differences in working memory capacity have been demonstrated to be good predictor of variance in both overall reading ability ... and specific reading skills (p. 244).' Similarly, Rumelhart (1984, p. 19) compares reading comprehension to 'a detective act' where the reader has to use his/her cognitive ability to connect all the relevant

information in the situation to solve the problem. This implies that one's degree of comprehension is not only affected by his/her background, cultural, and language knowledge but also by his/her cognitive abilities including intelligence. The resulting comprehension is the product of the interaction among these and other factors.

4.7.1.2. Affect

Another aspect of the reader is his/her affective state. The affective state refers, on the one hand, to a reader's purposes, perspectives, motivation, emotional mood, etc., prior to reading. On the other hand, it refers to the affective responses brought about by the text while reading. A reader may read the same text at different times for different purposes, and with different kinds of attention paid to information on the page, and therefore, with different degrees of comprehension. The influence of one's goals, views, and emotional states on reading comprehension has been documented by Garner (1987), Lorch and van den Broek (1997), Pressley (1998), and Alderson (2000, p. 80).

4.7.1.3. Culture/Background knowledge

The effect of culture and background knowledge (or prior knowledge) has been investigated extensively both in native-language and second/foreign language. Some even define comprehension as the 'use of prior knowledge to create new knowledge' (Adams & Bruce, 1982, p. 23). Prior knowledge, it has been claimed, is one of the most powerful influences on text comprehension (Adams & Bruce, 1982;

Alexander & Jettson, 2000; Beck & McKeown, 1999; Gordon, 1992; Pearson & Stephens, 1994). Carrel and Eisterhold (1983) consider language background knowledge an important factor in comprehending a text; they express this importance as follows: 'Efficient comprehension requires the ability to relate the textual material to one's own knowledge. Comprehending words, sentences, and entire texts involves more than just relying on one's linguistic knowledge' (p. 12). Further, these researchers talk of two types of background knowledge: formal and informal. Formal knowledge refers to the reader's knowledge of the rhetorical organizational structures of different types of texts; content knowledge refers to the content area of a text. They also believe that reader's failure to provide the proper formal and, particularly, content knowledge (schema) would result in various degrees of non-comprehension.

Prior knowledge can come from informal sources, like family, the home environment, or peers; or more formal sources like school, where subject matter, domain and topic knowledge are learned (Alexander, Kulikowich, & Schulze, 1994).

It is the author's initial responsibility for comprehensibility of text. The match or mismatch between what a writer presumes the intended readers know and that reader's background knowledge about the topic affect to a great extent comprehension. Reading comprehension is closely dependent on the quality of fit between reader and text. Consider for example the case a reading passage intended for secondary school students in an Algerian school on astronomy, a topic that is inexistent in any of the syllabuses throughout all their schooling. It is very likely that comprehension will not be easy.

Within the same line of thought, some texts have been reported to lead to incomprehension or misunderstanding because they do not provide the reader with effective bottom-up clues to activate the related schemata (Rumelhart 1981, 1984).

Studies in first language comparing experts with novices have made it clear that people with high domain knowledge comprehend a text better than those who lack that knowledge. Adams, Bell, and Perfetti (1995) investigated the relation between children's knowledge of a specific domain (football) and reading skill in text comprehension. They concluded that reading skill and domain knowledge make complementary contributions to reading comprehension and reading speed. High-skilled readers with little domain knowledge compensate for their lack of knowledge by relying on their general reading skill, and low-skilled readers with high domain knowledge compensate for poorer reading skills by relying on their specific domain knowledge. They also concluded that domain knowledge and vocabulary knowledge have independent effects on comprehension, and that these effects are on what is comprehended as well as on how much is comprehended.

Research has also provided evidence for a substantial role of background knowledge in reading comprehension in a second language. In a number of studies, it was found that the cultural background of the reader may affect the type of information remembered. Bugel and Buunk (1996) demonstrated that the text topics of a foreign-language reading comprehension examination gave an advantage to boys, because the topics of the texts were of more interest to boys than girls. Steffensen (1987) found that

the religious background of ESL children affected their reading comprehension scores. Children with a Christian background turned out to have higher comprehension scores on a text describing a Christmas celebration. Lipson (1983) confirmed this finding. The study compared the reading comprehension of children in relation to their religious affiliation and found an effect of religious affiliation on reading comprehension when children read texts about a topic dealing with aspects of their familiar or unfamiliar religion. To sum up, the studies mentioned have asserted that having background knowledge of the topic of a text contributes by a great deal in reading comprehension.

4.7.2. Text variables

The second important variable which considerably affects reading comprehension is text variable. It is a truism that without written material, there will be no reading at all. The text is then at the centre of the reading process. Text, or written discourse, is the product of the writer's thought expressed through some visible shapes, whether printed or hand-written, in Braille, written on a piece of paper, carved on a hard surface, or displayed on a computer or mobile phone screen. The term text variable refers to the textual features of learning materials which influence comprehension and memory. Many aspects of text might facilitate or make the reading process difficult. Although the language of the text is known to be the major variable, there are other factors ranging from aspects of layout, text content, to text types, text organization, and sentence structures which have been shown to have an effect on students' learning and comprehension. Below is a brief consideration of these aspects.

4.7.2.1. The Effect of Typographical features on Reading Comprehension

Typographical features pertain to text legibility mentioned in Chapter 3 (see under section 3.2) which constitutes a variable that affects directly comprehension. Obviously, if the print on the page is difficult to read either because it is too small or the font is an odd one, then this will contribute to reading difficulty. Clear design and layout is also important and again the reader must be taken into consideration. What may be suitable for a younger reader (comic book style, large print, etc) would be patronising for others. No doubt then that the first text characteristic is whether the symbols and the code used are familiar to the reader or not, i.e., whether it uses Latin symbols, Arabic symbols, Chinese symbols, etc., and whether the code or language is familiar. Although there should be a complete match between the symbols and the language used in the text and those possessed by the reader, this is by no means a guarantee for the reader's success. Goodman (1973) believes that because readers read for meaning and with minimal use of graphic cues, the differences in L1 and L2 in terms of direction of reading (right to left in Arabic, top to bottom in Japanese, left to right in English) will not affect 'the basic reading process' considerably, while Grabe (1991) agrees that direction of reading will not influence reading comprehension. In the Algerian schools as far as reading in English is concerned such an issue should not be a problem as learners are used to the Latin scripts since the primary school where they learn a twin language, French.

4.7.2.2. The Effect of Text Cohesion on Reading Comprehension

First, of all we have to accept Carrell's (1982) assertion that cohesion is not coherence. Simply put cohesion is the way in which ideas are expressed clearly and how they relate to each other. Therefore, cohesion, assumes that different occurrences in a text have a relation with each other. Cohesion refers to grammatical and/or lexical features that link the component parts of a text together. It is the relationship between different sentences or different parts of a sentence (Harris and Hodges, 1981).

Halliday and Hasan (1976, p. 123) defined cohesion as the 'set of semantic configuration that is typically associated with a particular class of context of situation, and defines the substance of the text.' They recognized five types of cohesive devices in English. They are reference, substitution, ellipsis, conjunction, and lexical cohesion. Reference, substitution, and ellipsis are grammatical; lexical cohesion is lexical; conjunction stands on the border line between the two categories.

Several studies have established a link between cohesion and comprehension. For example, higher cohesion has been found to facilitate comprehension and recall (Beck & McKeown, 1991). For example, Ozuru et al. (2005) reported that students comprehended high cohesion biology texts better than low cohesion biology texts. Linderholm et al. (2000) reported that improving the causal cohesion of a text improved recall. Nevertheless, high cohesion is not always better. The optimal level of cohesion depends on the knowledge level of the reader. McNamara and colleagues (McNamara, Kintsch, Butler-Songer, & Kintsch, 1996; McNamara & Kintsch, 1996) found that high

knowledge readers performed better on comprehension tests after reading low cohesion texts as compared to high cohesion texts. In contrast, low knowledge readers benefited from high cohesion texts, as most researchers would expect. These results clearly indicate that there is a complex and subtle relationship between the characteristics of the text and the characteristics of the reader. Specifically, cohesion gaps may force a high knowledge reader to process the text more deeply, resulting in improved comprehension and recall.

4.7.2.3. The Effect of Text Coherence on Reading Comprehension

Another text-related factor important for comprehension is coherence. According to de Beaugrande and Dressler (1981, pp. 3-4) 'coherence relates to 'how the configuration of concepts and relations which underlie the surface text, are mutually accessible and relevant' '. These authors explained that:

Coherence, assumes that a text cannot convey a meaning unless there is an interaction between world knowledge and text knowledge, that is to say, there should be an interaction between what people have in mind and what the text presents as knowledge. (id.)

The role of coherence is then central when readers construct meaning from a text. However, this effect seems to depend on reader's prior knowledge of the text content: low knowledge readers benefit from coherence marking, whereas high knowledge readers benefit from a more implicit text (McNamara, et al., 1996).

Coherence has been recognized as more crucial than cohesion because, as most researchers have pointed out a text which is not coherent will be nonsense, although it may be cohesive. Coherence markers include connectives (because, so, however, although) and lexical cue phrases (for that reason, as a result, on the other hand). These markers explicitly guide the reader in interpreting the text and in connecting ideas with other ideas (Britton, 1994; Sanders and Noordman, 2000) and can, therefore, play an important role when readers construct meaning from discourse. Coherence markers have often been examined in combination with prior knowledge. Roller (1990) suggested an interaction between text structure and prior knowledge, in the sense that text structure helps the reader when the topic of a text is unfamiliar; but when the topic is familiar, the relations are already known to the reader.

4.7.2.4. The Effect of Text Type and Genre on Reading Comprehension

Text type and genre have also been recognized as factors affecting reading comprehension (Alderson, 2000, pp 83, 127). Basically and with the advent of technology and the internet there are two categories of texts: Linear and non-linear text, or hypertext. Linear text is the traditional category of text found in journal, magazines, and books. It is permanent and author-driven. Non-linear text is found on the internet and is therefore connected with a computer database. It can have audio, video, animated graphics, and other elements to enhance presentation. They are reader-driven as they allow the reader to direct his/her own learning through the use of links. The reader can browse freely around the document, and the author cannot predict the routes that the

reader will take. In case of comprehension problems, the reader has many resources available. Deficits in reading skills can be compensated for by resorting to supplemental resources available on the web. This ability to browse gives the reader certain autonomy as he/she navigates through hypertext. Any of the two categories of texts can belong to one of the traditional types of texts: narrative, expository, descriptive, argumentative, and instructive. Narrative text is known for having a literary or aesthetic purpose. Fictional stories, novels, and dramas are generally thought to fall within this genre. The purpose of expository text is to describe events or objects to inform, explain, or persuade (Just & Carpenter, 1987). Examples include textbooks, encyclopaedias, newspaper articles, and health promotional materials. The effect of these types of texts on reading comprehension has been investigated. It is not the purpose of this section to provide a detailed review of these studies but it may prove useful to consider their findings.

It has been stated that expository texts are less familiar, less predictable, and less 'considerate' than narrative text (Gordon, 1992), though there was a strong debate on the issue (Alexander, 1997; Gordon, 1992). Alexander (1997) suggests that expository text, when well written, can provide emotional and cognitive interest, while many others, most notably Rosenblatt, disagree (Rosenblatt, 1994). Alderson (2000) stated that expository texts are harder to process than narrative texts, perhaps because of the greater variety of relationships among text units, or possibly due to greater variety of content types. A large number of empirical studies have demonstrated that narratives

typically have a hierarchical structure, that readers are sensitive to such structure, and when the structure is used to guide comprehension and recall, both are facilitated (Glenn, 1978; Mandler, 1984). In addition, narrative texts are more likely to induce visualization in the reader as part of the reading process than expository texts (Dennis, 1982).

4.7.2.5. The Effect of Text Organization on Reading Comprehension

Text organization or text structure refers to the characteristics of written material and the way ideas in a text are constructed and organized. The impact of different kinds of text structure has also been investigated as a potential variable that affects reading comprehension. The results of these studies, for example performance (Carrell, 1984) showed that certain more highly structured English rhetorical patterns were more facilitative of meaningful recall for non-native readers in general, indicating an interaction between a reader's prior knowledge of and processing strategies for text structure and the rhetorical organization of the text. Johnson (1981, p. 169) noted that organization of ideas in a text affects reading comprehension more than its language complexity. Johnston (1983, p. 25) cites a study by Freebody (1980) in which 'the order in which subjects read passages affected their comprehension.'

Carrell (1984) conducted a study with readers from Spanish, Asian, and Arabic language backgrounds in which different cultures favoured different organisational structures, which in turn affected readers' recall and comprehension. Carrell's study

showed that Arabs remembered best from expository texts with comparison structures, next best from problem-solution structures and collections of descriptions, and least well from causation structures. Similarly, Maxwell (1994, p. 68) reports a study in which the change of rhetorical organisation affected the degree of reading comprehension. Singhal (1998, p. 4) confirms that ‘differences in text structure can lead to differences in reading. The review of the literature confirms that there is a strong connection between reading comprehension and understanding text structures. Knowledge of text structure allows the reader to select and attend to the right details so comprehension can occur.

4.7.2.6. The Effect of Text length on Reading Comprehension

In Chapter 3 it was noted that the notion of comprehensibility of a text is closely related to a most familiar notion in reading comprehension, namely, readability. It was also noted that it depended on several factors among which the number of new words a passage contains. Thus, another factor which can contribute to the level of difficulty of a passage is its length. It is a common and intuitive belief among language teachers that the longer a passage, the more difficult it will be. With regard to this point, Chastain (1988, p. 234) contends that ‘language teachers usually favour short reading passages’. This is probably due to a desire not to overload the students for whom the process of deciphering print of unfamiliar linguistic code may already be laborious. Another reason may be the belief that the longer the text, the more unknown vocabulary and complex grammatical structures it will contain. The result would be that students focus

too much attention on language as opposed to meaning, and they read much too slowly and inaccurately. However, Stanovich's (1980) disagreed with such a view. This researcher recommended that students read longer passages. Stanovich's (1980) interactive-compensatory. Stanovich's model implies that when a reader cannot comprehend a text based on inside-the-head or top-down processing, he/she can resort to other sources of information such as text-based information. A second closely related implication is that the more text-based information the reader has access to, the easier the comprehension of the content of the text will be. Anderson (1984, p. 94) supports this position, saying that 'research on reading miscues has demonstrated that short items are harder to read than longer ones because reading involves building up expectations on the basis of redundancies.' Likewise, Feldmann and Stemmer (1987, p.255 cited in Wang, 2011) assert that, 'the more clues the learners are able to pick up, because of the natural redundancy of a text, the more developed is their foreign language competence and the better they will accomplish the task.' Alderson, (2000, p. 153-154) also maintains that, 'the length of input is clearly very important in reading. Texts may be a single word, a phrase (as in the case of public warnings and notices), single sentences (in advertisements), or they may be whole novels, 20-page academic articles, two-page memoranda, or two paragraphs on a postcard.' He further states that, 'whilst length is related to the amount of interpretation needed, this is not necessarily direct: a post card may be very enigmatic, a novel very explicit' (p. 154).

4.7.3. Writer variable

Since the reading is considered as an interactive process between the reader, the text and the writer, there must be an influence of some sort of the writer on the reader. Undoubtedly, the essence of a text owes much to its producer, i.e., the writer. It is obvious that the writer as a variable affecting reading comprehension cannot have a direct influence as the other variables do, he/she still contributes to the reading act, indirectly though. The interaction that takes place between the reader and the writer is different from that of the reader and the text or the interaction that occurs between a speaker and a listener as in this case both of the communication or interaction parties are present. However, the reader and the writer can be supposed to have some abstract form of interaction/communication, because when producing the text, the writer most certainly takes the characteristics of his/her readers into account (Nuttall, 1982, p. 14; Johnston, 1983, p. 14). In reading, the reader cannot consult the writer to clarify ambiguities, which is why readers may not get the intended message fully. Taylor (1985) argues that there is an active co-operation between the reader and the writer in that the reader tries to get the meaning out of the text by bringing meaning to the written material.

From another point of view, a writer has always an idea, or at least assumptions about who the reader will be as generally any piece of writing is intended to a given audience. However, the writer's assumptions about the reader will not always come true, and because no two people may have exactly the same background knowledge,

‘there always is a mismatch of some kind’ between the writer’s and the reader’s background and expectations (Nuttall, 1982, p. 7). The reader may make inferences based on prior knowledge which the writer assumes to be known by the reader and does not provide explicit information. Hence the reader may draw wrong conclusions and comprehension is hampered. Therefore, the degree of reading comprehension depends on the ‘active collaboration between writer and reader’ (Taylor, 1985, p. 5) on the one hand, and the interaction of other factors previously described on the other. Furthermore, incorrect inferences may be made because of a lack of relevant vocabulary knowledge, an issue considered in the foregoing chapters.

4.7.4. Context variable

Although researchers have mainly emphasized the role of reader and text and also the interaction between the two, there seems to be little research addressing the issue of the role of context. Context refers to something beyond the text itself. Generally context variable refers to all variables external to reader, writer, and text. These are environmental and situational elements, and the larger socio-economic context. Comprehension may, therefore, be affected by the time of reading, whether it be early in the morning, after a day’s work, etc.; and also by the place of reading, i.e., in the library, in the classroom, in an exam session, in a car, etc. Although such environmental elements may be considered trivial, they still have potential to affect comprehension. Another contextual factor is the social setting in which the act of reading takes place. Reading does not take place in a vacuum (Alderson, 2000, p. 25) and the situation in

which it occurs may have an impact on how it is comprehended. To cite Widdowson's (1978), the 'stop sign' at the entrance to a street means different things depending on whether one is driving or walking. While the sign has a single significance, it has different values depending on settings in which it occurs.

Conclusion

This chapter considered reading comprehension in a second/foreign. It started by pointing out that is reading comprehension one of the most researched issues as this is the ultimate goal of foreign language instruction. It was aimed at giving the reader into the nature of reading comprehension process. For this purpose, theories that describe reading comprehension were exposed as well as the variables that affect it.

To reiterate, four broad variables were discussed as factors affecting the process and the product of the act of reading: the reader, the text, the context, and the writer, each with some other inter-related elements. It should be made clear that none of these variables or their elements act separately; rather, the resulting outcome whether it be comprehension or non-comprehension, is the product of the interaction among all the variables on the one hand, and all the factors within each variable on the other hand. To arrive at comprehension, therefore, all variables and elements inside them have a share with different degrees of contribution. In the case of L2 learning, in addition to the above variables, another variable enters the scene, i.e., the relationship between L1 and L2 on the one hand (similarities or differences), and the degree of L2 readers' literacy

in L1 on the other. The success in reading comprehension owes much to the interplay of all the factors discussed above.

Although, as it has been shown above, there is an array of variables that come into play in the reading comprehension act, vocabulary knowledge emerges as a critical component of reading comprehension. Such an issue, being central to the present study, the next chapter (Chapter 5) will be devoted to it.

CHAPTER 5

VOCABULARY AND WORD KNOWLEDGE

Introduction

As argued in the introduction to this dissertation, academic success is closely related to reading proficiency which in turn is largely determined by the mastery of vocabulary. It was shown in the previous chapter that in the past few decades the field of language teaching and learning has witnessed a renaissance in research into the role of vocabulary (Bachman, 2000; Read, 2000) because of the determining role it plays in reading, listening, speaking, and writing and also because it has been shown to contribute greatly to achievement in school education (Schmidt, 1999). That is, students' word knowledge has been shown to be linked strongly to academic success. The term vocabulary being central to this study, it is necessary to delineate it and examine how it has been defined and used in linguistics. Consequently, it deserves a chapter to remove any ambiguity as to the use of the terminology. This chapter begins with considering how the concept of word has been defined and the criteria that are used to identify word as well as defining a few concepts and key words that will be used throughout this work. It aims to clarify what is meant by vocabulary and the different types of vocabulary. Finally, it reviews the literature on what is meant by word knowledge.

5.1. What is a word?

Right from its onset it should be pointed out that it appears from the review of the literature that the definition of the term 'word' is controversial and that an unambiguous and universal definition of 'word' is not available or possible for many reasons. The main one is that a word is a complex organization of different linguistic and extra-linguistic aspects. However, attempts have been made to define word in the context of a particular language. For obvious reasons, the present discussion is restricted to the definition of the term in the English language.

Because of the difficulties one faces in defining word in languages, attempts have been made to identify certain concepts that might solve the problem of identification of the word. One such attempt is that of Harold Palmer, and his concept of *monolog*, *miolog*, and *phiolog*, where monologs are words in conventional sense, represented graphically by a group of letters beginning and ending with a space functionally independent unit; miologs as being components of monologs, and recognizable linguistically as derivational and inflexional affixes and are made up of two or more monologs; and phiologs as being units containing monologs but representing in some way a semantic entity. In this classification, the miolog refers to morphology and the phiolog to syntax.

In linguistics, it is usual to distinguish four definitions of 'word'. The definition of word obeys criteria relating to orthography (spelling), pronunciation (phonology), and meaning (semantics). A detailed review of this different perspective is beyond the scope

of this study. However, a brief consideration is useful to delineate what is meant by ‘word’ and the different terms found in the literature such as vocabulary, lexis, lexical vocabulary, lexeme, lexicon, and other terms.

Furthermore, it should be pointed out that this study is primarily concerned with vocabulary in its written form. Hence, the focus will be on orthographic words as opposed to phonological words.

5.1.1. Definition by orthography

For most people, a word is a sequence of letters that we write consecutively, with no spaces. Conventionally, a linguistic item preceded and followed by a space is identified as word because of the convention of leaving space between two words in writing or in printing. Such a view relies on the spatial representation of word that is, the distribution of words in printed space. This criterion relates to the written representation of the language.

Thus, the paragraph above starts with the four words right, from, the, and onset. However, relying on spelling to identify a word poses a problem with compound words such as ice cream, common sense and thousands of others. Do they constitute one word each, or two separate words? They could be written with a hyphen ice-cream, commonsense but a glance at any dictionary would reveal that they appear as headwords in the first form.

5.1.2. Definition by phonology

Another criterion used to identify a word is its pronunciation. Referring to phonology, a word would be a sequence of sounds which can be pronounced on its own, with pauses on either side. Trask (n.d.) provided the following definition: “a phonological word is a piece of speech which behaves as a unit of pronunciation according to criteria which vary from language to language”. Harley (1996), states that:

A phonological word is sequence of sounds which is identified as a unit on the basis of how it is pronounced – a collection picked out by the phonology of a language. *Can't, bendable* and *dog* are phonological words... Phonological words that don't carry any meaning whatever. (p. 10)

She goes on to explain that:

While it's clear to most speakers of English what the phrases *kit and caboodle, to and fro, have (one's) druthers,* and *run the gamut* mean (respectively, “everything,” “back and forth,” “get one's way,” and “vary as widely as possible”), hardly any speakers know what the words *caboodle, fro, druthers,* or *gamut* mean in these expressions (no one would ever say “*Do you like John's druthers?*” (p. 12)

Phonological rules of English tell us that a word contains only one main (or primary) stress. Consider the following sentence, as pronounced it would be pronounced

by a native speaker in relaxed colloquial speech: ‘The boards of examiners are pleased’. This sentence, therefore, contains three primary stresses and, therefore, three phonological words. One obvious way of breaking up the utterance into phonological words is as follows: [*The boards of*] [*examiners are*] [*pleased*]. One might argue that it can be broken up in a different manner, but still the result would be much the same. It is clear then that a phonological word is totally different from an orthographic one. Obviously, in normal speech the breaks don’t occur after each word and consequently such a definition is to be discarded.

5.1.3. Definition by meaning

The third criterion that used to identify a word is meaning. This criterion is based on the belief that each word has a meaning, and that in a language, each unit of meaning or segment of meaning could be identified and separated from other units of meaning. Each item thus separated is called a word. This, as argued by Trask (n. d.), cannot be used as an independent criterion to define word because the meaning factor used here poses a problem in distinguishing morpheme and word. When this criterion alone is applied to identify a word, it is not possible to distinguish between a word, a phrase and an idiom. Even groups of words that combine to give a single meaning will come under this category. Consider for example the case of compound words and inflected forms. Applying this criterion will lead to the identification of these inflected forms as words, although these inflected forms may be more than words in some sense.

Consider the example provided by Schmitt (2000). The researcher gives the following list: *die – expire - pass away - bite the dust - kick the bucket - give up the ghost*. The six examples are synonyms, with the meaning “to die”. The term *synonym* here is used as referring to words that have approximately the same meaning. However, for the layperson, these synonyms are made up of from one to four words. *Die* and *expire* are single words, *pass away* could probably best be described as a phrasal verb, and the last three are idioms, an *idiom* being a string of words which taken together has a different meaning than the individual component words. Similarly, a phrasal verb is made up of a verb plus one or more other words, which also has an idiosyncratic meaning compared to the component words. Thus, Schmitt argued that there is not necessarily a one-to-one correspondence between a meaning and a single word. In English, meanings are in many cases represented by multiple words. Is, for example, ‘washing machine’ a single word or two? That is the question.

In order to circumvent the problem of defining ‘word’ and this problem of multiword units, the term *lexeme* (also *lexical unit* or *lexical item*⁵) was coined. These three interchangeable terms are all defined as “an item that functions as a single meaning unit, regardless of the number of words it contains.” Thus, all of the six examples above are lexemes with the same meaning. The entire store of lexical items in a language is called its *lexis*.

⁵ The three terms will be used interchangeably throughout this study.

Bluntly, in linguistics, *lexis* is the total word-stock or lexicon having items of lexical rather than grammatical, meaning” (*Cambridge Advanced Learner's Dictionary*).

In addition another problem is whether *talk*, *talked*, *talking*, and *talks* should be counted as a single word or four. Likewise, are *regulate*, *regulator*, *regulatory*, and *regulation* the same word? In these examples, there is a *base*, *root*, or *stem* word that is the simplest form of that word. To this stem, affixes are added. If the purpose of the affixes is grammatical, then the resulting word is called an *inflection*. *Talked*, *talking*, and *talks* are inflections of the root word *talk*. However, if the affixes change the word class of a stem, the result is a *derivative*. Thus *regulatory* (adjective) and *regulation* (noun) are derivatives of *regulate* (verb). It is clear that although these words have different *orthographic* (written) shapes, they are closely related in meaning. Due to the importance of the term *lexeme* to this study, it deserves to be considered in more details. However, a definition, albeit short, of the key terms and concepts that are directly relevant to the present study and that would serve to remove any ambiguity is necessary.

5.2. Key terms and concepts

The following is a presentation of the key terms and concepts and that will be used throughout this study. However, two terms deserve a more detailed discussion, due their importance in this study. These are *vocabulary* and *lexis/lexical*.

Grapheme

In writing, a sequence of letters separated from other sequences of letters by a space. This corresponds to a popular understanding of a 'word'. The word *postbox* is one grapheme, while *post box* consists of two graphemes. Most people would say that these are simply two alternative written forms of a single word. (The term orthographic word is also used in the same sense. The equivalent term for speech would be phonological word, although there are, of course, no 'spaces' between words in speech.)

Word form

Nouns and verbs typically have more than one form. For example, speech has the plural form speeches, while speak has the forms speaks, speaking, spoke and spoken. Adjectives have comparative, superlative and adverbial forms (quiet, quieter, quietest, quietly). Because these forms affect grammatical function, but not meaning, we tend to think of them as different word forms, rather than different words.

Word family

A word family is usually held to include the base word, all of its inflections, and its common derivatives. It is then a group of related word forms, such as *speak, speaks, speaking, spoken, speaker, and speech*. There is evidence that the mind groups the members of a word family together, giving a psychological justification for using word families as a unit for counting and teaching (Nagy et al., 1989).

Lemma

A lemma is the representative form of a lexeme (a set of forms) that has the same meaning. It consists of the set of different forms of a word, such as the inflected forms of a verb. In English, *run*, *runs*, *ran* and *running* are forms of the same lexeme, with *run* as the lemma.

The technical term is lemma which generally means the same as "headword."

Grammatical words

These are words which have a grammatical function, rather than a sense or meaning. These include most conjunctions, determiners, pronouns, prepositions and some verbs (also called function words). Grammatical words are sometimes described as a closed class (there only around 150 of them in English), while lexical words are an open class.

Word class or parts of speech

The term refers to the typical grammatical function of a word - e.g. whether it is a noun, verb, adjective, etc. Some words belong to more than one word-class: (e.g., *talk* can be both a verb and a noun). Most linguists consider *talk*, verb, and *talk*, noun, to be different lexical items.

Morpheme

A morpheme is a unit of meaning or grammatical function. Act is a morpheme, because it cannot be broken down into smaller meaningful parts (e.g. *a-* and *-ct*, or *ac-* and *-t*). React is composed of two morphemes because it can be broken down into smaller meaningful parts: *re-* and *-act*. Both of these morphemes contribute meaning to

the word react. The -s on the end of reacts is also a morpheme. But this -s contributes to grammatical function, rather than meaning.

Derivation

It is the process of creating new words (derivatives) by combining morphemes. This is most often done by adding prefixes and/or suffixes to a stem. The word reaction is formed by a combination of re- (prefix), -act (stem) and -ion (suffix). Prefixes and suffixes that contribute to the meaning of a word (e.g., re- in reaction) or change word class (e.g., -ion in reaction) are called derivational morphemes.

Inflection

Inflection is the process of varying word-forms for grammatical reasons. The ‘third-person present-tense’ -s at the end of reacts and the plural -s at the end of reaction, for example, are inflectional morphemes. There are only 8 inflectional morphemes in English.

Polysemy

Words with more than one sense (i.e. more than one ‘out-of-context’ meaning) are polysemous words. For example, fair (adj.) can mean ‘light in colour’, ‘reasonable’, or ‘average’. Fair is also a noun, but as we have already noted, linguists would consider fair (adj.) and fair (n.) to be different lexical items, rather than different senses of a single polysemous lexical item.

Compound

A compound is a fixed term formed by combining two existing English words: usually an adjective and a noun (e.g. black hole), or a noun and a noun (e.g. rocket engine). Short compounds may be written in three different ways, which do not correspond to different pronunciations. The "solid" or "closed" forms in which two usually moderately short words appear together as one. Solid compounds most likely consist of short (monosyllabic) units that often have been established in the language for a long time. Examples are housewife, lawsuit, wallpaper, etc.

In the hyphenated form two or more words are connected by a hyphen. Compounds that contain affixes, such as house-build(er) and single-mind(ed)(ness), as well as adjective-adjective compounds and verb-verb compounds, such as blue-green and freeze-dried, are often hyphenated. Compounds that contain articles, prepositions or conjunctions, such as rent-a-cop, mother-of-pearl and salt-and-pepper, are also often hyphenated. The open or spaced form consisting of newer combinations of usually longer words, such as distance learning, player piano, lawn tennis, etc are not hyphenated. Usage in the US and in the UK differs and often depends on the individual choice of the writer rather than on a rule; therefore, open, hyphenated, and closed forms may be encountered for the same compound noun, such as the triplets container ship/container-ship/containership and particle board/particle-board/particleboard.

In addition to this native English compounding, there is the classical type, which consists of words derived from Latin, as horticulture, and those of Greek origin, such as photography, the components of which are in bound form (connected by connecting

vowels, which are most often -i- and -o- in Latin and Greek respectively) and cannot stand alone (Wikipedia).

Idiom

Most often this term refers to a phrase or expression that cannot be understood by knowing what the individual words in the phrase mean. The *Cambridge Advanced Learner's Dictionary* defines the term idiom as: “a group of words in a fixed order that have a particular meaning that is different from the meanings of each word understood on its own. To "have bitten off more than you can chew" is an idiom that means you have tried to do something which is too difficult for you. Another example is "to roll out the red carpet" meaning to extravagantly welcome a guest; no red carpet is needed.

Collocations

Collocation is the relationship between two words or groups of words that often go together and form a common expression. A collocation is then a familiar grouping of words, especially words that habitually appear together and thereby convey meaning by association. Examples of collocations are: *iron will* and *nerves of steel*, *like two peas in a pod*, and *absolutely convinced*. We talk of high mountains and tall trees, but not usually of tall mountains and high trees. Collocations are remembered as wholes and often used together.

Collocation is ‘the readily observable phenomenon whereby certain words co-occur in natural text with greater than random frequency’ (Lewis, 1997, p. 8).

Furthermore, collocation is not determined by logic or frequency, but is arbitrary,

decided only by linguistic convention. Some collocations are fully fixed, such as "to catch a cold," "rancid butter," and "drug addict," while others are more or less fixed and can be completed in a relatively small number of ways, as in the following examples: blood / close / distant / near(est) relative; learn by doing / by heart / by observation / by rote / from experience; badly / bitterly / deeply / seriously / severely hurt.

Institutionalized utterances

These are expressions such as: "I'll get it", "I'll see", "That'll do", "If I were you", "Would you like a cup of coffee?" etc.

Sentence frames and heads

These are expressions such as: "That is not as...as you think", "The problem was".

Text frames

These are Expressions such as, "In this paper we explore...; Firstly...; Secondly...; Finally..."

5.3. What is a lexical item/lexeme?

A lexeme is an abstract unit. It can occur in many different forms in actual spoken or written sentences, and is regarded as the same lexeme even when inflected. For example, in English, all inflected forms such as *give*, *gives*, *given*, *giving*, *gave* would belong to the one lexeme *give*. Similarly, such expressions as *bury the hatchet*, *hammer and tongs*, *give up*, and *white paper* (in the sense of a government document) would

each be considered a single lexeme. In a dictionary, each lexeme merits a separate entry or sub-entry (*Dictionary of Language Teaching and Applied Linguistics*).

Much of the following discussion of the term *lexis* is based on Trask (n. d.). According to this researcher, a *lexical item* (or *lexeme*) is an abstract unit of the lexicon (vocabulary) of a language, with a more or less readily identifiable meaning or function. It is a word in the sense in which a dictionary contains words.

It is a word or words representing a single unit of sense or meaning. No matter how we write it, *post-box* or *post box* represent a single unit of sense. The terms *vocabulary item* and *lexeme* are used in the same sense. Lexical items made up of two or more graphemes are often called *multi-word units* or *multi-word lexical items*.

A lexical item is represented in speech or writing by one of the possibly several forms it can assume for grammatical purposes. For example, if we want to mention canine animals, we must use either the singular form *dog* or the plural form *dogs*. But these two grammatical forms both represent the same single abstract unit, the same lexical item. We can conveniently represent that lexical item as *dog*. Then *dog* and *dogs* are the two possible forms of the lexical item *dog*.

In the same way, we recognize a lexical item *take*, which can be represented by any of the five grammatical forms *take*, *takes*, *took*, *taken* and *taking*. A dictionary provides entries for lexical items. So, for example, we do not expect to find separate entries in the dictionary for *dog* and *dogs*: we expect to find only one entry for the lexical item *dog*. Likewise, we expect to find only one entry for the lexical item *take*,

and not five entries for the five forms of that lexical item. In English, the lexical item *BE*, uniquely, has eight different forms, but no other English lexical item has more than five, and most have fewer than this – sometimes only one.

5.3.1. Citation forms

The *citation form* (or *dictionary form*) of a lexical item is the particular grammatical form of it which we use in naming it, talking about it, and entering it in a dictionary. In English, almost every English lexical item has one form which carries no grammatical marking at all. This form, the base form, is the natural choice for the citation form. For a noun like *dog*, the citation form is the singular *dog*, and not the plural *dogs*. For the verb *take*, the citation form is the infinitive *take*, and not an inflected form like *taking* or *took*. For the adjective *big*, the citation form is the positive form *big*, and not an inflected form like *bigger*. For the preposition *with*, there is nothing to talk about: this lexical item has only the single form *with*, and this is therefore its citation form. Some lexical items are defective. A *defective* lexical item is one that lacks some of the forms normally exhibited by a lexical item of its class. For example, the nouns *furniture* and *happiness* are defective, since they have no plural forms, while the nouns *oats* and *police* have no singular forms. The verbs *must* and *beware* both lack a number of the forms exhibited by most verbs.

5.3.2. Grammatical word-forms.

A *grammatical word-form* (or *grammatical form*) is one of the several forms that may be assumed by a lexical item for grammatical purposes. so, for example, *dog* and

dogs are grammatical forms of the lexical item *dog*, and *take*, *takes*, *took*, *taken* and *taking* are all grammatical forms of the lexical item *take*. Sometimes a lexical item exhibits only one grammatical form. For example, the lexical items *oats*, *police*, *beware* and *with* have only one grammatical form apiece: *oats*, *police*, *beware* and *with*.

Nevertheless, it may still be useful to distinguish the lexical item *oats* from its sole grammatical form *oats*, and likewise for the others. We have already seen that an English lexical item usually has one grammatical form which carries no grammatical marking at all, its *base form*. All the other grammatical forms of a lexical item, the ones that carry grammatical markings, are the *inflected forms* of that lexical item. so, for example, *dog* has the base form *dog* and the inflected form *dogs*, while *take* has the base form *take* and the inflected forms *takes*, *took*, *taken* and *taking*. But *police*, *beware* and *with* have no inflected forms at all, while *oats* has *only* the inflected form *oats* (a complication: since the form *police* is plural, it must bear an abstract inflection for plurality. some linguists might therefore prefer to say that *police* is an inflected form, even though it carries no overt marking.) In some analyses, the number of grammatical forms exhibited by a lexical item may be larger than the number of overtly distinct forms. Here is an example: some English verbs exhibit different forms for the past tense and the past participle. The verb *take* is one of these: *she took the exam*, but *she has taken the exam*. Many verbs, however, present identical forms for the past tense and the past participle, an example is *finish*: *she finished the exam*, but *she has finished the exam*. Since the distinction between the past tense and the past participle is essential for

English verbs in general, most linguists prefer to say that past-tense *finished* and past-participle *finished* are two *different* grammatical forms of the verb *finish*.

Here is a slightly less obvious example. For any English verb, the present tense forms (usually apart from the third-singular form) are identical to the infinitive. We can see this with *take*: *you should take an umbrella* (infinitive *take*), but *I always take an umbrella* (present-tense *take*). Even though there is *no* English verb which exhibits different forms for the infinitive and the present tense, some linguists would still prefer to say that we are looking at two different grammatical forms here, because of the difference in grammatical behaviour. A further third example is the *-ing* form of an English verb which has at least three different grammatical functions. This can be illustrated with taking. First, *taking* can be a participle: *taking him for a stranger, I didn't greet him*. Second, it can be part of a progressive (continuous) verb-form: *Paul is taking photos*. Third, it can be a gerund: *taking breakfast gives energy for the whole morning*. Other linguists would prefer to say that all three of these functions represent *different* grammatical forms of *take*, even though the three forms are always identical for every verb. Still there is there is one more item of the form *taking* as in the following example from The *Hachette-Oxford Dictionary*⁶ “the money was there for the taking”. But this *taking* is not a grammatical form of the verb *take*, and in fact it's not a verb-form at all: it's a noun, and it represents a distinct lexical item *taking*. This noun *taking* is related to the verb *take* by a derivational process, much as the noun *arrival* is

⁶ All the examples that follow have been inspired from The *Hachette-Oxford- Dictionary* (soft copy).

related to the verb *arrive*. It should be noted here that many confuse verbal nouns in *-ing* with gerunds, and apply the label “gerunds” to all of them indiscriminately. The gerund *taking* is a verb form (for example, it can take an object, as in the example above), and it is an inflected grammatical form of the verb *take*. However, the verbal noun *taking* is not a verb form at all, but a noun: it behaves like a noun, and it has no verbal properties at all.

5.3.3. Inflection and derivation

We have seen that a lexical item can appear in several grammatical forms, some of which carry inflections (overt grammatical markings). It is essential to distinguish the process of inflection from the quite different process called derivation. Inflection is the variation in form of a lexical item for grammatical purposes. Derivation is the construction of a new lexical item from another lexical item, usually by the addition of an affix (a prefix or a suffix). Take the lexical item *cat*, which has two grammatical forms: *cat* and *cats*. But *catty*, as in *a catty remark* is not another inflected form of *cat*. Instead, it is a different lexical item, *catty*. *Cat* is a noun, and all its grammatical forms are nouns, while *catty* is an adjective. The relation between *cat* and *catty* is one of derivation: we say that *catty* is derived from *cat* by the addition of a suffix. Since *catty* is a different lexical item from *cat*, we find *catty* entered in the dictionary as a separate entry. In the same way, *doglike* is a different lexical item from *dog*, and it too has an entry in the dictionary. A few more examples of derivation by suffixation are *happiness* from *happy*, *arrival* from *arrive*, *singer* from *sing*, *slowly* from *slow*, *kingdom* from

king, *motherhood* from *mother*, and *reddish* from *red*. Some examples involving prefixation are *rewrite* from *write*, *unhappy* from *happy*, *non-smoker* from *smoker* and *pre-war* from *war*. Some examples of derivation are irregular, such as *destruction* from *destroy*, *calculable* from *calculate*, *comprehensible* from *comprehend* and *medical* from *medicine*. Thus, every lexical item obtained by derivation is expected to appear found as separate entry for. In practice, however, some dictionaries often do not do so to save space especially when a derived lexical item has a regular form and a transparent meaning, as with *unhappy*, *happiness*, *rewrite* and *slowly*, a single entry is provided. A large dictionary may simply list the word without definition, either at the end of the entry for its source word (look for *slowly* under *slow*), or at the bottom of the page (look at the words in *un-* in a big dictionary), while a small dictionary may omit the word altogether. English exhibits an interesting type of derivation called zero-derivation, or conversion, by which a lexical item is simply shifted from one word-class to another, without adding any material. For example, the adjective *brown*, as in *brown shoes*, can occur as a verb, as in *brown the meat*, and the noun *doctor*, as in “*that's just what the doctor ordered!*” has been converted into a verb, as in *the party in power doctored the elections*. English allows this with some freedom, and examples can be elaborate. For example, the noun *smoke*, as in “*full of tobacco smoke*”, has been converted into a verb, as in *Paul smokes*, and this verb has been converted into a different noun, as in *Paul was having a smoke*. In all these cases, we must regard the zero-derived word as

belonging to a different lexical item from the source, because it belongs to a different word class.

5.3.4. Multi-part and discontinuous words

These are items which appear to be words by some criteria, but which exist in two or three pieces, sometimes separated by other items. The most obvious English examples are the phrasal verbs. A phrasal verb consists of a simple verb plus a particle, or rarely two particles. Examples are *make up*, *take off*, *turn on*, *ring up* and *put up with*. A phrasal verb represents two (or three) orthographic and phonological words but it is usually considered a single lexical item, since its meaning is often unpredictable from the meanings of its components. Consider the several items of the form *make up*: *She made up a story*; *she made up her face*; *she made up the numbers*. Many phrasal verbs can occur discontinuously: *she took off her dress* or *she took her dress off*; *she turned on the light* or *she turned the light on*; and so on. It appears, then, that we must recognize lexical items which occur in pieces, sometimes in pieces which are separated by other items. English has another class of verbs, the prepositional verbs. A prepositional verb consists of a simple verb (possibly plus an adverb) plus a preposition. Examples are *call on*, *look at*, *take to*, and *look down on*: *We called on Paul*, *Look at me*, *Paul looks down on strangers*. Prepositional verbs differ from phrasal verbs in ways that are explained in reference grammars of English. In these examples, sequences like *on Paul* and *on strangers* appear to be syntactic units, prepositional phrases. Nevertheless, it appears that we must recognize *call on* ('visit') and *look down*

on ('regard with contempt') as single lexical items, because of their meanings. In this regard Trask concluded:

We are thus faced with conflicting analyses: the item *on* is syntactically part of a prepositional phrase, but lexically part of a multi-part verb. Examples like these present formidable problems of analysis, and they stretch our conception of words to the breaking point. (n.d.)

5.3.5. Content words and grammatical words

Lexical items are commonly divided into content words and grammatical words. A content word is a lexical item which has semantic content – that is, which has a readily identifiable meaning. A grammatical word (or function word) has little or no identifiable meaning, but has one or more grammatical functions. A content word can be defined, and it can often be translated into another language with some ease. English content words include *house*, *school*, *play*, *huge* and *randomly*. A grammatical word cannot be defined, and looking for an equivalent in another language is often pointless. English grammatical words include *of*, *the*, *and*, *have*, *who* and *if*. A dictionary cannot provide a definition of one of these, but can only give an account of its grammatical functions: the “meaning” of the word *is* cannot be provided. The English phrase *a bottle of milk* contains the two grammatical words *a* and *of*.

5.3.6. Clitics

A *clitic* is a grammatical form which cannot stand on its own in an utterance. It needs to co-occur with another form which either precedes or follows it. It is a morpheme that has syntactic characteristics of a word, but shows evidence of being phonologically bound to another word. It often has grammatical rather than lexical meaning and belongs to closed classes like pronouns, prepositions, auxiliary verbs, and conjunctions. Since a clitic cannot form a phonological word by itself, it must be phonologically bound to a more substantial item, its *host*, with which it forms a phonological word, possibly together with other clitics. Among the English clitics are the articles *a(n)* and *the*, which are bound to a following host, as in *a book* and *the man*, and the auxiliary *'ll*, which is bound to a preceding host, as in *John'll do it*. To take an example from French, a language which has large number of clitics, in the sentence *il te le donnera* (he'll give it to you), the pronouns *il* 'he', *te* 'you' and *le* 'it' are all clitics bound to the verb that follows.

Clitics come in two types: proclitics and enclitics. Proclitics attach themselves to the beginning of a word. Enclitics attach themselves to the end of a word. English enclitics include the abbreviated forms of be: *'re* as in *you're*, *'s* as in *she's*, the abbreviated forms of auxiliary verbs: *'ll* as in *they'll*, *'ve* as in *they've*; the genitive case (or "possessive") marker *'s* as in *John's car*, the negative marker *n't* as in *couldn't*. English proclitics include: the articles *a*, *an*, *the*.

5.3.7. Short forms

English exhibits a number of short forms of several kinds, and these short forms merit some discussion from the point of view of “wordhood”.

5.3.7.1. Abbreviations

An abbreviation is a short way of writing a word or a phrase, using only letters of the alphabet and possibly full stops. An abbreviation is strictly a written form. It has no pronunciation of its own, and it can only be pronounced by pronouncing the full form which it abbreviates – or, in some cases, by spelling it out letter by letter. Familiar English abbreviations include *Prof.* for *Professor*, *Sgt* for *Sergeant*, *Dr* for *Doctor* (only as titles accompanying names), *BC* for *before Christ*, *mph* for *miles per hour*, *kg* for *kilograms* and *C* for *degrees Celsius*. The forms *Mrs* and *Ms* are treated as abbreviations, even though they do not have a longer form. An abbreviation consisting of only one or two letters may sometimes be spelled out letter by letter. This is commonplace with *a.m.* (/eɪ em/) and *BC* (/bi: si:/). In other cases, this practice is usual only in reading out a written text, as with *e.g.* (/i: dʒi:/) and *i.e.* (/aɪ i:/). An abbreviation is usually considered as an orthographic word (as defined above), but it is not a word in any other sense. Abbreviations are not usually used in spoken language.

5.3.7.2. Logograms

A logogram is a written character which is not a letter of the alphabet but which conventionally represents a word, or rarely a sequence of words (Trask, n.d.). Familiar logograms include the digits, like '5' for *five*, monetary symbols like '£' for *pounds*, and arithmetical symbols like '+' for *plus*, '=' for *equals* and '%' for *per cent*. A logogram is a representation of a lexical item or of a grammatical word-form. It fits the definition of an orthographic word given above, but some linguists might prefer to restrict that definition to cases consisting of letters of the alphabet, in which case a logogram would not be an orthographic word.

5.3.7.3. Contractions

A *contraction* is a conventional brief way of pronouncing a sequence of two (or rarely three) words which often occurs together. A contraction always has a distinct written form. Typical contractions are *I'm* for *I am*, *it's* for *it is* or *it has*, *she'll* for *she will*, *couldn't* for *could not*, *won't* for *will not*, *hadn't* for *had not*, and (more informally) *she'd've* for *she would have*. Some contractions have unexpected forms, such as *won't* for *will not*. A contraction is always a single orthographic word and a single phonological word. But it represents two (or three) lexical items, and two (or three) grammatical word-forms.

5.3.7.4. Acronyms and Initialisms

In contemporary English, new words have been coined in the following way: the initial letters of the most important words of a phrase are extracted. The sequence of letters extracted are put together to form a new word – which almost always has the same meaning as the original phrase. A few examples are: *BBC* for *British Broadcasting Corporation*, *FBI* for *Federal Bureau of Investigation*, *NATO* or *Nato* for *North Atlantic Treaty Organization*, *AIDS* or *Aids* for *Acquired Immuno-Deficiency Syndrome*. Two kinds of outcome may be distinguished here, and there exist differences in terminology between British and American English. Sometimes the resulting form can only be pronounced by spelling it out letter by letter, as with *BBC* and *FBI*. In British English, a formation of this kind is Initialism. In other cases the resulting form can be pronounced like an ordinary word, as with *NATO* and *AIDS*. In the British English, a formation of this kind is an *acronym*. However, in American English, the label “acronyms” is applied to all such formations without distinction, and the term “initialism” is not used. There is a further difference between British and American English. In all varieties of English, Initialisms like *BBC* and *FBI* are written entirely in capital letters. The Americans also write most acronyms (in the British sense) entirely in capitals, and so they write *NATO* and *AIDS*. But the British commonly write these things with only an initial capital, and the usual British forms are *Nato* and *Aids*. But not all acronyms are written with capitals. Some acronyms have become perfectly ordinary lexical items, and they behave accordingly. For example, the phrases *self*

contained underwater breathing apparatus and *light amplification by the stimulated emission of radiation* have given rise to the acronyms *scuba* and *laser*.

In constructing an acronym, small grammatical words are ignored in order to obtain a result which can be easily pronounced. A good example is *radar*, from *radio detection and ranging*. Initialism or an acronym is a lexical item. An acronym does not differ from any other lexical item, except perhaps in its unusual written form. Initialism has both an unusual written form and an unusual pronunciation, but otherwise it is an ordinary lexical item. Initialisms are sometimes confused with abbreviations, but they are not abbreviations. First, initialism always has its own pronunciation, distinct from the pronunciation of the longer form which it represents. Abbreviations do not usually have their own pronunciations. Second, initialism, being a lexical item, can appear in a structural position in a sentence in which a lexical item is appropriate. For example, we can say or write *the BBC's decision*, in which the initialism *BBC* bears the possessive suffix -'s.

5.3.7. 5. Clipped forms

A *clipped form* is an item which is obtained by extracting a piece from a longer word or phrase. The process of extraction is *clipping*. In all but very rare cases, a clipped form has the same meaning as the longer form from which it is obtained. Contemporary English uses a lot of clipped forms. We find *gym* from *gymnasium*, *flu* from *influenza*, *fridge* from *refrigerator*, *phone* from *telephone*, and *gator* from *alligator*. As these examples show, any convenient part of a longer word may be

clipped. Even discontinuous pieces can be clipped, as with *sci-fi* for *science fiction*, *sitcom* for *situation comedy*, *biopic* for *biographical picture*, and British *maths* for *mathematics* (compare American *math*), and *hi-fi* for *high fidelity*. These clipped forms appear in dictionaries (for example the *Hachette-Oxford Dictionary*) as separate entries. It should be stressed that a clipped form is not an abbreviation. It is a genuine lexical item, just like any other lexical item. A clipped form accepts the grammatical inflections which are typical of its word class. For example, nouns obtained by clipping can pluralize: *gyms*, *fridges*, *phones*, *gators*. Verbs obtained by clipping can take ordinary verbal inflections. For example, the noun *disrespect* has given rise to a noun *diss*, which in turn has yielded a verb *diss*⁷ ‘treat with disrespect’, and this verb behaves like any other verb: *He was dissing me*. A clipped form can enter into compounds like any other lexical item: *gym shoes*, *phone book*. Clipped forms are entered in dictionaries like other lexical items. Sometimes a clipped form displaces its original longer form. For example, the clipped forms *piano* and *bus* have completely supplanted their sources, *pianoforte* and *omnibus*, and *mob* has displaced its source *mobile vulgus* ‘the fickle crowd’. The words *bra* and *cello* are so close to replacing their source words *brassiere* and *violoncello* that longer forms are rarely used. Note the following pair of examples. As observed above, English has an abbreviation *Prof.* (with dot) for *Professor*, as in the written form *Prof. Chomsky*. The other abbreviation *prof* (without dot and without capital letter) behaves like a lexical item *prof*, obtained from *professor*

⁷ Example given by Trask (n.d.) but does not exist neither in the *Hachette-Oxford Dictionary*, nor in *The American Heritage Dictionary*.

by clipping, as in “*This board of examiners is composed of two profs*”. The abbreviation *Prof.* and the lexical item *prof* are not the same item at all, and they should not be confused. Still, the abbreviation *prof.* stands for *professional*.

5.3.8. Types of words

There exists in the literature a plethora of types of words, some of which are:

Neologism: A neologism is “a purely arbitrary combination of words, not derived in whole or in part from any existing word” Stockwell and Minkova (2001, p. 7). An example of neologism is the word *blurb* created in 1907 and which means the embellished descriptions on the jackets of books.

Loanword: A loanword (or loan word) is a word belonging to a foreign language and “borrowed by another language, completely or partially naturalized such as *hors d'oeuvre, à la carte, adieu, garage, vinaigrette, brunette, façade, canapé, and brochette* from French; *alcohol, algebra, azan, Bedouin, bulbul caliph, couscous,* and *halal* from Arabic; *broccoli* and *spaghetti* from Italian; *aficionado* (an enthusiastic admirer or follower; a devotee or a fan) from Spanish and the list is long.

Nonce word: A nonce word is a word invented for a particular occasion or situation. Obviously words belonging to this category must be rare and do not have room in a textbooks for English as a second/foreign language. An example of this (from *The American Heritage Dictionary*) is the word *mileconsuming* in “*the wagon beginning to fall into its slow and mileconsuming clatter*” (William Faulkner).

Nonsense word: A nonsense word is a word or part of a word that do not exist in the language. They are usually found in child language or coined by researchers for research purposes.

Homograph: One of two or more words that have the same spelling but differ in origin, meaning, and sometimes pronunciation are called homographs. An example of such a words are *bow* /bau/ meaning the front of a ship, *bow* meaning a loop made in a string or ribbon, and *bow* /bəu/ meaning a device used to shoot arrows (*Cambridge Advanced Learner's Dictionary*).

Homonym: Two words that sound the same but have different meanings are known as homonyms. *No* and *know* are homonyms.

Blending or portmanteau word: A blending or portmanteau word is formed by merging the sounds and meanings of two different words (usually the first part of one word and the second part of the other). Examples of such words are: *smog*, a blend of *smoke* and *fog*; *motel* from *motor* and *hotel*; *chortle* from *chuckle* and *snort*; and *brunch* from *breakfast* and *lunch*. Technical terms are often created by blending. *Medicare*, form *medical care* is such an example. Other examples are: *urinalysis*, *guesstimate* (approximate estimation), and *spamouflage*.

Umbrella word: This type of word describes a single concept, idea, etc., that contains or covers a lot of different others, etc. For example, *corn* is an umbrella word for wheat, barley and oats (*Collins' English Language Dictionary*).

Having considered the various terms and concepts having a relation with the definition of word, we now turn to consider what is involved in knowing a word. But before doing so we need to consider what is meant by *vocabulary*.

5.4. What is Vocabulary/lexicon?

The term *vocabulary* is often equated with *lexicon*. Most dictionaries provide the following definitions for the term *vocabulary*:

1. The stock of words used by or known to a particular people or group of persons;
2. A list or collection of the words or phrases of a language, technical field, etc., usually arranged in alphabetical order and defined.
3. The words of a language.

Broadly defined, vocabulary is knowledge of words and word meanings.

However, vocabulary is more complex than this definition suggests. First, words come in two forms: oral and print. Oral vocabulary includes the words used in listening and speaking. Print vocabulary includes the words used in reading and writing. Second, word knowledge also comes in two forms, receptive and productive. Receptive vocabulary includes words we recognized when we hear or see them. Productive vocabulary includes words used in speaking and writing. Receptive vocabulary is typically larger than productive vocabulary, and may include many words to which we assign some meaning, even if we don't know their full definitions and connotations – or ever use them ourselves as we speak and write.

The vocabulary items of a language can also be classified under various groups on the basis of the features shared by the vocabulary items of individual groups. Each vocabulary item may have its own characteristics in terms of meaning and grammatical function (Mallikarjun, 2002).

5.4.1. Types of Vocabulary

Vocabulary is generally divided into several broad groups.

5.4.1.1. General vocabulary

The general vocabulary consists of such words as *man*, *nice*, and *go* which can be used in a variety of situations. The special vocabulary, on the other hand, consists of words with specific meanings that can be used only in certain situations (Hurlock 1972). Some of the special vocabulary groups are: colour vocabulary, number vocabulary, time vocabulary, money vocabulary, slang vocabulary, and swearing vocabulary. However, motivated and guided by pedagogical needs, another classification has been established: *active and passive vocabulary*.

5.4.1.2. Active and Passive Vocabulary

Child (1973) considers the 'number of words we actually use' as active vocabulary and the 'larger number we are able to understand' as passive vocabulary.

According to Finocchiaro (1958), the active vocabulary is the one that is learned very intensively with respect to form, meaning, and use in such a way that the learner will be able to use it in all the listening, speaking, reading, and writing activity. In

contrast, the passive vocabulary is the one that is understood by the students in a spoken or written context, but the student cannot reproduce the same on his own.

This definition of Finocchiaro provides a pedagogic dimension to active and passive vocabulary. But this type of vocabulary classification does not help much to solve the pedagogic problems. The same vocabulary item which is in active use today may become a passive vocabulary item tomorrow; the reverse of this is also true. That is, passive vocabulary item of today may become an active one tomorrow.

It is important to stress again that the number of passive vocabulary items of vocabulary of an individual is very high, compared to his/her active vocabulary. This is so because passive vocabulary could include also the active vocabulary. The active vocabulary items of an individual may be more frequency used in the language when compared to the frequency of the passive vocabulary items that an individual has. This obviously may not be true for all vocabulary items and for all individuals and the contexts. However, it is assumed to be true in the context of the present study.

5.4.1.3. Recognition and Reproduction Vocabulary

Recognition vocabulary is defined as that which is recognized or identified in listening or reading. The reproduction vocabulary is identified either in speaking or writing. The recognition vocabulary of an individual is much larger than his reproduction vocabulary.

5.4.1.4. Writing, Reading, and Speaking Vocabulary

The terminology itself reveals the criterion used to classify the vocabulary. Here, language skills are used as a criterion to classify the vocabulary items. Writing vocabulary includes all the words one uses or should be able to use in writing. Reading vocabulary is that which comprises the words one uses or should be able to use in reading. Speaking vocabulary is that which is usually quite different from the reading vocabulary. It is typically more informal. It includes many words that are not frequently written or found in reading materials.

5.4.1.5. Academic and Non-Academic Vocabulary

The vocabulary items that are acquired informally and that do not cause any difficulty in learning are called non-academic vocabulary. Vocabulary items that are formal, used more in writing than in speaking, and are associated with sciences, humanities, and other areas of formal learning are called academic vocabulary.

5.4.1.6. Form and Content Vocabulary

Leaving aside all the extra-language criteria for the classification of vocabulary, Michael West (1953) took language structure as the major criterion to classify vocabulary. He classified the words broadly into two major categories, with some sub-categories under each major category.

1. Form words:
 - (a) Essential words
 - (b) General words

2. Content words:

- (a) Common environmental words
- (b) Specific words

Form words are words that we speak with, and are liable to be used in any discourse on any subject. They make up the structure of the language. According to West, the essential words are nearly 150 in number and these are pronouns, conjunctions, etc. The general words are also structural in function but they are more refined, more limited in meaning and usage; and the less frequent general words are more stylistic.

Content words are the words that speakers use to talk about objects, events, etc. They constitute the substance of their talk, whereas common environmental words are words connected with things that are a part of the environment of all men.

Specific words are words that are peculiar to one group of men. The important characteristics of content words that as identified by West (1953) are as follows:

1. Most of the content words are simple in structure. They have only one or a few related meanings.
2. They are innumerable.
3. They are unpredictable.

Nisbet (1960) agreed with West, and followed the same vocabulary classification. He suggested 300 words as essential words. Words like *say*, which is a general word for

reply, ask, declare, but *food* and *eat* are common environmental words. *Chalk, pencil*, etc., are specific words.

Some other characteristics of form words are as follows:

1. They can be comparatively more easily omitted in telegrams, and now with the new technology in text messages and e-mails.
2. They are not an open set because easy addition to this group is not possible and also it is difficult to replace them by another alternative form by creation.

The characteristics of content words are as follows:

1. They cannot be easily omitted in telegrams, text messages and e-mails.
2. They are an open set because easy addition to this group is possible and also it is possible to replace them by another alternative form by coinage or borrowing.

5.4.1.7. Concrete and Abstract Vocabulary

The concrete vocabulary represents the concrete entity, the entity that can be physically perceived through the sense either in the form of objects or in the form of events. In turn, vocabulary items representing the entities like *love, soul*, and *fear* are called abstract vocabulary because the entities represented by these vocabulary items cannot be easily and physically perceived but need to be imagined mentally. They are only psychological realities and not the physical entities.

5.4.1.8. Basic Vocabulary

Basic vocabulary is more widely talked about than any other type of vocabulary. Basic vocabulary is assumed to be the core of the total vocabulary of a language. That is, it consists of all the vocabulary items that are found in daily use in listening, speaking, reading, and writing contexts of the speakers of that language. It excludes all other vocabulary items that are not found in their everyday language activity. Thus, the most frequent words that occur in the daily language activities are the basic vocabulary of the language. Further, it is assumed that the basic vocabulary consists only of the root words and not the derivatives.

The underlying assumption is that if one is taught the basic vocabulary and the rules of manipulation of these vocabulary items in appropriate contexts, both linguistic and social, a learner would be capable of operating with a much wider vocabulary range than covered by the basic vocabulary. The question of the utility of basic vocabulary arises in the context of teaching a language as a second language, and also in the context of teaching a language as first language in the adult education programmes, because in these cases much has to be achieved within the shortest duration.

In the case of basic vocabulary research also, while one scholar considers a particular vocabulary item as an item of basic vocabulary, another scholar may consider the same item as a non-basic vocabulary item. This is due to differences in their approach to the study of the vocabulary of languages. More details are provided in Chapter Seven.

5.4.1.9. Graded Vocabulary

The graded vocabularies are those that are graded in accordance with age or school grade of the pupil. Wherever possible, the lists of graded vocabularies try to accommodate both these variables, merge them, and form a single criterion to classify the vocabulary items of the language.

Though there is not any gradation inherent in the vocabulary of the language, it is possible to classify the vocabulary using these criteria.

This type of vocabulary classification has great pedagogical implications and benefits. It is expected that the vocabulary items used by the children of a particular school grade should find place in the teaching, learning, and reading material that is meant for that particular age and grade. These vocabulary items can also be used in the construction of aptitude tests, and other vocabulary tests intended to measure verbal intelligence. It is also true that differences in gradation may be found even when two researchers have employed the same set of criteria variables.

5.4.1.10. Recall Vocabulary

The concept of recall vocabulary serves pedagogical purposes well, when recall and the content to recall are decided and defined on the basis of the contexts in which recall vocabulary are used. However, the items that constitute a recall vocabulary set may differ from one individual to another.

Let us consider the following situation (Mallikarjun, 2002) in which a language is to be taught to an individual or to a group of people as second language. The objective

in such a situation being to train these people to interact in situations 'a, b, c...' using second/foreign language. The native speakers of that language are asked to recall all the vocabulary items they prefer to make use of in the situations 'a, b, c...' The vocabulary items collected in this way form the recall vocabulary of the language 'x' for 'a, b, c...' situations. Such vocabulary is then used in the construction of instructional material intended for specific groups of people to handle specific situations. Another way to draw the list of recall vocabulary items is to list as many semantic categories (like animals, vegetables, etc.,) as possible and ask the native speakers of that language to recall the vocabulary items under each of the semantic categories given to them.

5.4.1.11. Dialect Vocabulary

Vocabulary items that are identified with a particular socio-economic group of speakers of a language, or identified with the speakers in a particular geographic region, are called *dialect vocabulary* and these vocabulary items are normally absent in the standard language. These can be found in the regional literature produced by writers belonging to a particular class or region. In the context of the present study the dialect vocabulary is irrelevant as Algerian textbooks are supposed to deal only with Standard English.

5.4.1.12. Archaic Vocabulary

Due to many reasons, linguistic and non-linguistic reasons, vocabulary items may fall out of use from the current spoken and written language and thus become obsolete.

Such vocabulary items that fall into disuse are called archaic vocabulary. These vocabulary items are naturally found in old written literature and also in folk literature. Again, to relate this to the context of the present study the archaic words are irrelevant as Algerian textbooks are supposed to deal only with modern or contemporary English.

5.4.1.13. Technical Vocabulary

Technical vocabulary is a special vocabulary. A technical vocabulary item belonging to a particular technical subject or technical context indicates specific meanings, the same item with the same physical features. But, in the context of another technical subject or technical context, it may indicate different specific meanings. At times, a vocabulary item used as a technical term may indicate one meaning in the technical content and another different meaning in the language of daily life. By way of example, the common word table has the following meanings among many others (*The American Heritage Dictionary*):

1. An article of furniture supported by one or more vertical legs and having a flat horizontal surface.
2. *Anatomy*. The inner or outer flat layer of bones of the skull separated by the diploe.
3. *Architecture*. A raised or sunken rectangular panel on a wall.
4. *Music*. The front part of the body of a stringed instrument.
5. *Sport* (also league table). Ranking.
6. A system of laws or decrees; a code: the tables of Moses.

5.5. What constitutes word knowledge?

A great deal has been written on the topic of what it means to “know” a word. According to Beck and McKeown (1991): “People’s vocabulary knowledge is called incremental: knowledge of a word is to be seen as a continuum from “not knowing” to rich knowledge of a word’s meaning, its relationship to other words, and its extension to metaphorical uses”. Vocabulary knowledge in the mother tongue as well as in a foreign language continues to deepen throughout lifetime: as we grow older, we continue to learn nuances and subtle distinctions conveyed by words. Much of what has been written on word knowledge goes back to the vocabulary knowledge framework of Richards (1976). Richards identified seven aspects of word knowledge. In his view, “knowing a word” means:

- a) Knowing the degree of probability of encountering the word in speech or print;
- b) Knowing the limitations imposed on the use of the word according to function and situation;
- c) Knowing the syntactic behaviour associated with the word;
- d) Knowing the underlying form of a word and the derivations that can be made of it;
- e) Knowing the associations between the word and other words in the language;
- f) Knowing the semantic value of the word, and
- g) Knowing many of the different meanings associated with the word.

Put differently and in more pragmatic way, knowing a word means:

- a) To understand it when it is written and/or spoken.

- b) To recall it when you need it.
- c) To use it with the correct meaning.
 - d) To use it in a grammatically correct way.
 - e) To pronounce it correctly (i.e. with an acceptable pronunciation).
 - f) To know with which other words you can use it.
 - g) To spell it correctly.
 - h) To use it in the right situation (i.e. appropriate).
 - i) To know if it has positive or negative association (i.e. connotations).

Applied linguists seem to agree that the same continuous idea of incremental expansion of vocabulary knowledge also applies to the transfer from receptive to productive mastery. The learning of a word is thought to progress from receptive to productive knowledge. This means that a word that can be correctly used is assumed to be understood by the user, when heard or seen. The opposite however, is not necessarily true. Passive vocabulary size is thus considered to be larger than the active size even though it is not clear how much larger it is (this point is treated in the next chapter). In Nation's (1990) framework for vocabulary knowledge, he therefore distinguishes eight types of word knowledge that are specified both for receptive and productive knowledge. According to Nation (2000), there are nine aspects of knowing a word that include form, meaning and uses:

1. Spoken form;
2. Written form;

3. Word parts;
4. Connection of form and meaning;
5. Conceptual meaning;
6. Association with related words;
7. Grammatical functions;
8. Collocation behaviours;
9. Word usage constraints; appropriateness.

Nation relates these nine aspects to the receptive and productive knowledge of words. Receptive knowledge is important for recognizing the meaning of a word without the need to produce the word again. On the other hand, productive knowledge is the knowledge that learners need to “produce language forms by speaking and writing to convey a message to others” (p. 24). Also, Ruddell (1994) divided knowing a word into five categories:

1. Knowing the word meaning aurally
2. Knowing the word meaning but not expressing it
3. Knowing the meaning but not the word
4. Knowing the partial meaning of the word
5. Knowing a different meaning of a word

Thus, one’s knowledge of a word does not have to include both receptive and productive control to perform certain tasks. Schmitt & Meara (1997) argue that native speakers do not master all types of word knowledge. They only master a limited

number of word knowledge categories for most of their lexicon, and have only the receptive knowledge of some low frequency words. Second language learners, like native speakers, do not have to have a full knowledge of all vocabulary to function in the language. This knowledge varies depending on the task the person is performing (Qian, 2002). Some activities involve only receptive knowledge while others require productive knowledge as well. In reading tasks, only receptive knowledge is required. Second/foreign language learners do not need full knowledge of a word meaning for every reading comprehension task, partial knowledge may suffice. Through this partial or incomplete knowledge learners can gain additional knowledge about a word, as suggested by Henriksen (1999) when he proposed three components in vocabulary development:

1. Partial to precise knowledge;
2. Depth of knowledge;
3. Receptive to productive ability.

As indicated, second/foreign language readers need to develop both receptive and productive vocabulary knowledge and to increase their vocabulary size. “When readers increase their vocabulary size, their use of language skills implicitly increases and their knowledge of the world also becomes broader” (Huang, 1999, p. 43). A larger vocabulary enhances other language skills of second/foreign language learners. Language learners must acquire as much vocabulary as possible in order to effectively

read in the language (Bernhardt & Kamil, 1995). graves et al. (2001) listed five stages of vocabulary knowledge:

1. Learning to read a known word;
2. Learning new meanings of known words;
3. Learning new words that represent known concepts.
4. Clarifying and enriching meaning of known words;
5. Moving words from receptive (listening and reading) to expressive (speaking and writing) vocabulary (p. 81).

Conclusion

From the above, it can be concluded that although *word* and *vocabulary* may appear simple concepts to the layman, it is not so from the linguist's point of view as different criteria have been used to identify *word*. Defining *word* from three different perspectives (semantics, phonology, and orthography) bears witness on the difficulty and disagreements over working out a universal definition of *word*. Various definitions have also been put forth and there exists different types of vocabulary. The distinction between these types has proved useful in second/foreign language pedagogy. What constitutes word knowledge has also been a discussible issue. However, the positions considered above agree on a certain minimal number of aspects of knowing a word which are believed to constitute the minimal ultimate objective of English teaching in Algeria at least:

1. Knowing the meaning of the word, be it total or partial, preferably total;

2. Knowing how to use the word appropriately;
3. Knowing how to pronounce the word in speaking;
4. Knowing how to spell the word in writing; and
5. Understanding the meaning of the word in listening.

As for teaching and learning vocabulary, depending on the aim assigned to the language learning syllabus, one type of vocabulary or another may be favoured over another such as in cases where English is taught for specific purposes. Technical English will take the bulk of vocabulary for a technology-oriented textbook, for example. Academic vocabulary is obviously useful for any learner pursuing academic studies but certainly not for someone learning English with a different goal in mind. It should be mentioned that this classification cannot be a rigid one as many words can belong to more than one category as shown above with examples of words which have a certain meaning in a certain field and a different meaning in another field. Except may be in the case where English is the object of study as in philology, archaic and dialect vocabulary are not included in textbooks.

On the whole, the above considerations strengthen the argument that there is a relationship between vocabulary and reading comprehension. A pertinent issue related to vocabulary is lexical coverage. Such issues will be tackled in the next chapter.

CHAPTER 6

LEXICAL COVERAGE AND READING COMPREHENSION

Introduction

This chapter provides an overview of research concerning vocabulary learning and comprehension on the one hand, and vocabulary learning and lexical text coverage, or simply lexical coverage, on the other. It examines the issues of how some aspects of vocabulary might affect the stability of text coverage. It has been stressed throughout this work that the importance of vocabulary in second/foreign language learning goes uncontested and that vocabulary is indispensable for successful communication.

The first section of this review is about the relationship between vocabulary and comprehension. It will discuss both theoretical and empirical basis for the connection between understanding the meaning of individual words and understanding text.

The second section of this review starts with a definition of a few useful terms then it moves on to the central issue in this work, lexical coverage. A brief historical overview of the evolution towards recognition of the importance of lexical competence within second/foreign language learning will be briefly sketched. Finally the bulk of this chapter is devoted to the review of the literature on lexical coverage.

6.1. Reading comprehension and vocabulary knowledge

According to Hunt and Belgar (2005,) ‘the heart of language comprehension and use is the lexicon’. Widdowson (1989, p.136) called ‘to shift grammar from its pre-eminence and to allow the rightful claims of lexis to arrive on the working agendas of researchers’. The relationship between vocabulary knowledge and comprehension has a

long tradition of study in the field of reading research (for example, Davis, 1944, 1968; Thorndike, 1973). Both vocabulary knowledge and reading comprehension have been shown to be closely related, and this relationship is not one-directional, since vocabulary knowledge can help the learner to comprehend written texts and reading can contribute to vocabulary growth (Chall, 1987; Nation, 2001; Stahl, 1990). Numerous studies have shown a strong correlation between vocabulary and comprehension (Davis, 1944; 1972; Farr, 1969; Harrison, 1980; Stahl & Fairbanks, 1986). Although the relationship appears to be simplistic in nature or even a truism—the more words know the easier it is to understand a passage, the interaction is actually complex. Consider the large number of factors discussed in Chapter 4 (see section 4.7.) which bear witness of such complexity. In addition, factors involved in “knowing” a word (see Chapter 5, section 5.5) reveal a complex relationship between comprehension and vocabulary leading the RAND Reading Study Group to conclude that: “the relationship between vocabulary knowledge and comprehension is extremely complex, confounded as it is by the complexity of relationships among vocabulary knowledge, conceptual and cultural knowledge, and instructional opportunities (RAND, 2002, p. 35). There is little doubt then that the relationship exists, the question for researchers has been to determine the nature of the relationship, exploring the extent to which vocabulary knowledge and reading comprehension affect each other. Questions regarding the nature of the relationship have been explored through research offering empirical evidence for the

relationship and theoretical suggestions describing the interaction between vocabulary and comprehension.

6.1.1. Empirical Evidence

As stated, empirical evidence exists indicating that vocabulary knowledge may impact comprehension (Stahl & Fairbanks, 1986; National Reading Panel, 2000).

Carver (1994) showed that the relative difficulty of the text is a function of the percentage of unknown vocabulary words. Following this reasoning, teaching unknown words prior to reading would directly assist the reader in reducing the relative difficulty of a text, thus enhancing comprehension. Other researchers, such as Sternberg (1987) asserted that ability to comprehend a text can be predicted based on vocabulary knowledge. The greater the store of vocabulary and understanding of words and concepts the more likely one is to comprehend. Therefore, increasing an individual's vocabulary knowledge should directly increase their ability to comprehend. Sternberg stated the relationship through his assertion that the level of vocabulary knowledge of a reader may determine their level of comprehension.

Studies have shown that instruction in vocabulary may affect comprehension of text passages that include the vocabulary terms (Mezynski, 1983; Graves, 1986; Stahl & Fairbanks, 1986). In a review of several studies exploring the relationship between vocabulary learning and comprehension, Graves (1986) cited studies by Beck, Perfetti, & McKeown (1982); McKeown, et al. (1983); and McKeown, (1985) as particularly

convincing in establishing evidence between teaching new vocabulary and comprehension.

6.1.2. Theoretical Perspectives

A variety of hypotheses have been put forth to explain the relationship between vocabulary and comprehension (Anderson & Freebody, 1981; Mezynski, 1983; and Ruddell, 1994). A review of the literature on this relationship which focused on why such a strong correlation exists between the two areas reveals four basic assertions regarding this relationship. One is that vocabulary learning and comprehension have a reciprocal relationship. Two, that some factors, such as background knowledge are the causal link between vocabulary knowledge and comprehension ability. A third view is that vocabulary knowledge is the result of comprehension. The last view, which is the opposite of the previous one, holds that comprehension is the result of vocabulary knowledge.

Researchers have suggested several models to describe the relationship between vocabulary knowledge and reading comprehension. These are formulated in the form of the following hypotheses: *the Instrumentalist Hypothesis*, *the Knowledge Hypothesis*, *the Aptitude Hypothesis* Anderson and Freebody (1981), and *The Access Hypothesis* Mezynski (1983).

6.1.2.1. Instrumentalist Hypothesis

According to the instrumentalist hypothesis there is a causal connection between vocabulary size and reading comprehension, i.e. knowing more words makes a learner a better reader (Anderson & Freebody, 1981). For showing the validity of this hypothesis it suffices to imagine a reader who has encountered a text replete with unknown words. Here, the reader's vocabulary knowledge can be a strong predictor of his or her success in comprehending the text. A number of studies have shown that teaching words can improve reading comprehension (Beck & McKeown, 1991; Stahl & Fairbanks, 1986). In sum, this view holds that comprehension is the result of vocabulary knowledge and posits a causal relationship between vocabulary and comprehension, that is the extent of one's knowledge of word meanings directly affects how much is understood. In other words, knowledge of individual words encountered while reading a text is the necessary and prerequisite condition needed in order to understand a text as a whole. A natural consequence then of learning new vocabulary would be increases in comprehension. Hence, because vocabulary controls comprehension, to improve understanding it is necessary to increase the number of word meanings that are known. In a nutshell, the instrumentalist view sees that good vocabulary knowledge enables good comprehension.

6.1.2.2. Knowledge Hypothesis

The second hypothesis advanced by Anderson and Freebody (1981), as an alternative to instrumentalist hypothesis, is knowledge hypothesis. This hypothesis

intimates that it is along with world knowledge that word knowledge ameliorates comprehension. Connecting words and students' prior knowledge and experiences is one of the salient characteristics of an effective vocabulary instruction (Stall, 1986). This prior knowledge enhances reading comprehension because the reader must bring as much information to the text as the reader expects to get from it. The knowledge hypothesis emphasizes the role of vocabulary knowledge within the framework of a schema theoretical view of reading comprehension discussed in Chapter 2. The more developed one's background knowledge, the easier it is to integrate new concepts (i.e. words) into existing schema. Thus schema, or background knowledge, of a particular concept enhances vocabulary knowledge. It is difficult to read about a topic, say astrophysics, if you know nothing about it.

Nagy and Herman (1987) offer support for the knowledge hypothesis by asserting that the correlation between vocabulary and comprehension is just really a result of individual background knowledge. This hypothesis holds that the more background knowledge one has about a topic the greater their comprehension of the text, and therefore also the greater their ability to learn new vocabulary related to the text. Thus, the greater the background knowledge, the more likely vocabulary can impact comprehension (Nichols, 2007). Worded differently, as students learn new words their background knowledge grows, thus allowing them to read more difficult passages.

6.1.2.3. Aptitude Hypothesis

The Aptitude Hypothesis sees vocabulary knowledge as dependent among other factors on cognitive abilities especially intelligence. Other abilities might include the ability to understand oral explanation.

6.1.2.4. Access Hypothesis

According to Mezynski (1983), the gist of access hypothesis is that, to make best use of word knowledge in reading comprehension, students should have a quick and easy access to the words they already know. Put simply, both depth and breadth of word knowledge are involved in reading comprehension. In fact, besides knowing more words, students should be able to come up with the precise meanings of words quickly. The pedagogical implication of this hypothesis is that words (at least some special words) should be taught comprehensively. McKeown, et al. (1985) suggested that students need to encounter a word as many as 12 times before they know it well enough to ameliorate their comprehension. In other words, for bolstering students' understanding of a set of words, multiple exposures to those words in different texts should be highly considered. As Anderson and Freebody (1981) indicated, these hypotheses are not mutually exclusive and all are probably to be at least part of the truth.

The access view of the relationship between vocabulary knowledge and reading comprehension, like the instrumentalist view, sees vocabulary as having a causal relationship with comprehension provided that the vocabulary can be easily accessed.

Access can be improved through practice. This access can involve several factors including fluency of lexical access, speed of coping with affixed forms, and speed of word recognition. For EFL/ESL learners, the relationship between vocabulary knowledge and reading comprehension are even more complicated. These complications can arise from the learners being able to read in their first language and the common situation of beginning to read the target language with virtually no vocabulary knowledge in the target language. That indicates for EFL/ESL learners the vocabulary knowledge plays a more important role in reading comprehension.

The issue to address next, then, is the relationship between reading comprehension and vocabulary knowledge.

6.2. Lexical Coverage and Reading Comprehension

Before surveying the literature on lexical threshold, it is necessary to clarify the terminology whose understanding is essential for the foregoing discussion of reading and vocabulary. The terms involved are: *word frequency/list*, *sight vocabulary*, *type/token*, *lexical density*, and *word family*.

6.2.1. Key terms and concepts

6.2.1.1. Word frequency/list

This term refers to the frequency with which a word is used in a text or corpus. One of the principles by which learner vocabulary is often organized is that of word frequency. It stands to reason that there is a relatively small number of words that are

used much more often than others. Apart from the function words, articles, modals, prepositions and so on — there are certain content words that occur with much greater frequency than others. Carter (1987) discusses the idea of a core vocabulary for English; that is to say, the notion. Thus based on the idea of a definable set of words that could stand as an attainable goal for language learners, researchers have attempted to make lists, called frequency lists, of the basic and most important words in a language generally intended for use as a basis for language teaching or for the preparation of teaching materials. Lists such as Thorndike and Lorge's *Teacher's Word Book* (1944), and West's *General Service List* (1953, see Appendix 2) were basically compiled for that purpose and have provided the basis for numerous studies and other more recent word lists (for example, Quirk, 1982; Nation's 1996 Vocabulary Lists). More details about these lists and others are provided in the next sections.

6.2.1.2. Sight vocabulary

The term “sight vocabulary” is used to refer to ‘those words which a child can recognize at sight in a reading passage or text and which he or she does not need to decode using phonic or other reading skills’ (*The Longman Dictionary of Language Teaching and Applied Linguistics* 2002, Third Edition). In second/foreign language learning a quite similar definition has been provided by Laufer (2010). This researcher defines ‘sight vocabulary as:

The term “sight vocabulary” is used to refer to words whose meaning is so familiar to a person that they can be understood out of context.

Therefore, when encountered in a text, these words are recognized and decoded quickly and without any cognitive effort. (p. 16)

For example, if a reader encounters the word ‘dissertation’ in a text, and if this word is part of his/her sight vocabulary, there is no need to have recourse to the surrounding context to understand its meaning. Consequently, having a large sight vocabulary enhances to a great extent reading fluency and frees cognitive effort for higher level reading processes that is engaging with comprehending the text content and its implications (Mezynski, 1983; Segalowitz, 2007).

6.2.1.3. Type/token

Simply put, tokens or running words are the total number of individual words occurring in a text. In linguistics, a distinction is sometimes made between classes of linguistic items (e.g. phonemes, words, utterances) and actual occurrences in speech or writing of examples of such classes. The class of linguistic units is called a type and examples or individual members of the class are called tokens. For example, *hello*, *hi*, *good morning* are three different tokens of the type “Greeting”.

In mathematical linguistics the total number of words in a text may be referred to as the number of *text tokens*, and the number of different words as the number of *text types*. Text tokens are orthographic words as defined in Chapter 5 under 5.1.1. that is, a linguistic item preceded and followed by a space or a sequence of letters that we write consecutively, with no spaces. For example, in the sentences ‘the person wearing

sunglasses and the person wearing eyeglasses professors' there are 10 tokens and 7 types.

6.2.1.4. Lexical density

Lexical density (also Type-Token Ratio, concept load) is the measure of the ratio of different separate words (type) to the total number of words in a text (tokens). It is used as a measure of the difficulty of a passage or text. Lexical density is normally expressed as a percentage and is calculated by the formula:

$$\text{lexical density} = \frac{\text{number of separate words}}{\text{total number of words in a text}} \times 100$$

One of the uses of Type-Token Ratio has been in corpus linguistics to measure the proportions of frequent and infrequent words. A number of widely available computer programmes dedicated to linguistic analysis, such as Systematic Analysis of Language Transcripts (SALT) (Miller and Chapman, 1993), the Oxford Concordance Program (Hockey, 1988) and the Computerized Language Analysis (CLAN) programs (MacWhinney, 1995; MacWhinney and Snow, 1990) have relied on Type-Token Ratio in their analysis.

6.2.1.5. Word family

A word family, as has already been defined earlier in this dissertation (see Chapter 5, section 5.2.) consists of a base word and its inflected forms and derivations (Nation, 2001, p. 8). For example, *AID* has the following family members *AIDED*, *AIDING*, *AIDS*, and *UNAIDED*.

Having provided such clarifications, and since lexical coverage is the focus of the dissertation, we will now first define what is meant by lexical coverage and then survey studies that investigated it, focusing on the interaction between coverage, and reading comprehension that was considered adequate in each study.

6.3. What is text lexical coverage?

Since The Threshold Hypothesis (Ulijn & Salager-Meyer, 1998, p. 81), has been widely accepted as discussed in Chapter 4 (section 4.6.1..), and since the importance of vocabulary has been widely acknowledged and has recently been a particular focus in the field of reading comprehension as evidenced by the body of studies that approached the issue (Appelt, Ken. 2006, Beglar, 2009; Campbell & Weir. 2006 ; Crossley et al., 2008; Crossley & McNamara, 2008; Davis, 1972; Hirsh & Nation, 1992; Hu & Nation, 2000; Huckin and Bloch, 1993; Klare, 1974-75; Laufer, 1985, 1989a, 1992c, 1997, 1999, 2001, 2010; Laufer & Nation, 2001; Laufer & Waldman, 2011; Nation, 1994, 2001, 2006, 1990, 2009; Nation & Waring, 1997) there has been continuing interest in whether there is a language knowledge threshold which marks the boundary between having and not having sufficient language knowledge for successful language use (Bensoussan and Laufer, 1984; Holley, 1973; Hu and Nation, 2000; Laufer 2001 , 2010; Nation, 2001), or how much unknown vocabulary can be tolerated in a text before it interferes with comprehension? Determining the amount of unknown vocabulary has been considered crucial in order to know if it enables the foreign language learner to understand what he/she reads, that is to say determining the lexical threshold.

One approach to address this issue has been to research how the amount of vocabulary a reader knows affects reading comprehension. Put in the form of a question, this yields the following: How many words in a text must a reader know in order to understand what is being read? “The percentage of running words in the text known by the readers” (Nation, 2006, p. 61) is referred to as *lexical or text coverage*. Technically, it is calculated as “the number of the words known in a text, multiplied by 100 and then divided by the total number of running words, i.e. tokens in the text” (Nation, 2001, p. 145). This can be put in the form of the following formula:

$$\text{lexical coverage} = \frac{\text{number of known words in a text}}{\text{tokens in a text}} \times 100$$

The assumption made behind lexical coverage is that there is a lexical knowledge threshold which marks the boundary between having and not having sufficient vocabulary knowledge for adequate reading comprehension.

6.4. How has text coverage been assessed?

Three ways have been used to assess text coverage, that is the percentage of words likely to be unknown. The first way is teacher judgment of text, based on knowledge of students and of the syllabus and what vocabulary they are likely not to know. Second, researchers have relied on computer software to check text against general frequency or a specific syllabus wordlist. The third way is relying on learner judgment as to whether a text is at the right level for them? Hu & Nation (2000). More details are provided in the next chapter. Reviewing the literature on text coverage,

Nation (2000) asked and attempted to answer a certain number of questions in relation to text coverage. These questions are considered below as outlined by this researcher.

6.5. Which words should second/foreign language learners know?

The current trend in language pedagogy is that teaching of vocabulary is crucial and needs to be structured. Such structuring needs to be done on the basis of word frequency and text coverage (Meara, 1993). Consequently, it seems evident that the more frequent words are most useful and should be taught first, before spending time on less frequent words or words that only occur in specialized domains. Nation (1990, 2001) reports that frequency-based studies have shown that a small group of very frequent words cover a very large proportion of the running words in any spoken or written text and occur in all kinds of uses of language. Nation (2001) divided vocabulary into four categories: (1) high-frequency or general service vocabulary, (2) academic vocabulary, (3) technical vocabulary and (4) low-frequency vocabulary. High-frequency words refer to those basic general service English words which constitute the majority of all the running words in all types of writing. The most well-known general service vocabulary is West's (1953) General Service List of English Words (GSL). The GSL containing the most frequently-occurring 2,000 word families of English (3,372 word types) accounts for approximately 75% of the running words in non-fiction texts (Hwang, 1989) and around 90% of the running words in fiction (Hirsh, 1993). Academic vocabulary, also called sub-technical vocabulary, or semi-technical vocabulary, is a class of words between technical and non-technical words and usually

with technical and non-technical implications. Technical words are the ones used in a specialized field and are considerably different from subject to subject. About 5% of the words in an academic text are made up of technical vocabulary, with each subject containing roughly 1,000 word families (Nation, 2001).

Nation and Waring (1997) pointed out that the beginners of English learning should focus on the first 2,000 most frequently-occurring word families of English in the GSL, while for intermediate or advanced learners who usually study English for academic purposes, the command of the top 2,000 frequent words may no longer be their concern and the priority of their vocabulary learning may be shifted to the next level of vocabulary, i.e. sub-technical/academic vocabulary. In academic settings, ESP students do not see technical terms as a problem because these terms are usually the focus in the specialist textbooks. Low-frequency words are rarely used terms. Academic vocabulary with medium-frequency of occurrence across texts of various disciplines (somewhere between the high-frequency words and technical words) has some rhetorical functions. Acquiring these sub-technical words seems to be essential when learners are preparing for English for Academic Purposes. Alternatively, vocabulary based on Nation's (2001) four divisions can be learned in a systematic order. Students should learn first the 2,000 general words of English, followed by a set of academic words common to all academic disciplines. In line with Nation and Waring (1997), Coxhead (2000) compiled a corpus of around 3.5 million running words from university textbooks and materials from four different academic areas (law, arts and commerce as

well as science), and identified 570 academic word families (AWL), which were claimed to cover almost 10% of the total words in a general academic text. This research suggested that for learners with academic goals, the academic word list contains the next set of vocabulary to learn after the top 2,000-word level. To put it concretely, greater lexical coverage is gained by moving on to learning 570 academic words (10% coverage) than by continuing to learn the next 1,000 words (“3–5%” coverage for the 3rd 1,000, Nation, 2006, p. 79) after the top 2,000 word families on a frequency list.

Given the role played by high-frequency words in vocabulary learning the question arises if this group of words within a language is stable. According to Nation (2001, 15-16), ‘frequency lists may differ in frequency rank order of particular words but there generally is 80% agreement about what words should be included in the list, provided that the corpus has been well-designed’. With reference to word counts, Nation (2001) holds that knowing a word involves knowing the members of its word family and the number of members of the word family will increase as proficiency develops. A learner may be familiar with the word “rich”, “richly” and “richness” in an early stage and expand this word family with “to enrich” and “enrichment” in due time. Nation (2001) suggests that rather than talking about “knowing a word”, we should be talking about “knowing a word family” (p. 47). A frequency-based approach to vocabulary learning hinges upon the assumption that frequency is strongly related to the probability that a word will be known. According to Anderson and Freebody (1981)

this hypothesis is supported by evidence from a number of L1 areas. Hazenberg and Hulstijn (1996) researched the relationship between word frequency and word knowledge. They attempted to find out to which extent word frequency can be used to predict word knowledge. It may be that the most frequent words are known by all students, whereas more infrequent words are known only by particular individuals, depending on variables such as hobbies, work and experiences, culture, age, and other factors. They concluded that the relationship between word frequency and word knowledge appears to depend on vocabulary size (or breadth). When individuals have a relatively large vocabulary there is no significant relationship. But when individuals have a relatively small vocabulary, word frequency can be used as a criterion to predict word knowledge.

6.6. How many words should second/foreign language learners know?

From a pedagogical perspective it is useful to know how much vocabulary is needed before learners reach the vocabulary threshold level which is necessary for the comprehension of written texts. Thus, a major concern in reading comprehension studies has been the issue of the number of words a reader should know in order to read effectively in a foreign language. Researchers distinguish *breadth* or *size* of knowledge (the number of words of which the learner knows at least some significant aspects of the meaning) from depth of knowledge, with which they refer to the quality of vocabulary knowledge, namely how well a particular word is known. Although both measures are considered important a lot of work on vocabulary testing has focused on

vocabulary size. Meara (1996) asserts that the basic dimension of lexical competence is size. He states that:

All other things being equal, learners with big vocabularies are more proficient in a wide range of language skills than learners with smaller vocabularies, and there is some evidence to support the view that vocabulary skills make a significant contribution to almost all aspects of L2 proficiency. (p. 37)

Reviewing the literature on vocabulary studies, Mallikarjun (2002) noted that vocabulary has been studied following three underlying principles: the first principle is that the frequency of a vocabulary item in a language gives the quantum of use to which the speakers of the language put the vocabulary. The second principle is that only nearly 3000 words make up 95% of the total vocabulary of the individual speaker, and that words like I, the, and, to, a, you, of, in, we, and for constitute 25% of the writing vocabulary, and that only 100 words form 60% of the words of an individual speaker. The third principle is that in the teaching and learning context, most frequent words should be utilized. Aiming at collecting words that should be of use to the American children learning to read English, Thorndike (1921) collected 10,000 most frequent English words. These 10,000 words were found to be frequent in 41 sources used for the study. Nearly 625,000 running words of the children's literature, nearly 3,000,000 running words of the Bible and English classics, nearly 300,000 running words of the Elementary school text-books, nearly 50,000 running words of books about cooking,

farming, sewing, trade, etc., nearly 90,000 running words of daily news papers, and the rest 500,000 running words from correspondences were used as source running words for this 10,000 words word count.

Thorndike (1931) revised the 1921 list and expanded it to 20,000 words relying on 238 sources with nearly 500,000 running words. Again in 1944, Thorndike, in collaboration with Large, revised and expanded the list to 30,000 most frequent words.

Thorndike made a distinction between the *range* of a word and the *frequency* of a word. The *range* is the occurrence of the word in different sources, and *frequency* is the occurrence of the word in the total number of running words. Thorndike considered the word with *wide range* as more important than the word with *wide frequency*.

Other researchers focused on word count. Similarly, Horn (1926) identified 25,000 words as basic writing vocabulary. He collected these words from personal and business letters through word count. Ogden (1932) collected 850 basic words needed for expressing one's ideas and desires. He obtained these after eliminating all synonyms and non-essential verbs. Out of these 850 basic words, 6 are nouns, 150 are adjectives, 18 are verbs and others are prepositions, pronouns, adverbs and conjunctions. According to Ogden, it is possible for one to express himself adequately by deriving different words from these basic words, and by using the same word in different parts of speech with different meanings. He claimed that these 850 words do all the work of 20,000 words.

On the basis of Thorndike's and Horn's list, Faucett and Maki (1932, cited in Gilner, 2011) prepared a vocabulary list. They grouped the vocabulary into four categories:

- i) Indispensable words (about 360) for four-fold mastery; i.e., for understanding, reading, speaking and writing, oral mastery being important.
- ii) Essential words (about 1,198) for three-fold mastery; i.e., for understanding, reading and speaking them, oral mastery being attempted only if time permits.
- iii) Useful words for two-fold mastery, i.e., reading and understanding only.
- iv) Special words for reading mastery only.

Another pioneer of vocabulary studies, Palmer (1930) developed a system called 'head-word' system. He devised this as a help to arrive at the core vocabulary with the help of which one could learn a language. That is, one learns a finite set of words with which an infinite number of words can be derived and used. According to Goulden, Nation and Read (1990), a well-educated adult native speaker of English has a vocabulary of around 17,000 words. This large number of English words, however, is a learning goal far beyond the reaches of foreign language learners like ours.

In relation to second/foreign language learning, there has been a research interest in defining the vocabulary size required for second language reading. The studies looking at the relationship between vocabulary size and reading comprehension support the threshold hypothesis, which, it should be recalled from Chapter four (see under 4.6.2.), that there is a threshold of vocabulary necessary for comprehending written

texts (Alderson, 1984). The size of a student's vocabulary has been found to correlate closely with reading comprehension (Laufer, 1992c; Beglar, 1995; Qian, 1999) as well as with writing ability (Astika, 1993; Laufer, 1998; Laufer and Nation, 1995).

Laufer (1992c) has suggested that a vocabulary of 3000 word families of general English is enough for a good understanding of a general English text such as a novel. Hirsh & Nation (1992) estimate that for pleasure reading knowledge of 5000 word families is necessary. According to Waring and Nation (1992), the number of words needed for the reading of technical texts such as science texts, or newspapers is larger than for less formal texts. One reason, according to these authors, is that there is a high proportion of technical words. Chung & Nation (2003) found that 38% of the running words in an anatomy text and 17% of the words in an applied linguistics text were technical words.

Nation and Hwang (1995, cited in Matsuoka & Hirsh, 2010) showed that knowledge of the 2,000 most frequent word families enables SL/FL readers to recognize 84% of the words in various types of authentic texts. Laufer (1997) suggested that the 95% lexical threshold in L2 reading would be required to enable L2 readers to apply their L1 reading strategies.

While research suggests that comprehension of texts may be achieved with a vocabulary size of 3,000 word families, a vocabulary size of 5,000 word families is needed to attain 98% lexical coverage of texts, allowing for more pleasurable reading and more accurate guessing of unknown words in their context (Huckin et al., 1993;

Hirsh & Nation, 1992; Laufer, 1997). Hu and Nation (2000) investigated the effect of four different text coverage levels (80%, 90%, 95%, and 100%) on unassisted reading comprehension of a fiction text and found that reading comprehension increased as density of known words increased. They found that while some readers required 95% lexical coverage for adequate comprehension, most required 98%. They also hold that EFL learners must have around 98% coverage of the words in the text to be able to read. They found that there was an expected relationship between level of unknown words and comprehension. They concluded that as the number of unknown words increases, comprehension falls.

6.7. Repetition and Vocabulary Learning

In considering the use of textbooks for vocabulary development, one important issue has been brought to light, that is repetition. Nation (20010) asserts that word repetition is a favourable condition in vocabulary learning.

Research on the effect of word repetition on vocabulary learning has focused on three variables: (1) the number of repetitions, (2) spacing of repetitions, and (3) types of repetitions.

Regarding the number of repetitions, according to Huckin & Coady (1999), no set number of repetitions of a word guarantees its learning. A research by the author of the present study (Torki, 2001; 2011) revealed that the frequency of occurrence of lexical item in classroom discourse does not guarantee intake. However, other research have shown that it is necessary to encounter a word in a variety of contexts a number of

times, at regular intervals, in order for the learner to have a realistic chance of learning the word (Nation, 2001, Schmitt, 2000). Many researchers (Nation & Wang, 1999; Saragi, Nation, & Meister, 1978; Webb, 2007) have suggested a target of 10 repetitions for learning of unknown words. Single long continuous texts such as novels (Hirsh & Nation, 1992), separate texts related through topic such as newspaper reports on the same event (Hwang & Nation, 1989), and separate expository texts on a similar narrow theme (Gardner, 2008) provide more favorable repetitions of unknown vocabulary than is the case in unrelated texts, and thus provide improved conditions for vocabulary acquisition. This is due to the effect of incremental acquisition of repeated vocabulary while reading (Matsuoka and Hirsh, 2010). An analysis of the textbook can show if target vocabulary occurs frequently enough and is given enough repetitions over time to provide optimum vocabulary-learning conditions. The results can guide teachers in deciding how best to supplement the text with activities that will give learners exposure to target vocabulary that is not sufficiently presented in the textbook

Greene (1992) notes that “The saying ‘practice makes perfect’ illustrates the accepted fact that repetition is critical for learning. All other things being equal, our memory for information will depend on the number of times that we have encountered or studied it” (p.132). This suggests that from the perspectives of vocabulary learning, the more times learners encounter vocabulary items, the better chance they have of learning them. According to Barcroft (2004), learners begin to acquire new words by having the words presented to them frequently and repetitively in the input. The author

also pointed out that vocabulary learned through recycled words fosters long-term learning rather than short-term retention. According to Nation (2001), vocabulary is not about knowing words, but also knowing them so well that they can be fluently used.

Regarding the second variable, spaced repetition, research in memory (Baddeley, 1990) and second language vocabulary learning (Bloom & Shuell, 1981; Dempster, 1987) has found that spaced repetition (i.e., the spreading out or ‘spacing’ of repetitions of a word throughout a text) is more conducive to the learning of vocabulary than repetition that is massed (i.e., the concentration of repetitions of a word in only one part of a text). Massed repetition refers to repeated attention to a word over a continuous period of time, say 6 minutes, while spaced repetition refers to giving the same amount of attention to a word over a longer period of time, such as 2 minutes on three occasions over a 2-week period. The study time of the word is 6 minutes in total in both cases, but the repetitions are ‘massed’ or ‘spaced’ depending on the approach.

6.8. What vocabulary does a language learner need?

The previous sections of this chapter have suggested there is vocabulary threshold that should be given high priority in second/foreign language learning, the vocabulary in question consists of high frequency words, and to quote again Nation & Waring (1997, p. 11) ‘...there is little sense in focusing on other vocabulary until these are well-learned’

In this section we look at some useful vocabulary lists based on frequency and review the research on the adequacy of the General Service List (West, 1953). Most

counts consider *range*, that is the occurrence of a word across several subsections of a corpus.

6.8.1. The Teachers Word Book

The Teachers Word Book of 30,000 words (Thorndike and Lorge, 1944): This list of 30,000 lemmas or about 13,000 word families (Goulden, Nation and Read, 1990) is based on a count of an 18,000,000 word written corpus. Its value lies in its size. It is based on a large corpus and contains a large number of words. However, it is old, based on counts done over seventy years ago.

6.8.2. The American Heritage Word Frequency Book

The American Heritage Word Frequency Book (Carroll, Davies and Richman, 1971) This comprehensive list is based on a corpus of 5,000,000 running words drawn from written texts used in United States schools over a range of grades and over a range of subject areas. The list focuses on school texts and the frequency of each word in each of the school grade levels and in each of the subject areas.

6.8.3. The General Service List (West, 1953)

This list was developed in the 1953 and contains 2000 headwords. The frequency figures for most items are based on a 5,000,000 word written corpus. Percentage figures are given for different meanings and parts of speech of the headword. Although, it is now about seventy years since it was developed and it is still widely used.

The 2000 word *General Service List* (GSL) has been of practical use to teachers and curriculum planners as it contains words within the word family each with its own

frequency. The list consists only of headwords, which means that the word "be" is high on the list, but assumes that the person is fluent in all forms of the word, e.g. am, is, are, was, were, being, and been. The GSL was written so that it could be used as a resource for compiling simplified reading texts into stages or steps. West and his colleagues produced vast numbers of simplified readers using this vocabulary. This is actually a very old list being based on frequency studies done in the early decades of this century. Doubts have been cast on its adequacy because of its age (Richards, 1974) and the relatively poor coverage provided by the words not in the first 1000 words of the list (Engels, 1968). Moreover, due the boom in technology in the last three decades tens of words have been coined and by such do not appear in the GSL

Engels makes two major points. Even if a limited vocabulary covers 95% of a text, a much larger vocabulary is still needed to cover the remaining 5% (p. 215). However, Engels overestimates the size of this vocabulary. He suggests 497,000 words. His second point is that the limited vocabulary chosen by West (1953) is not the best selection. Engels examined 10 texts of 1000 words each. He found that West's GSL plus numbers covered 81.8% of the running words (this did not include proper nouns which covered 4.13%). Engels' definition of what should be included in a word family did not agree with West's and so Engels considered that West's GSL contained 3,372 words. This is because Engels considered *flat* and *flatten*, and *police* and *policeman* to be different word families. West gives separate figures for such items but indicates through the format of the GSL that they are in the same family. This difference

however does not influence results. Engels considered the first 1000 of the GSL to be a good choice because the words were of high frequency and wide range (p. 221).

Engels (1968) further criticized the low coverage of the words not in the first 1000 words of the list. He found that whereas the first 1000 words covered 73.1% of the running words in the ten one thousand word texts he looked at, the words in the GSL outside the first 1000 covered only 7.7% of the running words. Other researchers have found a similar contrast (Sutarsyah, 1993; Hwang, 1989; Hirsh, 1992). However, Billuroğlu and Neufeld (2005) confirmed that the General Service List was in need of minor revision, but the headwords in the list still provide approximately 80% text coverage in written English. The research showed that the GSL contains a small number of archaic terms, such as *shilling*, while excluding words that have gained currency since the first half of the twentieth century, such as *plastic*, *television*, *battery*, *okay*, *victim*, and *drug*.

The question to be asked is: after the 2000 high frequency words of the GSL, what vocabulary does a second language learner need? The answer to this question depends on what the language learner intends to use English for. If the learner has no special academic purpose then the learner should work on the strategies for dealing with low frequency words. If however the learner intends to go on to academic study in upper high school or at university, then there is a clear need for general academic vocabulary. This can be found in the word list called the *University Word List* (Xue and Nation, 1984; Nation, 1990).

6.8.4. The University Word List (UWL)

The *University Word List* (UWL) is a list of 836 words that do not occur on the GSL, but are common in academic texts. It was developed by Xue and Nation and first published in 1984. It consists of words that are not in the first 2000 words of the GSL but which are frequent and of wide range in academic texts. Wide range means that the words occur not just in one or two disciplines like economics or mathematics, but occur across a wide range of disciplines. The word *frustrate* for example which is in the UWL can be found in many different disciplines. The purpose behind the setting up of the UWL is to create a list of high frequency words for learners with academic purposes, so that these words can be taught and directly studied in the same way as the words from the GSL can. Newspapers and magazines which are more formal make use of more of the UWL. Very formal academic texts make the greatest use of the UWL. The UWL is thus a word list for learners with specific purposes namely academic reading. It has been replaced by the *Academic Word List* (AWL).

6.8.5. The Academic Word List

The Academic Word List (AWL) came as an alternative to the University Word List. It was developed by Coxhead (2000). The AWL is shorter than the UWL (570 words), and provides more coverage of academic texts. It consists of 570 different words which provide coverage of 10% of an academic text, it means that 10% of an academic text (10% of all word tokens) consists of the AWL words. From the reader's perspective, this means that the knowledge of the AWL will assure the comprehension

of 10% of the vocabulary in an academic text. Nation (2001) suggested that, after the 2,000 general-purpose word level, further study of Coxhead's Academic Word List (Coxhead 2000) improves the coverage of academic articles considerably.

6.8.6. Core vocabulary

One of the principles by which learner vocabulary is often organized is that of word frequency. It is commonsense that only a relatively small number of words that are used much more often than others. Apart from the obvious "function" words - articles, modals, prepositions and so on — there are certain "content" words that occur with much greater frequency than others. As mentioned earlier. Carter discusses the idea of a core vocabulary for English - the notion that there exists a basic nucleus of words that can be used to communicate successfully in the language. Although he warns that there are different core vocabularies relating to different contexts (and that a proficient speaker will in fact have recourse to several such sets for different purposes) the idea of a definable "minimum requirement" that could stand as an attainable goal for language learners is very attractive.

As an approach to the daunting volume of new words that must be learned before effective communication can take place, Sinclair and Renouf (1988) recommend teachers making more use of the words a student already knows: "There is far more general utility in the combination of known elements than in the addition of less easily usable items" (p.155). In other words, aiming for thorough knowledge of a core

vocabulary is more useful than spending time learning a wide range of low frequency lexis.

How much core vocabulary it is necessary to learn for incidental vocabulary learning to be viable depends on the nature of the text and how well the learner understands the context of the new word. For authentic materials most research indicates that knowing the most frequent 5,000 words should be enough (Schmitt, Schmitt & Clapham, 2001). Hazenberg and Hulstijn (1996) found that a 10,000 word vocabulary might be required to cope with university study in a second language, given the demands academic language places on the learner. Nation (1999, p.13) suggests that students concentrate on the most frequent 2,000 words then master the academic vocabulary of the Academic Word List (AWL) (Coxhead, 1998) or the University Word list (UWL) (Xue and Nation, 1984) rather than spend the time on the third thousand words. This would give the learner 86.1% coverage of academic text: not quite enough for fluent reading, but certainly a good coverage for a much smaller investment of learning than 10,000 - or even 5,000 words - would require. The work of Nation and Laufer has been influential on the subject of vocabulary size and text coverage (the amount of text that can be understood by a learner with a vocabulary of a certain size). Nation and Waring (1997) suggest that 3,000 to 5,000 words are sufficient to provide a basis for comprehension, and as few as 2-3,000 words would serve for productive purposes. This is on the basis that, according to analysis of the Brown corpus (Francis & Kucera, 1982) the most frequent 1,000 words (lemmas) in English

cover about 72% of a text; while 3,000 words give 84% coverage and 5,000 words 88.7%. Nation draws the line at "3,000 or so" that are "immediate high priority and there is little sense in focusing on other vocabulary until these are well-learned" (Nation & Waring, 1997, p.11), in other words, it is the threshold vocabulary. Schmitt (2000) also stresses that the learning of the high frequency lexicon cannot be left to chance and must be explicit.

Conclusion

It becomes clear from the above that vocabulary and reading comprehension are closely linked. Both theoretical and empirical evidence attest to this fact. Such a relationship is reciprocal. However, a slight disagreement exists among researcher on this reciprocity, that is on which affects which, or which is the result of which. This has led somehow to a catch-22 situation. However, in any case what is of concern for the present study is that there does exist a causal relationship between vocabulary and reading comprehension. Confirmation of such an assertion has led researchers to conclude that the amount of unknown vocabulary is crucial for the foreign language learner to understand what he/she reads and research has been directed towards determining how much unknown vocabulary can be tolerated in a text before it interferes with comprehension, or lexical threshold.

In the pursuit of achieving this aim, lexical text coverage has emerged as a key concept. A key finding was that only a relatively small number of words are used much more often than others. Consequently, for effective language teaching pedagogy, it

became important to determine which and how many words should second/foreign language learners know as well as what vocabulary a language learner needs. High frequency words became the focus of investigation and several lists were developed, the most used ones are: *The General Service List* (West, 1953), *The University Word List* (Xue and Nation, 1984), and *The Academic Word List* (Coxhead, 2000).

CHAPTER 7

METHODS AND PROCEDURES

Introduction

This chapter will describe the steps that were taken the process of carrying out the study. Specific sections will be devoted to a discussion of the reading material (the textbooks), instruments, and procedures that were employed in this investigation.

The reader is reminded that the research question which are the focus of the study are:

1. What is the lexical coverage of textbooks used by Algerian middle and secondary school students?
2. What is the readability level of middle and secondary school textbooks?
3. Are Algerian EFL learners learning sufficient, useful and appropriate vocabulary items?

It should be pointed out here that as a methodology to answer these questions, the study combined qualitative and quantitative research methods. Statistical analysis preceded qualitative interpretation. Qualitative research methods were used to address research question that require explanation or understanding of phenomena.

Quantitatively the data consisted of a corpus made out of the selected textbooks.

7.1. Corpora and textbook corpus research

A corpus, plural corpora, has been defined as “a collection of materials that has been made for a particular purpose, such as a set of textbooks which are being analyzed and compared ...” (Richards et al., 1992, p.110). Schmitt (2000, p. 68) referred to

corpora as “simply large collections or databases of language, incorporating stretches of discourse ranging from a few words to entire books”. For the purpose of this research the following definition of corpus can be worked out: a corpus is a collection of texts in electronic form used for linguistic research analyzed by means of a computer-assisted analysis method.

The early corpora appeared in the first third of the previous century and were made manually. Examples of these are The Brown University Corpus (Kucera & Francis, 1967). It contains 500 samples of English-language text, totalling roughly one million words, compiled from works published in the United States in 1961.

As a counterpart of the BROWN Corpus of American English, The Lancaster-Oslo/Bergen Corpus (LOB Corpus) (Hofland & Johansson, 1982; Johansson & Hofland, 1989) was compiled by researchers in Lancaster, Oslo and Bergen between 1970 and 1978. It consists of one million words.

The British National Corpus (BNC), University of Oxford, is a 100 million word collection of samples of written and spoken language from a wide range of sources, designed to represent a wide cross-section of current British English, both spoken and written.

Another English corpus is the COBUILD, an acronym for Collins Birmingham University International Language Database, which is a British research facility set up at the University of Birmingham in 1980. It is also referred to as The Bank of English.

The Oxford English Corpus is also a text corpus centred only in the English language. It is thought to be the largest corpus of its kind, containing over two billion words.

However, a corpus is not necessarily compiled from a very huge amount of large data. A textbook or a set of textbooks can be viewed as a corpus. In literature, research has been done to analyze textbooks as a corpus.

7.2. Description of the Textbooks:

Below is a cursory description of the textbooks used in the Algerian middle and secondary schools and which constitute the materials of the present study. The number of books used is seven. All are published by The National Authority for School Publications.

7.2.2. Middle school textbooks

Following the reform of the Algerian Educational system launched from 2003 through 2008 textbooks for the teaching of English were published one after the other. They constitute the official textbook designed by the Ministry of Education for the pupils in their four years of English study in the Middle school. These manuals are: *Spotlight on English* in the first year, *Spotlight on English Book Two* in the second year, *Spotlight on English Book Three* in the third year, and *On the Move* in the fourth year. These textbooks are designed along the principles of the Competency-based approach which relies basically on project works, problem-solving situations and task-based-teaching or practices. Table 1 displays broad information about these textbooks.

Table 1

Middle school textbooks

Level	Book	First published	Number of files	Number of pages	Authors
First year	Spotlight on English One	2003	7	189	L. Merazga, K. Achour H. Ameziane, F. Bouhadiba W. Guedoudj, O. Mekaoui B. Riche, L. Tamrabet
Second year	Spotlight on English Two	2004	5	125	L. Merazga, Z. Torche F. Bouhadiba W. Guedoudj
Third year	Spotlight on English Three	2005 & 2009 (Rev. ed.)	4	188	H. Ameziane, N. Khouas K. Louadj, B. Riche
Fourth year	On the Move	2006	6	192	S.A. Arab B. Riche

7.2.2.1. Spotlight on English Book One (first year)

Spotlight on English contains a task-based content, organized along a structural/functional and notional/topical lines, and whose purpose is to bring the learners to engage in the acquisition of structures (e.g. how to form the present simple), functions (e.g. greeting, offering, asking permission), notions (e.g. time, cause, quantity) and topics (e.g. sport, family, travel) to develop their four language skills (listening, speaking, reading and writing). It has also adopted a competency-based approach by allowing the learners to develop skills through the elaboration of individual or group projects at the end of each unit studied. It has abandoned the

traditional presentation-based approach to English language teaching, or the *3Ps* approach, i.e. presentation, practice, and production.

7.2.2.1.1. General layout

Spotlight on English is the official textbook designed by the Ministry of Education for the pupils in their first year of English study in the Middle school. The manual is a pedagogical document that handles the new official syllabus adopted within the framework of the recent Education Reform. The syllabus is communicatively oriented and thematically organised. It is developed through a pre-file followed by seven files: *Hello, Family and Friends, Sports, In and Out, Food, Inventions and Discoveries, and Environment*. Each file is presented with varied colourful illustrations.

The book opens with an Introduction in Arabic addressing the students and giving information about the book and its contents. This is followed by a pre-file entitled "*You know English*" which includes basic notions and terms in English as well as the English alphabet. The alphabet is presented together with "schoolthings" (sic), and a few verbs of action used every day in a classroom such as *sit down, stand up, listen, read, etc.*

Each file is made up of three sequences and the following sections: *Listening Scripts, Learn about Culture, Check, and Your Project*. Supporting Texts in *Spring One* are related to scientific subjects or to introduce prominent scientists. They try mainly to present specific subjects the learners study elsewhere so as to facilitate their understanding. The aim of such texts is mainly to reinforce the idea of studying English as a means to gain access to scientific content. In *Spotlight on English* authentic texts

deal with a varied set of topics related to different cultural facets such as music and food.

7.2.2.1.2. File Structure Description

The structure of a file in *Spotlight on English One* is displayed in Figure 1. Each of the seven files of *Spotlight on English One* consists of three sections called sequences (Sequence 1, Sequence2, and Sequence 3). Each sequence consists of:

- *Listen and Speak* aiming at developing oral interaction in English by recognizing sound, making sense of what is heard, training the learner's tongue and lips to speak correctly and meaningfully, acting out guided dialogues, playing roles in situations similar to real life

Practise aims at practising the language points.

- *Produce* aims at developing the writing skill as students are given a written task.

These sections are then followed by other sections:

- *Listening scripts* which provides the transcripts of the listening activity material.
- *Learn about the culture* consists of various pieces of writing about Britain and other English-Speaking countries. Learners will compare and contrast them with Algeria through class discussion, reports and writing tasks. It is devoted to the discovery of world culture (flags of English-speaking and non-English-speaking countries, world monuments, greeting in different languages, sports in the world, Chinese horoscope, etc.),

- *Reminder* provides a summary of the main point covered so far in the file.
- *Check* aims at consolidating what has been covered so far in the file.
- *Your project* assigns a project and provides guidelines for carrying out the project suggested.

A glossary is at the end of the book.

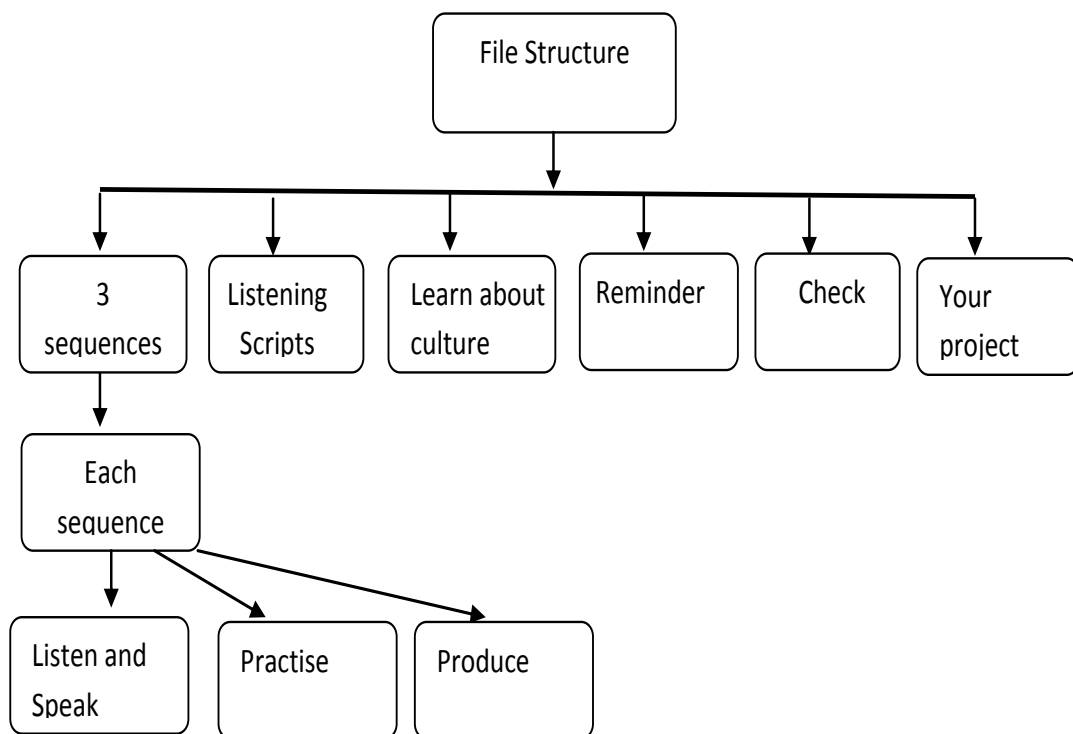


Figure 1: *File structure in Spotlight on English Book 1.*

7.2.2.2. Spotlight on English Book Two (second year)

7.2.2.2.1. General layout

Again, the book opens with an Introduction in Arabic addressing the students and giving information about the book and its contents. The objectives of the sequences are stated right at the beginning.

Spotlight on English Book Two is a textbook designed for second year pupils, i.e. pupils who have studied English for a year. The syllabus is built on a competency-based approach and seeks to consolidate and develop the three main competences pupils initiated to, namely: to interact orally in English, to interpret authentic oral or written documents, and to produce simple oral or written messages.

The textbook is made up of five files, each cantered around a theme: *A person's Profile* in file one, *Language Games* in file two, *Health* in file three, *Cartoons* in file four, and *Theatre* in file five. The five files are organized in a similar way, that is each consists of three sequences.

7.2.2.2.2. File Structure Description

Similarly to the first year book, this one follows roughly the same file structure. Figure 2 below displays the file structure in *Spotlight on English Book 2*. Each file consists of three sequences. All three sequences unfold as follow:

- *Objectives*: The overall objective of the file is the project pupils need to accomplish at the end of the file. Objectives are set up at the beginning of the three sequences. Each sequence in each file opens up with the objectives it sets out to achieve.
- *Listen and Speak*

It unfolds in the following pattern:

- *Conversation (listening)*
- *Pronunciation and spelling*
 - a- *Listen and repeat*
 - b- *Identify*
 - c- *Compare*
- *Practice stress and intonation*
 - a- *Practice*
 - b- *Go forward*

- *Discover the language:* This section is made up of the following sub-parts:
 - a- *Read*
 - b- *Practice*
 - c- *Reminder*
- *Listening scripts*
- *Learn about culture*

This section equips pupils with some cultural facts related to the topic of the file.

Extracts from real-life contexts initiate tasks and activities, oral or written, that help pupils discover some cultural aspects, and re-invest the items acquired during the *Learn the language* sequences.

- *Check*

This section is built around a series of tasks from which the teacher can select the most relevant ones. It aims at helping learners consolidate their acquisitions, and seeks to make sure that the items covered so far have been correctly acquired; some tasks are

carried out through both written and oral media. *Check* acts as a classroom evaluation of the learning process, and helps guide the teacher to some remedial work in case pupils encounter some learning problems.

- *Project*

Pupils through the project work are placed in front of a problem situation whose resolution requires the integration and re- investment of the knowledge, skills and capacities acquired along the file in order to come out with a tangible output or a solution , and in which pupils' attitudes are inferred from teacher's observation. For example, in file three the project involves making a herbal or medical guide; in file four it is writing a story and making it into a cartoon strip

- *Self-assessment*

It is a grid that closes up the file and in which pupils evaluate their acquisitions of file components. Self- assessment is one way of rendering pupils responsible of their learning. They decide by themselves what they learnt to understand, to say, to write and to do. Likewise, they express their points of view as to whether they enjoyed or not the topic of the file, the projects, the activities selected, working alone, with a partner or in groups.

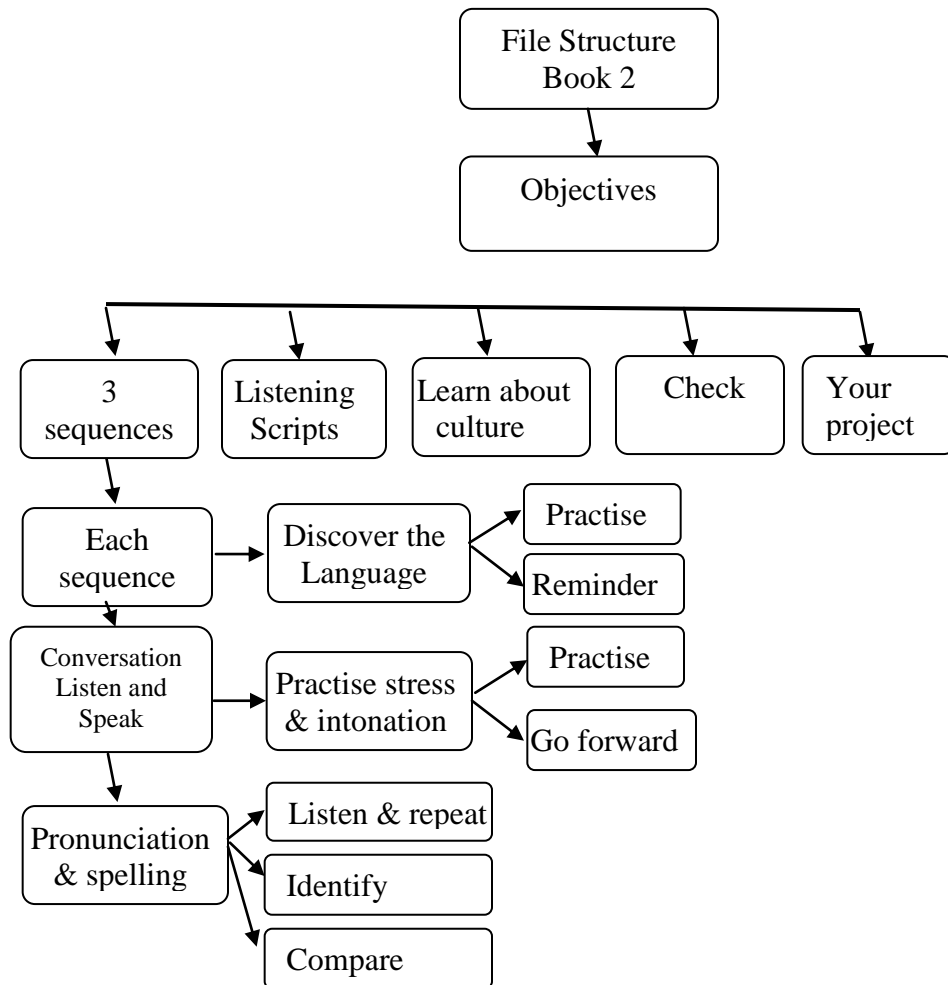


Figure 2: *File structure in Spotlight on English Book 2.*

7.2.2.3. Spotlight on English Book Three (third year, revised edition)

The book opens with a preface to the revised edition. It is said that the book builds upon the format of the preceding one and the general layout is roughly the same. The reading and writing skills are emphasized in this book and authentic texts are included.

7.2.2.3.1. General layout

This book follows roughly the same pattern as the following ones. It contains four files: *Communication* (file 1), *Travel* (file 2), *Work and Play* (file 3), and *Around the World* (file 4). Listening scripts, abbreviations used in electronic messaging, the phonetic symbols, a list of irregular verbs, and a few spelling rules are at the end of the book.

7.2.2.3.2. File Structure Description

Right at the onset of the file, the project is announced. Then, follows a section entitled *Preview* in which the language functions dealt with in each sequence are listed. Figure 3 displays the file structure in Spotlight on English Book 3. Each file contains three sequences. Each sequence contains:

- Listen and speak
 - Say it clear
 - Practice
 - Imagine
- Read and write

The sequences are followed by three other sections:

- Snapshots of culture
- Activate your English
- Where do we stand now?

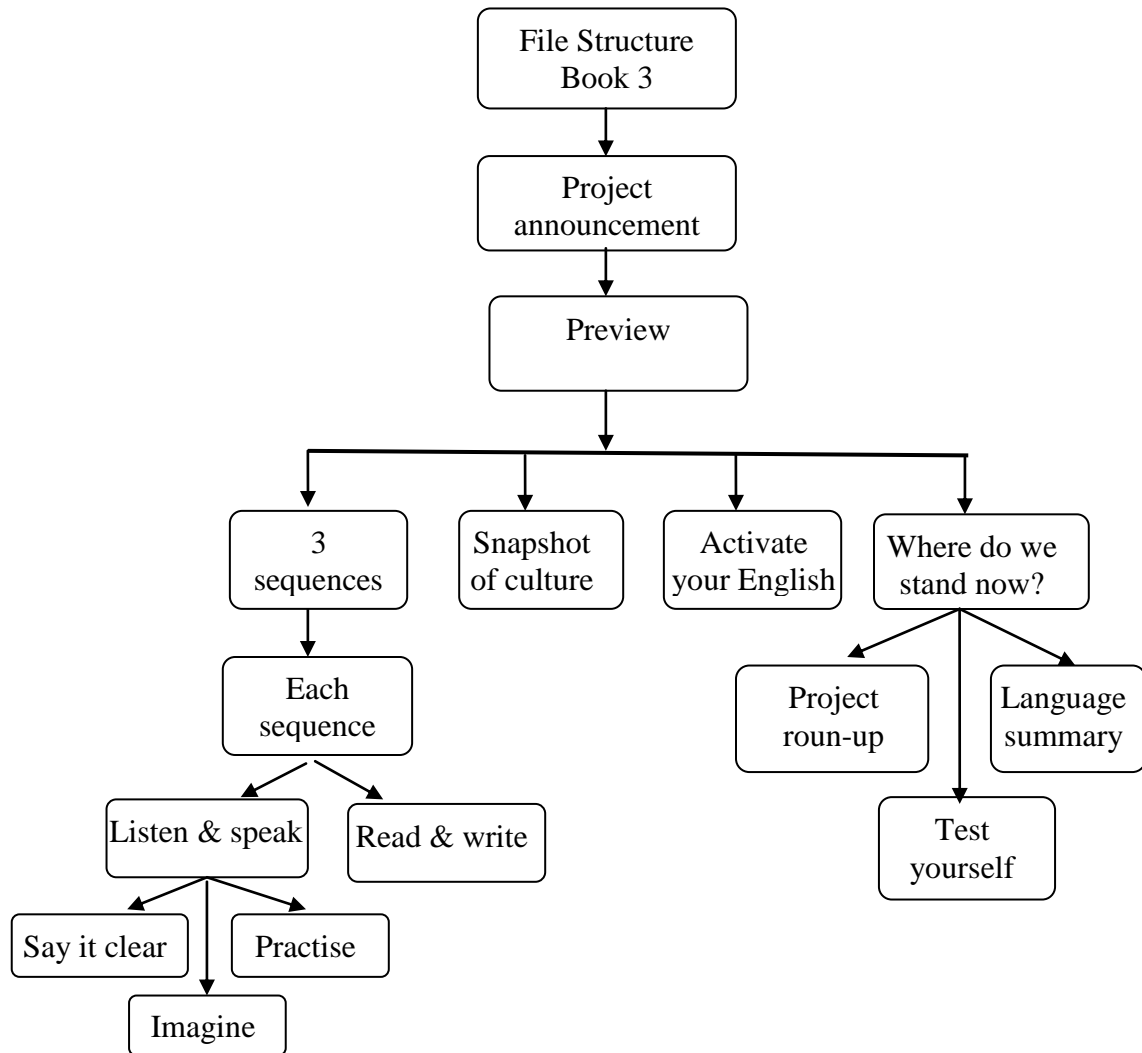


Figure 3: *File structure in Spotlight on English Book 3.*

7.2.2.4. On the Move (fourth year)

7.2.2.4.1. General layout

Contrary to the other books, no introduction in Arabic is provided here. It opens with a section entitled ‘From the authors of this book to the student’ followed by another one entitled ‘To the Teacher’.

Again the book consists of files. Their number is six: *It’s my treat* (File 1), *You can do it* (File 2), *Great expectations* (File 3), *then and now* (File 4), *Dreams, Dreams...* (File 5), and *Facts and fiction* (File six). The listening scripts appear at the end of the book, as well as a *Grammar Reference* section.

7.2.2.4.2. File Structure Description

Each file in *On the Move* is organized as follows:

1. **Preview**
2. **Food for thought:** two contrasted pictures aimed at setting the student thinking and serve as a warm-up.
3. **Listen and consider:** listening comprehension activities and practicing grammar in oral and written texts.
4. **Read and consider:** reading comprehension activities and practicing grammar.
5. **Words and sounds:** acquisition of new vocabulary related to the topic of the file.
6. **Take a break:** reviewing and relaxing
7. **Research and report:** engaging in a research task.
8. **Listening and speaking:** acquiring a good command of listening and speaking.
9. **Reading and writing:** acquiring a good command of reading and writing.

10. **Project round-up:** comparing student's project with the one given as an example and discuss it with peers.

11. **Where do we stand now?** Assessing achievement.

12. **Time for...** This section consists of songs, jokes, and relaxing activities.

7.2.3. Secondary school textbooks

The books in use in Secondary school are: *At the Crossroads* (first year), *Getting Through* (second year, and *New Prospects* (third year). Table 2 below provides information on these books.

Table 2

Secondary school textbooks

Level	Book	First published	Number of units	Number of pages	Authors
First year	At the Crossroads	2005	5	155	S.A. Arab, B. Riche H. Ameziane, H. Hami K. Louadj
Second year	Getting Through	2006	8	208	S.A. Arab, B. Riche M. Bensemene H. Ameziane, H. Hami
Third year	New Prospects	2007	6	272	S.A. Arab, B. Riche M. Bensemene

Contrary to the books of the middle school, where the didactic entity is the file, these books are organized in units. Below is a brief description of each book.

7.2.3.1. At the Crossroads (first year)

7.2.3.1.2. General layout

This textbook is basically designed for learners aged 15 to 16, who have already studied English for four years at Middle School level. The number of its pages is 155. This book starts with a ‘To the teacher’ section. It starts by spelling out that it is assumed that the students or whom the book is addressed have completed the preceding four books of the middle school. The author explains that the course book is designed to comply with the relevant Ministry of Education curriculum as laid down in January 2005 on the basis that these students have completed the four years of English in the new Middle school EFL syllabus. The textbook is intended for all streams and consists of five units.

At the crossroads is composed of:

- Contents (p. II)
- Map of the book (pp. III-VII)
- A note to the teacher (pp. VIII-IV)
- A note to the pupil. (pp. X-XI)
- Phonetic symbols (pp. XII – XIII)
- 5 Units (pp. 2-155)
- Scripts for listening appear at the end of the book.

The five files included in the book are: *Getting Through* (Unit 1), *Once Upon a Time* (Unit 2), *Our Findings Show ...* (File 3), *Eureka* (Unit 4), and *Back to Nature* (Unit 5). Then a ‘*To the Student*’ section follows.

7.2.3.1.3. Unit Structure Description

Each unit contains the following:

Preview: states the objective of each section

Sequence one: listening and speaking

Anticipate: pre-listening activities.

Listen and check: listening comprehension activities.

Say it clear: speaking activities.

Your turn: listening and speaking activities.

Sequence two: reading and writing

These two sections are of the same pattern, they aim at producing oral and written discourse).

Anticipate: pre-reading activities.

Read and check: reading comprehension activities.

Discover the language: focus on some language exponents.

Write it right: writing activities.

Sequence three: developing skills: the students combine the four skills in problem-situations.

Stop and consider: grammar, spelling, and pronunciation rules.

Sequence four consolidation and extension to expand on and consolidate social skills, and to make students aware of problem areas in pronunciation and stress.

Project workshop: (group work) the learners are assigned projects to carry out and are asked to follow a checklist of instructions for its realization.

Check and progress: self-assessment.

The figure on next page (Figure 4) displays the layout of a unit in *At The Crossroads*.

7.2.3.2. Getting Through (second year)

7.2.3.2.1. General layout

The book starts with an introduction addressed to the student having the layout of an email. It explains the content of the book and the objectives of each section. This is followed by another ‘email’ addressed to the teacher. The book consists of eight units: Unit 1: *Signs in the Time*; unit 2: *Make Peace*; unit 3: *Waste Not, Want Not*; unit 4: *Budding Scientist*; unit 5: *News and Tales*; unit 6: *No Man is an Island*; unit 7: *Science or Fiction?* Unit 8: *Business is Business*.

At the end of the book we can find transcripts of text for listening and a ‘Grammar Reference’ section.

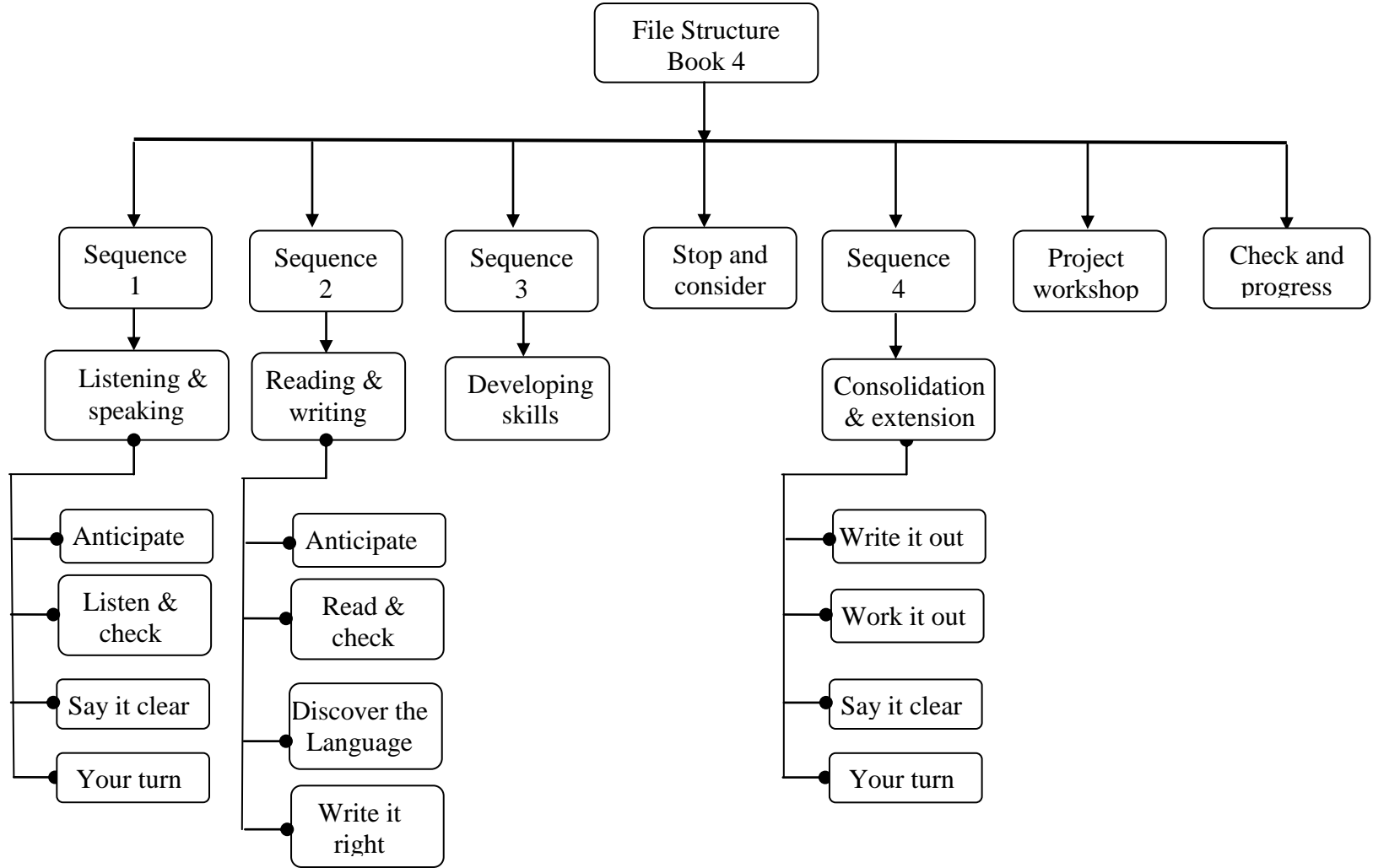


Figure 4: File structure in At the Crossroads.

7.2.3.2.2. Unit Structure Description

In each unit there are five stages. These are called: Discovering Language, Developing Skills, Putting Things Together, Where Do We Go From Here? and Exploring Matters Further.

The section *Discovering Language* is aimed at learning the vocabulary, spelling, pronunciation and grammar. *Developing Skills* as explained by the authors purports to help students develop language as well as intellectual such as thinking, guessing, making hypotheses analyzing, planning, etc.). In the next stage, the students are supposed to put into practice what they have learnt by making a project. The part *Where Do We Go From Here*' is for self-assessment. The last section *Exploring Matters Further* presents additional authentic material for extensive reading to further consolidate grammar structures and learn more vocabulary, the authors explain.

The longest sections are *Discovering Language and Developing skills*. These consist of subsections:

Discovering Language

- Before you read
- As you read
- After reading
- Practice
- Write it right
- Say it loud and clear
- Working with words

Developing skills

- Listening and speaking
- Reading and writing

7.2.3.3. New Prospects (third year)**7.2.3.3.1. General layout**

This book also opens with a *Foreword*. There are six units: Unit 1 *Exploring the past*; Unit 2 *Ill-gotten gains never prosper*; Unit 3 *Schools: Different and alike*; Unit 4 *Safety first*; Unit 5 *Are we alone?* Unit 6 *Keep cold!* Transcripts of the texts for listening as well as a series of texts for extensive reading appear at the end. Illustrations are rare comparatively to the other books.

7.2.3.3.2. Unit Structure Description

Each unit starts by stating the *Project Outcome* and comprises two main parts with two sequences each. The first part, Language Outcomes is divided into *Listen and Consider* and *Read and consider*. These, according to the authors, aim at getting the students internalize the thematic and linguistic tools they will need in the subsequent part. This part ends with *Take a Break*. The second part, entitled *Skills and strategies outcomes* is concerned with the structural and discursive aspects of the text. It starts with a section called *Research and Report* followed by two sequences: *Listening and speaking* and *Reading and Writing*, then by *Project Outcome, Assessment and Time For...*

7.3. Method of analysis

This section gives a detailed account of the methodology adopted to answer the research questions, object of the present study. The reader is reminded that one of the main aims of this study is to examine what vocabulary exists in learners' textbooks, and to determine if the learners are likely to meet the 95% comprehension criteria. It is believed that if such is not the case, educators must then provide the supplemental vocabulary to bridge this gap. Without this kind of bridge, reading would be a daunting task and learners would certainly resort very often to a dictionary. This would certainly deter the most perseverant reader.

To examine the vocabulary of the textbooks used in Algerian schools, it was necessary to compile all the textbooks to form a basis of comparison. A full list of the words occurring in each textbook is needed. For such a task a computer method was adopted involving the use of software tools. The methodology adopted is divided into two main steps: the first step consisted in corpus compilation and the second step consisted in the study of the corpus with software. The instruments used as well as the preparation of the corpus for analysis are described hereafter. For ease of presentation, the instruments (step 2) are described first.

7.3.1. Instruments

The study relied on four computer software for the data analysis two of which are vocabulary profilers. Vocabulary profiling is a measure of the proportions of low and high frequency vocabulary used in a written text. In addition to frequency information, a profiler designed for lexical analysis of texts often provides other information such as

the presence/absence of the set of words from the input text in other specialized word lists.

A vocabulary profiler divides the words of any text into four categories by frequency: (1) the most frequent 1000 words of English (level K1), (2) the second most frequent thousand words of English, i.e. 1001 to 2000 (level K2), (3) the academic words of English (the AWL, 570 words that are frequent in academic texts across subjects), and (4) the remainder which are not found on the other lists (off-list). In other words, VP measures the proportions of low and high frequency vocabulary used by a native speaker or language learner in a written text. A typical NS result is 70-10-10-10, or 70% from first 1000, 10% from second thousand, 10% academic, and 10% less frequent words. This relatively simple tool has been useful in understanding the lexical acquisition and performance of second language learners. Its latest version is known as Range

7.3.1.1. Range and Frequency programs for Windows based PCs

Range and Frequency programs for Windows based PCs (Heatley et al., 2002) is a vocabulary profiler (available at <http://www.victoria.ac.nz/lals/staff/paul-nation.aspx>) developed by Paul Nation and Alex Heatley (2002) of the Victoria University of Wellington and is freely downloadable. As such, this software, as described in the instructions which come with it, compares the words in a text file with three word frequency lists and classifies each word into one of four categories: a) a word occurring within the top 1000 high frequency words on West's 1953 General service List, b) a

word occurring within the 1001-2000 high frequency words range on West's 1953 General service List frequency words, c) a word from Nation's 1984 University Word List (Nation, 1990), or d) an unknown word (i.e., all words not included in the above lists). The programme includes also the British National Corpus High Frequency Word List (BNC HFWL 1st–14th 1,000) based on English words' occurring frequency and range.

RANGE is used to compare the vocabulary of up to 32 different texts at the same time. For each word in the texts, it provides a range or distribution figure (how many texts the word occurs in), a headword frequency figure (the total number of times the actual headword type appears in all the texts), a family frequency figure (the total number of times the word and its family members occur in all the texts), and a frequency figure for each of the texts the word occurs in. It can be used to find the coverage of a text by certain word lists, create word lists based on frequency and range, and to discover shared and unique vocabulary in several pieces of writing.

RANGE can also be used to compare a text against vocabulary lists to see what words in the text are and are not in the lists, and to see what percentage of the items in the text are covered by the lists, namely lexical coverage. According to its designers, it has been used to answer the following questions.

- What common vocabulary is found in all the texts?
- How large a vocabulary is needed to read a text?

-
- If a learner has a vocabulary of 2,000 words, how much of the vocabulary in the text will be familiar to the learner?
 - What are the words in the text which the learner is not likely to know?
 - How well does the course book prepare learners for the vocabulary in newspapers?
 - How rich a vocabulary do second language learners use in their free writing?

The RANGE software can be used to compare a text against certain base word lists to see what words in the text are and are not in the word lists, and to see what percentage of the vocabulary items in the text are covered by the lists, namely lexical coverage.

Three readymade base lists are included. The lists include both American and British spellings. The first (BASEWRD1.txt) includes the most frequent 1000 words of English. The second (BASEWRD2.txt) includes the 2nd 1000 most frequent words, and the third (BASEWRD3.txt) includes words not in the first 2000 words of English but which are frequent in upper secondary school and university texts from a wide range of subjects. All of these base lists include the base forms of words and derived forms. The first 1000 words list thus consists of around 4000 forms or types. The sources of these lists are *A General Service List of English Words* by Michael West (Longman, London 1953) for the first 2000 words, and *The Academic Word List* by Coxhead (1998, 2000) containing 570 word families. The first thousand words of *A General Service List of English Words* are usually those in the list with a frequency higher than

332 occurrences per 5 million words, plus months, days of the week, numbers, titles (Mr, Mrs, Miss, Ms, Mister), and frequent greetings (Hello, Hi etc).

The criteria used in the RANGE program to make word families were based on Bauer and Nation's (1993) six-level basic word building processes, which include all the affixes, inflected and derived forms. Word families are regarded as an important counting unit in terms of the learning load (Nagy et al. 1989). The concept of a word family is used to represent a group of words whose meanings can be inferred when the meaning of the base form in the group is known to a learner. Therefore, comprehending regularly inflected or derived members of a word family does not require much effort, namely, not having to learn each form separately. The word forms in the base lists are grouped into word families under a headword. For instance, the headword *accuse* is grouped with its members *accusing*, *accusingly*, *accuses*, *accused*, *accusation*, *accusations*, *accuser*, and *accusers* to form a word family. Thus, the four family members are counted as the same word *accuse*.

Below are some other examples of word families from the BNC HFWL 3rd 1,000.

ABSENCE

ABSENCES

ACCELERATE

ACCELERATED

ACCELERATES

ACCELERATING

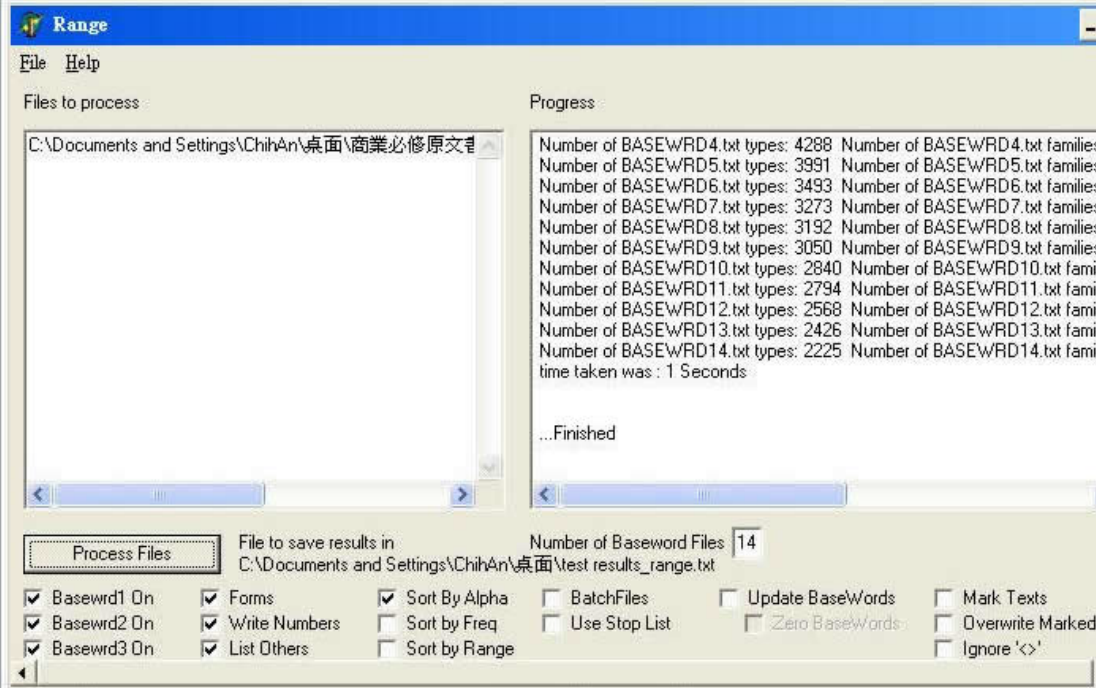
ACCELERATOR
ACCELERATORS
ACCELERATION
ACCELERATIONS
ACCORDINGLY
ACCOUNTANT
ACCOUNTANTS
ACCOUNTANCY
ACCUSE
ACCUSING
ACCUSINGLY
ACCUSES
ACCUSED
ACCUSATION
ACCUSATIONS
ACCUSER
ACCUSERS

Furthermore, RANGE can exclude words from the count provided that a list of them is supplied (called *stoplist*).

Figure 5 is a screenshot, which demonstrates the fourteen base word lists, i.e. fourteen 1,000 high-frequency English word families made from the British National

Corpus. For example, Base word list 1 includes 1,000 base forms, their inflected forms and derivatives, thereby making a total of 6,348 different words (types).

Number of BASEWRD1.txt types: 6348	Number of BASEWRD1.txt families: 1000
Number of BASEWRD2.txt types: 5593	Number of BASEWRD2.txt families: 1000
Number of BASEWRD3.txt types: 4517	Number of BASEWRD3.txt families: 1000
Number of BASEWRD4.txt types: 4288	Number of BASEWRD4.txt families: 1000
Number of BASEWRD5.txt types: 3991	Number of BASEWRD5.txt families: 1000
Number of BASEWRD6.txt types: 3493	Number of BASEWRD6.txt families: 1000
Number of BASEWRD7.txt types: 3273	Number of BASEWRD7.txt families: 1000
Number of BASEWRD8.txt types: 3192	Number of BASEWRD8.txt families: 1000
Number of BASEWRD9.txt types: 3050	Number of BASEWRD9.txt families: 1000
Number of BASEWRD10.txt types: 2840	Number of BASEWRD10.txt families: 1000
Number of BASEWRD11.txt types: 2794	Number of BASEWRD11.txt families: 1000
Number of BASEWRD12.txt types: 2568	Number of BASEWRD12.txt families: 1000
Number of BASEWRD13.txt types: 2426	Number of BASEWRD13.txt families: 1000
Number of BASEWRD14.txt types: 2225	Number of BASEWRD14.txt families: 1000



The screenshot shows a software window titled "Range" with a menu bar (File, Help) and two main panes. The left pane, "Files to process", contains a list of files: C:\Documents and Settings\ChihAn\桌面\商業必修原文書. The right pane, "Progress", displays the same data as the table above, followed by "time taken was : 1 Seconds" and "...Finished". Below the panes, there is a "Process Files" button, a "File to save results in" field with the path C:\Documents and Settings\ChihAn\桌面\test results_range.txt, and a "Number of Baseword Files" field set to 14. At the bottom, there are several checkboxes for processing options: Basewrd1 On, Basewrd2 On, Basewrd3 On, Forms, Write Numbers, List Others, Sort By Alpha, Sort by Freq, Sort by Range, BatchFiles, Use Stop List, Update BaseWords, Zero BaseWords, Mark Texts, Overwrite Marked, and Ignore '<'.

Figure 5: Number of word families and word types in the BNC high-frequency word lists

FREQUENCY

FREQUENCY is another program that runs on an ASCII text to make a frequency list of all the words in a single text. It can only run one text at a time. The output is an alphabetical list, or a frequency ordered list. It gives the rank order of the words, their raw frequency and the cumulative percentage frequency. Table 3 is a sample output from FREQUENCY provided in the instructions that accompany the programme.

Table 3

Sample output from FREQUENCY

Word Type	Rank	Frequency	Cumulative Percent
THE	1	271	7.55
OF	2	134	11.28
A	3	108	14.29
IN	4	101	17.10
TO	5	98	19.83
GROUP	6	88	22.28

In the example, the word type *a* is the third most frequent word. It occurs 108 times in the text, and along with *the* and *of* covers 14.29% of the text. On its own it covers 3.01% (14.29 minus 11.28) of the text. Further details are provided in Chapter 8.

7.3.1.2. The Compleat Lexical Tutor

The Compleat Lexical Tutor (available at www.lextutor.ca) is also a vocabulary profiler. It is a Canadian free web-based resource developed by Tom Cobb (2008) as a web based version of Nation's Range and Frequency program with a view to the

practical application of data-driven learning using principled approaches supported by published research. It includes 26 applications among which two, range and frequency, are of interest to the present study. Figure 6 below displays a snapshot of The Compleat Lexical Tutor

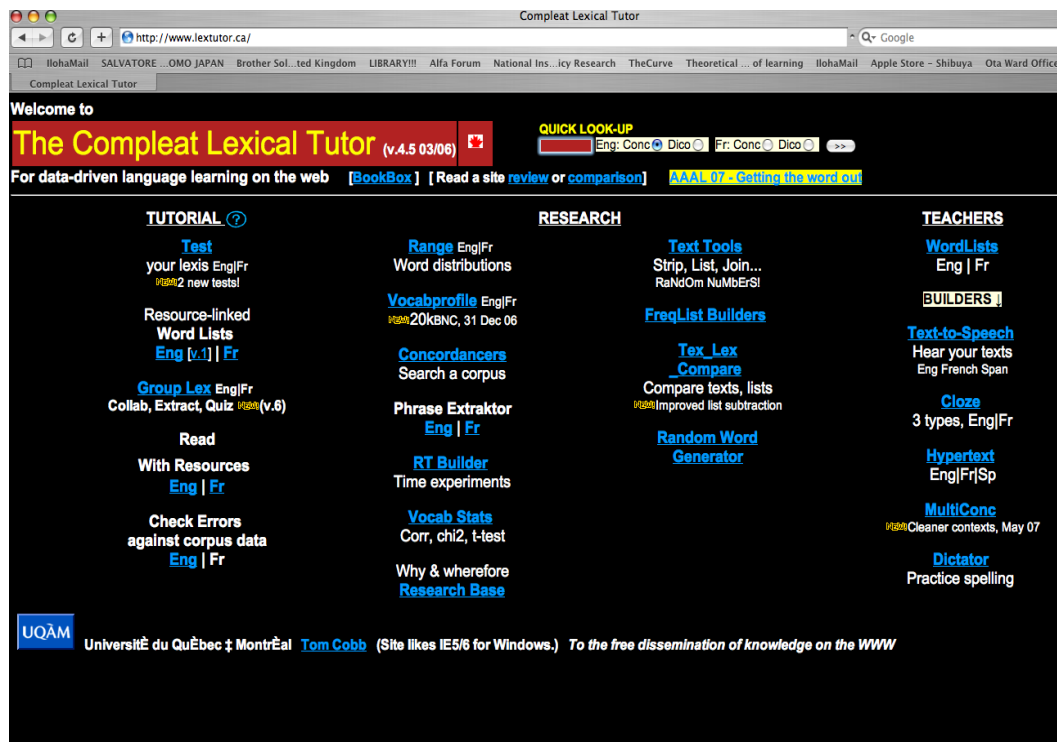


Figure 6: Snapshot of The Compleat Lexical Tutor

7.3.1.3. TextCompare software

This study was interested in the comparison of vocabulary that occurred in different textbooks. Therefore, software that could do the task was developed by the researcher in collaboration with colleagues⁸ from the Department of Computer Science,

⁸ Special thanks are due to Abdallah Khababa and Mabrouk Korichi, Department of Computer Science, Ferhat Abbas University, for their help to develop this software.

Ferhat Abbas University, Setif⁹ named *TextCompare software* This programme compares any two texts, say Text1 and Text2 and helps find out which words occurred in one text and not the other (Text1 minus Text2, or Text2 minus Text1), and which words occurred in both texts (the intersect).

7.3.1.4. TextMaster

Text Master (2009, available at <http://www.nontube.com/products/text-master/>) is a plain text editor freely downloadable from the internet. It is an alternative of NotePAD. Its chief feature is statistical analysis. For each text it gives the total number of characters, the number of unique characters, the total number of words, the number of unique words, character frequency, and word frequency. It can also generate a word list, make the text lowercase and uppercase, and reverse a sequence (for example, family can be turned into *ylimaf*).

7.3.2. Procedure

7.3.2.1. Preparing textbooks for analysis

7.3.2.1.1. Creating Word Lists from English Textbooks

The process of generating and preparing word lists for analysis was done in four steps described hereafter.

⁹ Special acknowledgments must be made here to Mabrouk Korichi for his help in developing this software program.

7.3.2.1.1.1. Step 1: Scanning the books

The first step in the analysis of the vocabulary in the textbooks under investigation was to create a corpus consisting of all the words occurring in the seven textbook. For this purpose, the contents of the seven textbooks were converted into machine-readable format text files, so that they could be “read” by the computer programme used to analyze them. This was done by scanning the books using a scanner with a built-in function called Optical Character Recognition (OCR). This function renders a text printed on paper into files ‘read’ by word processing software (Microsoft Word 2007). The outcome is a copy of each book, *as is*, with text and illustration.

All the texts from the material in textbooks were entered into the corpus including unit titles, section headings, and instructions, since students would encounter them all during their use of the textbook. Omitted from the corpus were acknowledgements, introduction (for teachers and students), table of contents, and glossaries.

7.3.2.1.1.2. Step 2: Cleansing the untreated text

As the study was interested only in text, the pictures, diagrams, caricatures and illustrations were unnecessary. Removing them was a very simple matter. The scanned material was simply saved in the word processor as ‘plain text’, that is files with the extension ‘.txt’. In this way the outcome consisted in text only. The scanning and file saving procedures were as follows:

All the running words in each file/unit of each textbook were extracted and converted to ‘electronically – recognizable’ files, (e.g. B1F1.txt). This was done using a

predetermined labelling system to avoid any mistake or confusion. Table 4 displays details of this system. Thus, for example:

B1F1 stand for book one file one, the first file in first year middle school textbook.

B3F4 stand for book three file four, the fourth file in third year middle school textbook (3AM).

B6U3 stand for book six file three, the third unit in third year secondary school textbook (3AS).

Table 4

Labelling System

Letter/numeral	Stands for
B	Book (textbook under analysis)
F	Textbook file (for middle school books)
U	Textbook Unit (for secondary school books)
1	First year of middle school (1AM)
2	Second year of middle school (2AM).
3	Third year of middle school (3AM)
4	Fourth year of middle school (4AM).
5	First year of secondary school (1AS)
6	Second year of secondary school (2AS).
7	Third year of secondary school (3AS).

However, scanning was not without imperfections. The outcome of the scanning operation needed to be sifted in order to generate word lists.

7.3.2.1.1.3. Step 3: Generating word lists from textbooks

The third step consisted of making lists of words that occurred in each textbook. Thus, the computer files of each textbook were made into a single computer file. The list of all words occurring in each file was then generated using computer software.

The last step was refining word list for analysis. The whole process of corpus compilation and preparation for analysis can be summarized in the following table:

Table 5
Corpus compilation process

Step	Operation	Instrument/tool	Outcome
1	Scanning the books	Mustek 1248UB scanner	Machine readable text with illustrations
2	Cleansing the untreated text	Microsoft Office Word	Gross untreated text
3	Generating word lists (headwords)	<ul style="list-style-type: none"> • Lextutor 	words in each file/unit and textbook
4	Refining word lists	<ul style="list-style-type: none"> • Lextutor • Text Master • Researcher 	Refined final lists

7.3.2.1.1.4. Step 4: Refining the word lists

Once lists of words occurring in each file/unit were generated as described above using Lextutor, word families (or headwords, henceforth) were extracted using the same software. It may prove useful here to reiterate that lexical analysis often uses concepts such as tokens (a count of every word in a text), types (unique words in a text), and

word families (a headword, its inflected forms and its closely related derived forms from affixation, etc.) (Nation, 2001). The software gives a list of headwords and another one of *off-list* words. These are all the words that are either misspelt, loan words, foreign words, in short all un-English words plus non-words and proper nouns, whether English or not . The off-list words were then further sifted by the researcher.

Spelling errors were corrected, logograms, acronyms (e.g., UNESCO, CONCACAF), proper names, articles, abbreviations, months, days, names of countries, nationalities, loan words, foreign words (e.g., Inuit, Tamazight, couscous, yo-yo), rare words or oddities, etc. were removed and a list of them was made. It should be noted here that Range and Frequency program has a function whereby it can ignore a list of words made into a file called *stop-list*. This list consists of function words (Appendix 3) such as he, she, will, should, if, prepositions, numbers, and any word added to the list. Hence, the stop-list available with the programme was completed by adding to it the words removed from the lists generated from the textbooks, a list of proper names copied from the internet, metalanguage (adjective, noun, verb, etc.), and text language. The stop-list (Appendix 4) is a list comprising 14099 words and is too large to be included here except as a soft copy. Table 6 (Examples of types of words excluded from the headword lists other than function word) below displays examples of types of words excluded from the headword lists. The outcome was a refined list consisting of only headwords.

Table 6

Examples of types of words excluded from the headword lists other than function word.

Types of words	Examples
Proper names	Abdullah, Abdallah, Acapulco, Alexandria, Abraham, Adrar, Averroes, Beniizguen, Umuofia
Foreign words/ Loanword	Sushi, moussaka, Couscous, baccalaureate, canapé, garage, vinaigrette, spaghetti, bourek, chaabi, twiza, yennayer, Tamazight, tamina, soleil, shikigari
Metalanguage / text language	Adjective, adjunct, verb, adverb, paragraph, diphthong, vowel, subheading, comma, slash, colon, semicolon, etc.
Acronyms/abbreviations	BBC, a.m., bio, bros., UNESCO, USMA
Numerals	1, 2, 568, i, ii, xv, xxi

7.4. How to answer the research questions

The procedure adopted to answer the research questions object of the present study is summarized the following table (Table 7 Answering the research questions).

Table 7
Answering the research questions

Question	How to answer the question
1. What is the lexical coverage of Algerian EFL textbooks?	Compute lexical coverage
2. What is the readability level of Algerian EFL textbooks?	<p>Assess readability using text coverage. Comparing lexical coverage to research standards, if:</p> <ol style="list-style-type: none"> 1. Lexical coverage $\geq 95\%$, then <ol style="list-style-type: none"> a) high readability b) textbook is below students' reading level 2. $95\% > \text{Lexical coverage} > 75\%$, then <ol style="list-style-type: none"> a) readability is medium b) textbook at the students' reading level (instructional level) 3. Lexical coverage $< 75\%$, then <ol style="list-style-type: none"> a) readability is low b) textbook above students' reading level (frustration level)
3. Are Algerian learners learning sufficient, useful and appropriate vocabulary items?	Compare lexical coverage to three criterion lists: Academic Word List (AWL), General Service List (GSL), British National Corpus High Frequency Words (BNC-HFW)

Conclusion

This chapter presented the textbooks which were the focus of the present study and the analysis of which will be exposed in the next chapter. It provided a thorough account of the research strategy and methodology adopted for this study. This included a detailed description of the procedures for data collection, data analysis and the instrument utilized. In sum, the methodology adopted fits within corpus research and was characterised by a multi-instrument computer-based approach involving computer software.

CHAPTER 8**DATA ANALYSIS AND RESULTS****Introduction**

In this chapter, the analysis of the data will be described and the results will be presented. The data will be processed in response to the questions asked and the hypotheses stated in Chapter One. It may be useful to recall that the questions were:

Research Question 1: What is the lexical coverage of Algerian EFL textbooks?

Research Question 2: What is the readability level of Algerian EFL textbooks?

Research question 3: Are Algerian EFL learners learning sufficient, useful and appropriate vocabulary items?

The first and second research questions are interrelated as readability in this work is deduced from lexical coverage. Hence in the data analysis these two questions will be treated together. The research will endeavour to answer these questions by determining the lexical coverage and then infer the answer related to readability. Middle school textbooks will be analyzed first then, the same procedure will be followed for the secondary school textbooks. Finally, results from all textbooks will be considered together. The methodological procedure followed is that outlined in Table 3 of the previous chapter.

8.1. Lexical coverage and readability in middle school textbooks

8.1.1. Data analysis

As already mentioned, the procedure here involves computing lexical coverage and then comparing the vocabulary of a textbook to the vocabulary occurring in the textbook of the following year. Hence, in order to determine lexical coverage, it was necessary to determine at each level the percentage of words which occurred in the previous levels. Put in the form of questions this yields the following:

(Recall that AM stands for year of study in middle school. For example, 2AM stands for second year middle school, etc.)

- a- What percentage of words in 2AM textbook occurs also in 1AM textbook?
- b- What percentage of words in 3AM textbook occurs also in 1AM and 2AM textbooks?
- c- What percentage of words in 4AM textbook occurs also in 1AM and 2AM and 3AM textbooks?

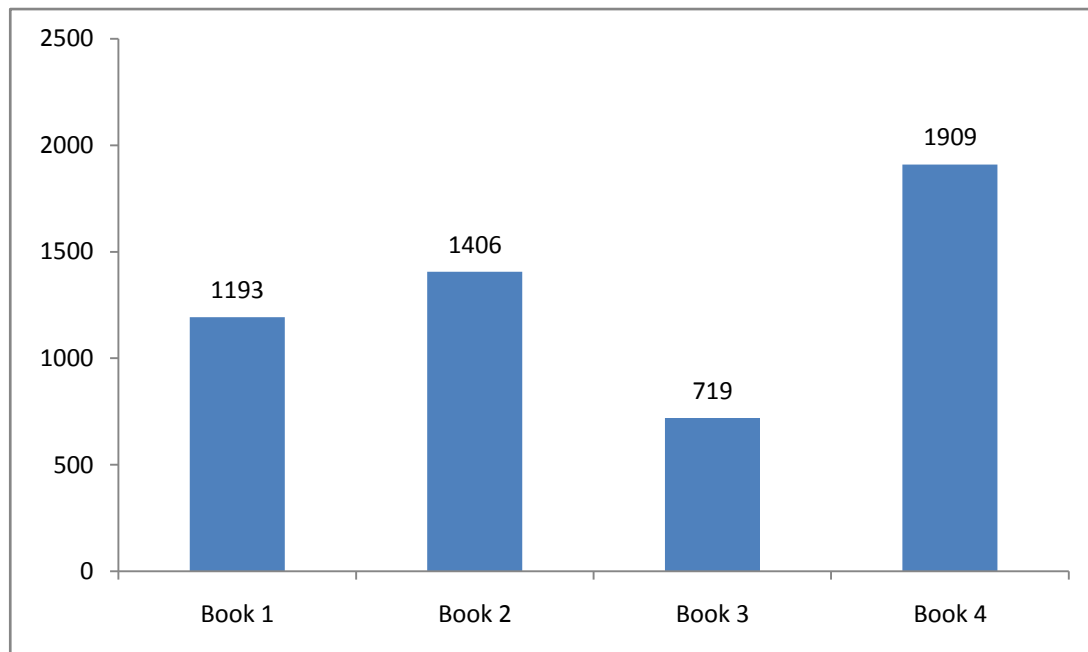
The first task in the analysis was to find out the number of word families in each textbook. Table 8 presents the distribution of word families in the middle school textbooks/levels. The number of files and the number of pages are also displayed for reference.

For ease of presentation and interpretation the figure in the table above have been made into a histogram (Histogram 1. Distribution of word families in the middle school textbooks/levels).

Table 8

Distribution of word families in the middle school textbooks

Textbooks	# files/units	# pages	# word families
Book 1: Spotlight on English Book One (1AM)	7	189	1193
Book 2: Spotlight on English Book Two (2AM)	5	125	1406
Book 3: Spotlight on English Book Three (3AM)	4	188	719
Book 4: On the Move (4AM)	6	192	1909
Total	22	694	5227



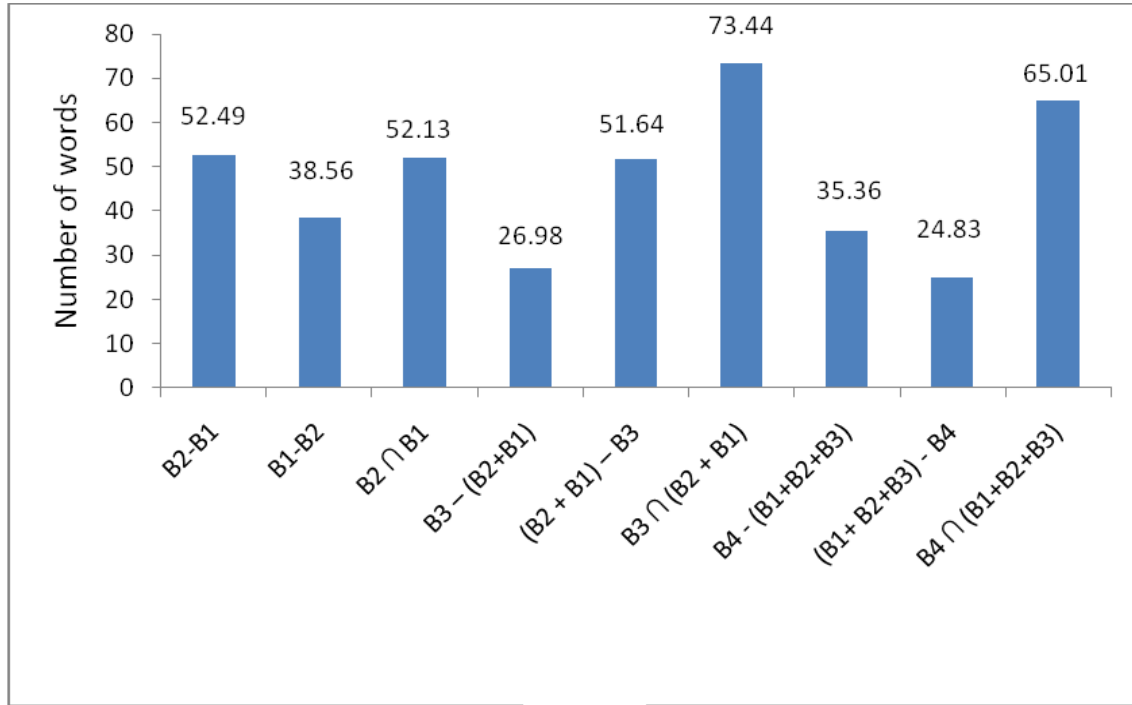
Histogram 1: Distribution of word families in the middle school textbooks/levels

The above table and histogram call for comments. Two striking facts appear from them. First, there is a rising pattern in terms of word families as we move from Book 1 to Book 2 and from Book 3 to Book 4. Second, there is a sharp drop by nearly half from Book 2 to Book 1, i.e., from second year middle school through third year middle school. The increasing number of word families across levels indicates that globally learners are exposed to more lexical items when studying at a higher level. It is predictable that more words are incorporated in the textbooks used in a higher level by textbook writers. When learners are promoted to an upper level, ideally speaking, new words are added in the textbook at the next level and old words are recycled to reinforce vocabulary learning. However, the drop in the number of headwords in Book 3 looks odd compared to the other books. To understand how many old words are common in two consecutive levels and how many new words are introduced in an upper level textbook, word lists prepared were compared. It was felt interesting to examine the common words, the newly added words, and the difference in numbers of word families in each next level of each textbook. It should be noted that in this section, the term *common words* refers to words occurring in two consecutive levels, for example, in Book One and in Book Two or in Book three and both in Books one and two. In other words, *common words* can be viewed as recycled lexical items. Worded differently, new words are words that are used in a certain level and are again included in the next one or ones. In the comparison of this study, it should also be stressed again, words refer to word families, not tokens.

Table 9 and Histogram 2 display the overall results of the middle school textbooks comparison in terms of headwords (Book One, Book Two, Book Three, and Book Four). The first column shows the books under investigation. The other column displays the number of words in each of the textbooks and not in the preceding one or ones. It also displays the intersect (percentage of common words) of each of the textbooks with the preceding one or ones (this is denoted with the mathematical symbol \cap). The results are expressed in percentage and in number of words. It should be pointed out that the percentage indicated refers to the textbook from which has been subtracted another one or ones. For example in the table below the number 52.49 in column two is the percentage of lexical coverage from Book Two. That is 52.49% of the words in Book Two did not occur in Book One.

Table 9
Overall results of middle school textbooks comparison

Books compared	Word distribution	
	%	# words
B2-B1	52.49 of B2	738/1406
B1-B2	38.56 of B1	460/1193
$B2 \cap B1$	52.13 of B2	733/1406
$B3 - (B2+B1)$	26.98 of B3	194/719
$(B2 + B1) - B3$	51.64 of (B2 + B1)	1342/2599
$B3 \cap (B2 + B1)$	73.44 of B3	528/719
$B4 - (B1+B2+B3)$	35.36 of B4	675/1909
$(B1+ B2+B3) - B4$	24.83 of (B1+ B2+B3)	824/3318
$B4 \cap (B1+B2+B3)$	65.01 of B4	1241/1909



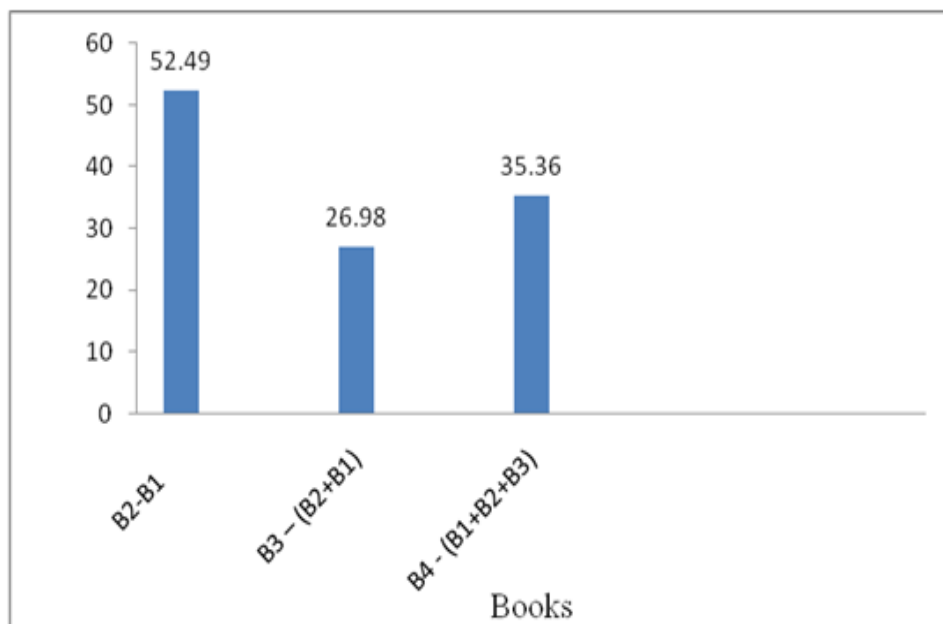
Histogram 2: Overall results of middle school textbooks comparison

A quick glance at Table 9 and Histogram 2 reveals a total absence of ‘systematicity’ in lexical coverage since the results display a jigsaw pattern as we move from one book to the other. The vocabulary load varies greatly in an unsystematic way. To present a clearer picture of this ‘unsystematicity’ vocabulary occurrence, only the difference in the number of lexical items of the four books under consideration is displayed in Table 10 and Histogram 3 below.

Table 10

New lexical items in middle school textbooks, Books 2, 3, and 4.

Books compared	% of new lexical items	
	%	# words
B2-B1	52.49 of B2	738/1406
B3 – (B2+B1)	26.98 of B3	194/719
B4 - (B1+B2+B3)	35.36 of B4	675/1909



Histogram 3: New lexical items in middle school textbooks, Books 2, 3, and 4.

This demonstrates clearly that as we move from Book 1 (1 AM) to Book 2 (2AM) and from Book 3 (3AM) to Book 4 (4AM), learners are likely to come across and have opportunity to learn far more lexical item than when we move from Book 3 (3AM) to Book 4 (4AM). Worded differently and in a much simpler way, these results can be looked at as expressing vocabulary learning opportunity, or lexical load of the textbooks.

Hence, (i) in the absence of reliable means to assess vocabulary learning, (ii) knowing that the choice of the target lexis is generally the sole decision of the teacher, and (iii) knowing that learning is idiosyncratic, one can only talk of vocabulary that is likely to be learnt and of learning opportunity.

Having said so, we can now turn to lexical coverage. Recall that in Chapter 5 lexical coverage was defined as “The percentage of running words in the text known by the readers” (Nation, 2006, p. 61), and this study adopted headwords as a unit of count instead of tokens.

The above table and histogram reveal that in middle school:

- Second year middle school students have opportunity to encounter 738 lexical new items which did not occur in the first year book and which represent a percentage of 52% of the total items occurring in the second year textbook.
- Third year middle school students have opportunity to encounter 194 lexical new items which did not occur either in first year book or in the second year book and

which represent a percentage of 26.98% of the total items occurring in the third year textbook.

- Fourth year students have opportunity to encounter 675 lexical new items which did not occur neither in first year book nor in the second year book, nor in the third year one and which represent a percentage of 35.36% of the total items occurring in the three previous textbook.

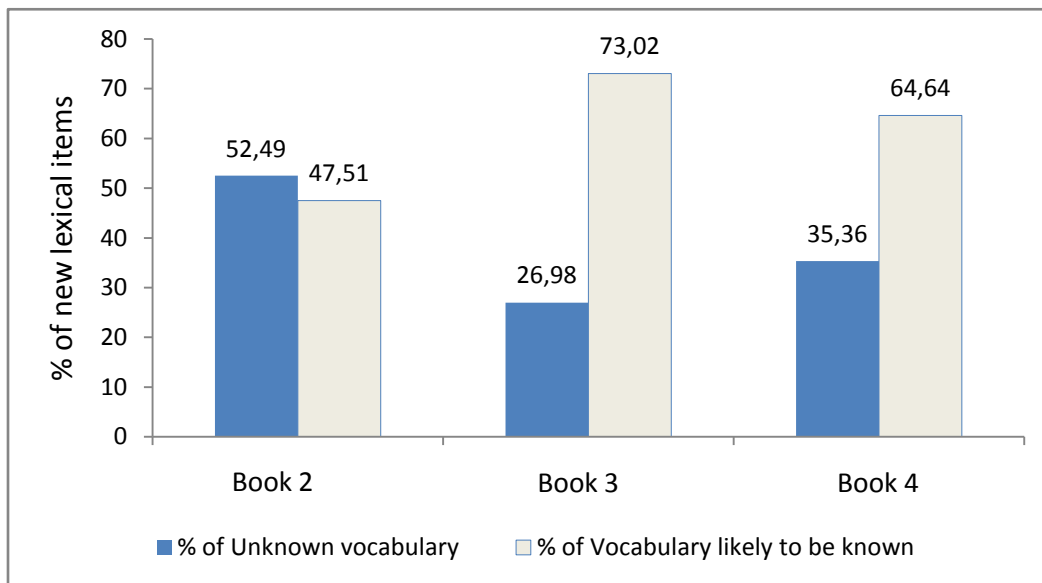
The discrepancy in the number of headwords is too conspicuous. Second and fourth year textbooks are close to each other in terms of number of new headwords with 738 and 675 lexical item respectively, whereas third year textbooks is far away with only 194 new lexical items. This trend in terms of number of lexical items should rather be inverted. The number of lexical items in a textbooks should grow larger as we move from one level to the upper level and not the other way round as it is the case in middle school textbooks.

Relying on the definition of lexical coverage, we can compute the words in the text known by the readers by subtracting the assumed known words (those occurring in the book of the previous level) from those of the current textbook. This would yield lexical coverage. The lexical coverage of the middle school textbooks is displayed in Table 11 and Histogram 4 below.

Table 11

Lexical coverage in middle school textbooks

Textbook	level	Lexical coverage (%)	
		unknown	Likely to be known
Book 2	2AM	52.49	47.51
Book 3	3AM	26.98	73.02
Book 4	4AM	35.36	64.64

**Histogram 4:** Lexical coverage in middle school textbooks

To have a clearer picture of the part of lexical coverage in each textbook, the results are displayed in the apple chart (Chart 1) and graph (Graph 1) below.

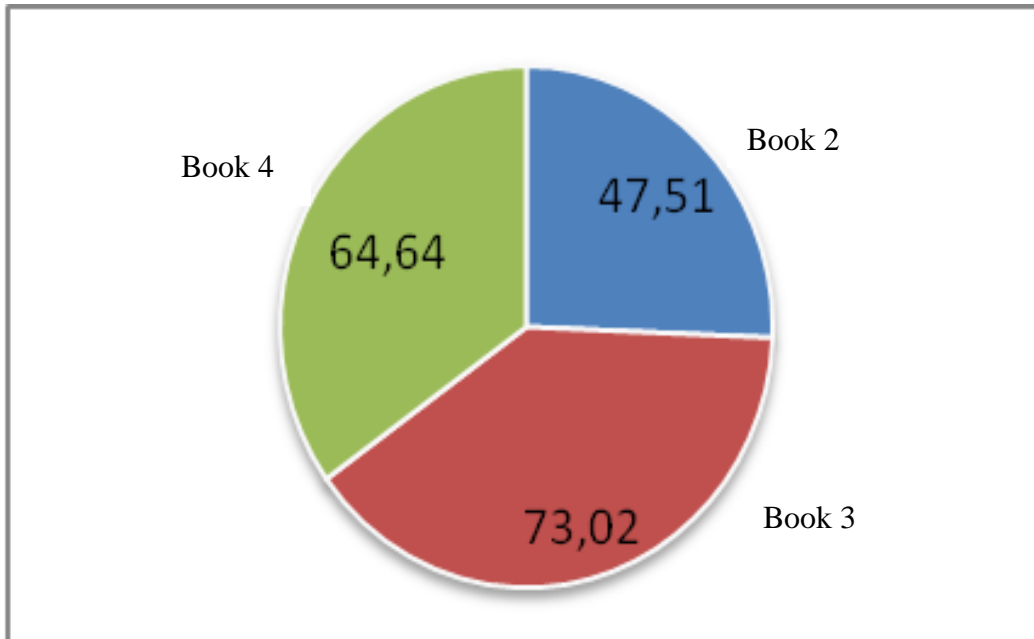
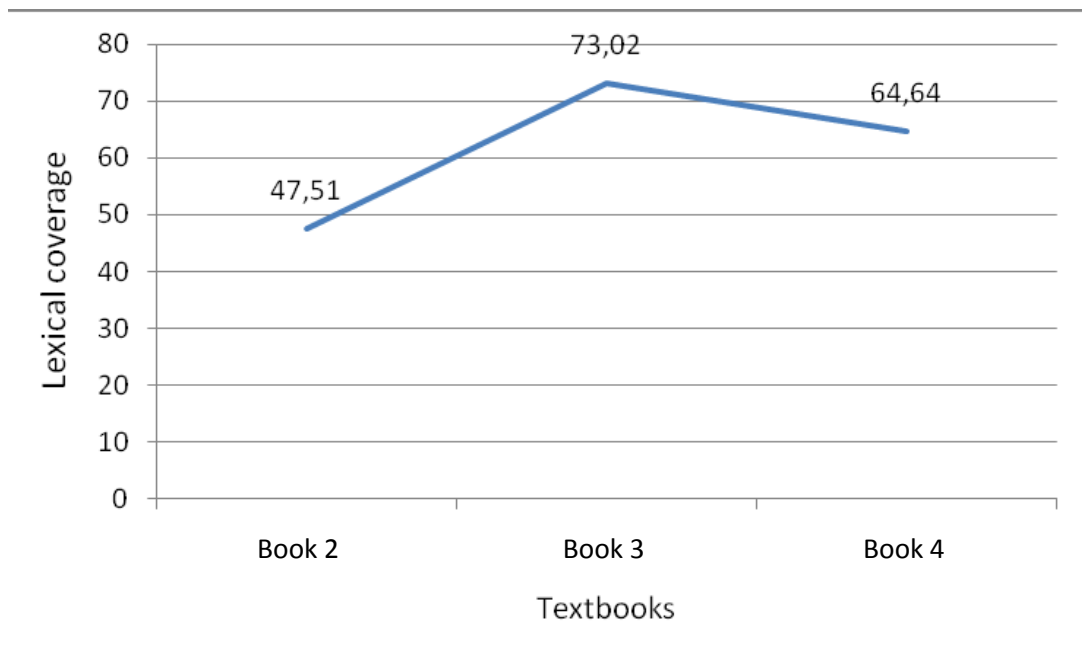


Chart 1: Lexical coverage in middle school textbooks



Graph 1: Lexical coverage in middle school textbooks

8.1.2. Interpretation

The results are interpreted in light of the hypotheses stated in Chapter One and the conditions posed in Table 7 (Chapter 7, page 258) and posited by ESL/EFL research (notably Laufer, 1997, 2010). Recall that the hypotheses were:

Hypothesis 1: the EFL textbooks used by Algerian students are **above** the students' lexical coverage. If so, then the textbook readability is **low** (reading material is **difficult**).

Hypothesis 2: The EFL textbooks used by Algerian students are at the students' lexical coverage. If so, then the textbook readability is medium (reading material is of medium difficulty).

Hypothesis 3: The EFL textbooks used by Algerian students are below the students' lexical coverage. If so, then the textbook readability is high (reading material is easy).

The conditions posed were as follows:

If a textbook has a lexical coverage equal to or higher than 95%, then the textbook is **below** students' reading level and has a **high** readability level. It is at the *independent reading level* and the comprehension rate would be 95%.

If a textbook has a lexical coverage equal to or lower than 95% and higher than 75%, then the textbook is *at students' reading level* and has a *medium* readability level. It is at the *instructional reading level*, that is the reading material contained therein is

challenging but manageable for the reader and the comprehension rate would be 90% (Cf. pp.20-21).

Finally, if a textbook has a lexical coverage below 75%, then the textbook is *above students' reading level* has a *low readability level*. This means that the reader is unfamiliar with more than one unknown word in every ten words. The texts in the textbook are considered difficult and the comprehension rate would be less than 90%. The textbook is said to be at the *frustration reading level*.

The results displayed in Table 11 and Histogram 4 above, suggest that in second year middle school lexical coverage is (47.51%), the learners are likely to be familiar with 47.51% of the lexis in the textbook; in third year lexical coverage is (73.02%); learners are likely to be familiar with only 73.02%, and in the fourth year lexical coverage is 64.64 %.

Hence, concerning middle school textbooks, **Hypothesis 2** and **Hypothesis 3** are rejected and **Hypothesis 1** is found to be valid. We can then safely infer that all three textbooks (Book 1, Book 2, and Book 3) have a **low readability**, they are **above students' lexical coverage** and consequently, they are **at the frustrational level** (reading material is **difficult**).

However, if we were to compare the three textbooks against each other, it emerges that Book 3 is the most readable followed by Book 4. Book 2 is the least readable.

If we admit the ideal or rather fallacious assumptions that:

- i) At the end of the first year a learner would know all the words contained in the textbook he had been using,
- ii) At the end of the second year, classroom lessons would have focused on half of the unknown words, i.e. a learner would know half of the lexis contained therein,

Then,

- i) Learners at the end of the second year would still be left with more than 26% (that is $52.49/2$) of unknown lexical items. Lexical coverage would be 74%.
- ii) Learners at the end of the third year would still be left with more than 13.5% (that is $26.98/2$) of unknown lexical items. Lexical coverage would be 86.5%.
- iii) Learners at the end of the fourth year would still be left with more than 17.5% (that is $35.36/2$) of unknown lexical items. Lexical coverage would be 82.5%.

Having provided the analysis of the middle school textbooks, we now turn to secondary school textbooks to investigate their lexical coverage and readability following the same procedure adopted for the middle school textbooks.

8.2. Lexical coverage and readability in secondary school textbooks

8.2.1. Data analysis

Again, in order to determine lexical coverage, it was necessary to determine at each level the percentage of words which occurred in the previous levels. However, it should be pointed out that at this stage lexical coverage in secondary school textbooks is linked to that of the middle school. For the first year secondary school it was

computed based on the middle school textbooks coverage. Put in the form of questions this yields the following:

(Recall that AS stand for year of study in secondary school. For example, 2AS stands for second year secondary school, etc.)

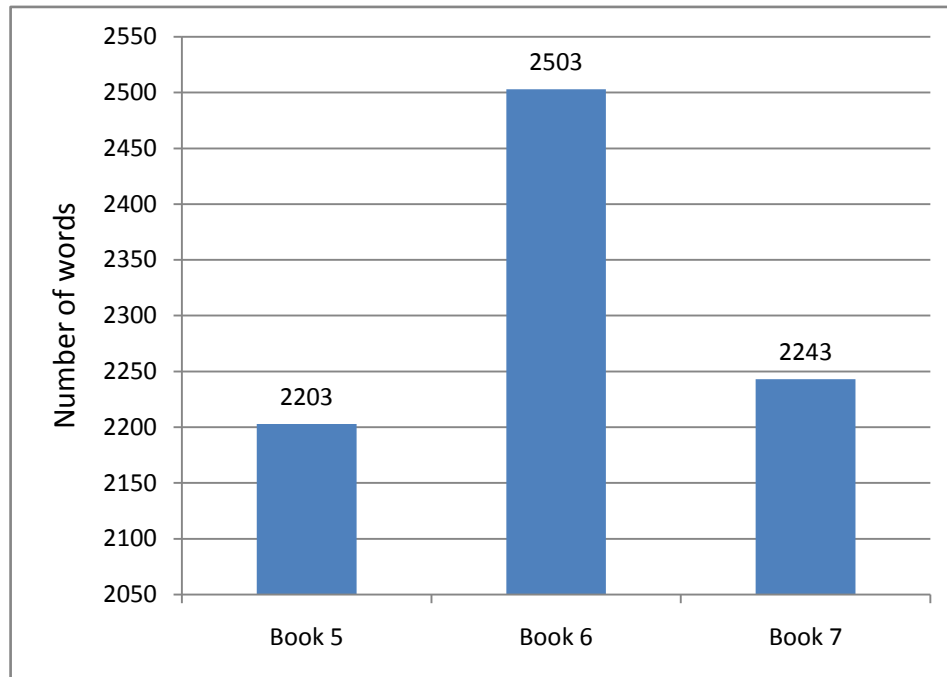
1. What percentage of words in 1AS textbook occurs in middle school textbooks?
2. What percentage of words in 2AS textbook occurs in middle school textbooks and in 1AS textbook?
3. What percentage of words in 3AS textbook occurs in middle school textbooks and in 1AS and 2AS textbooks?

Table 12 and Histogram 5 display the distribution of headwords in the secondary school textbooks. The number of units and the number of pages are also displayed for reference.

Table 12

Distribution of headwords in the secondary school textbooks

Textbooks	# units	# pages	# word families
Book 5: At the Crossroads (1AS)	5	155	2203
Book 6: Getting Through (2AS)	8	208	2503
Book 7: New Prospects (3AS)	6	272	2243
Total	19	155	6949



Histogram 5: Distribution of headwords in the secondary school textbooks.

Again, and similarly to middle school textbooks, the above table and histogram reveal that the number of headwords in secondary school textbooks is distributed unequally. The number of headwords in Book 6 (Getting Through, 2AS), is outstanding. The sharp rise in the number of headwords from first year to second year then the sharp fall from second year to third year is noticeable. It is also noticeable that the number of headwords in Book 5 and Book 6 is almost the same. It is predictable and understandable that Book 6 contains more headwords than Book 5 as it is one level beyond the former. But this is not the case when we move from book 6 to Book 7. The

drop in number of words Book 6 to Book 7 is sharp and unexpected. Similarly to the middle school textbooks, the headword lists extracted from secondary school textbooks were compared. Table 13 and Histogram 6 display the comparison of headwords distribution in the secondary school textbooks.

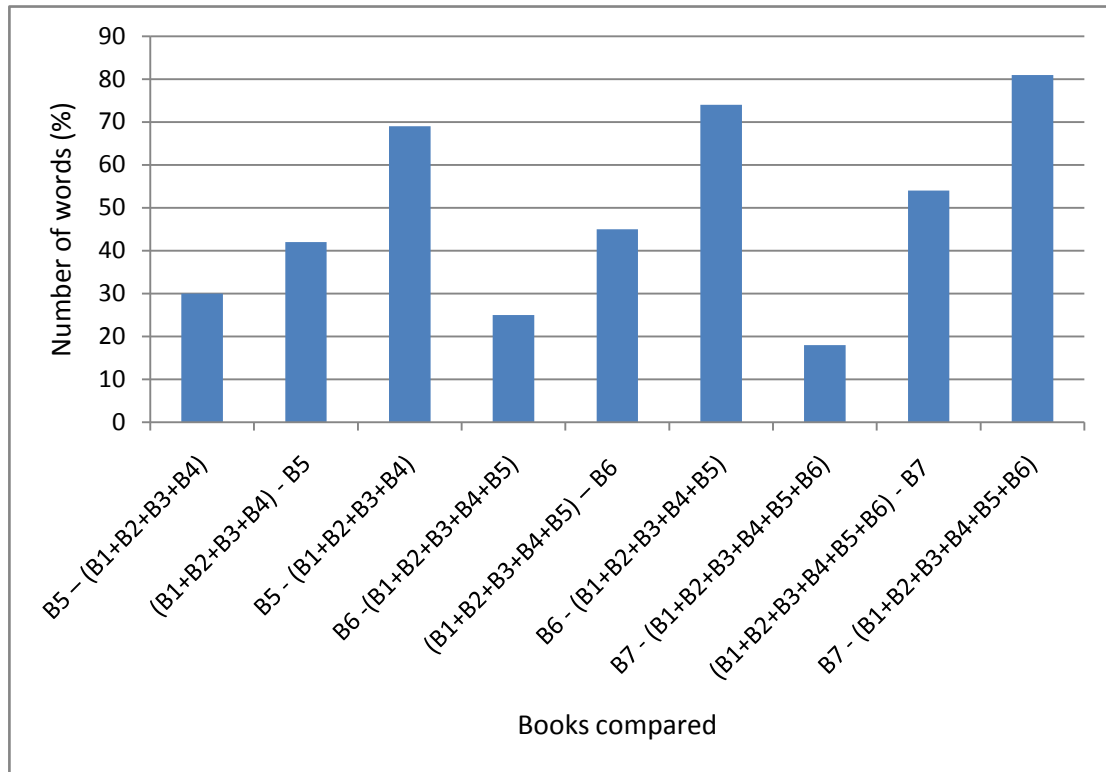
Table 13 and Histogram 6 display the comparison of headwords distribution in the secondary school textbooks.

Table 13

Overall results of secondary school textbooks comparison

Textbooks	Word distribution	
	%	# words
B5 – (B1+B2+B3+B4)	30	675
(B1+B2+B3+B4) - B5	42	1175
B5 - (B1+B2+B3+B4)	69	1564
B6 -(B1+B2+B3+B4+B5)	25	646
(B1+B2+B3+B4+B5) – B6	45	1552
B6 - (B1+B2+B3+B4+B5)	74	1867
B7 - (B1+B2+B3+B4+B5+B6)	18	412
(B1+B2+B3+B4+B5+B6) - B7	54	2220
B7 - (B1+B2+B3+B4+B5+B6)	81	1842

Note: the percentage in column 2 is computed out of the number of headwords in the book(s) from which has been subtracted the number of headwords in the other book(s). Thus, the figure 30% is out of B5.



Histogram 6: Overall results of secondary school textbooks comparison

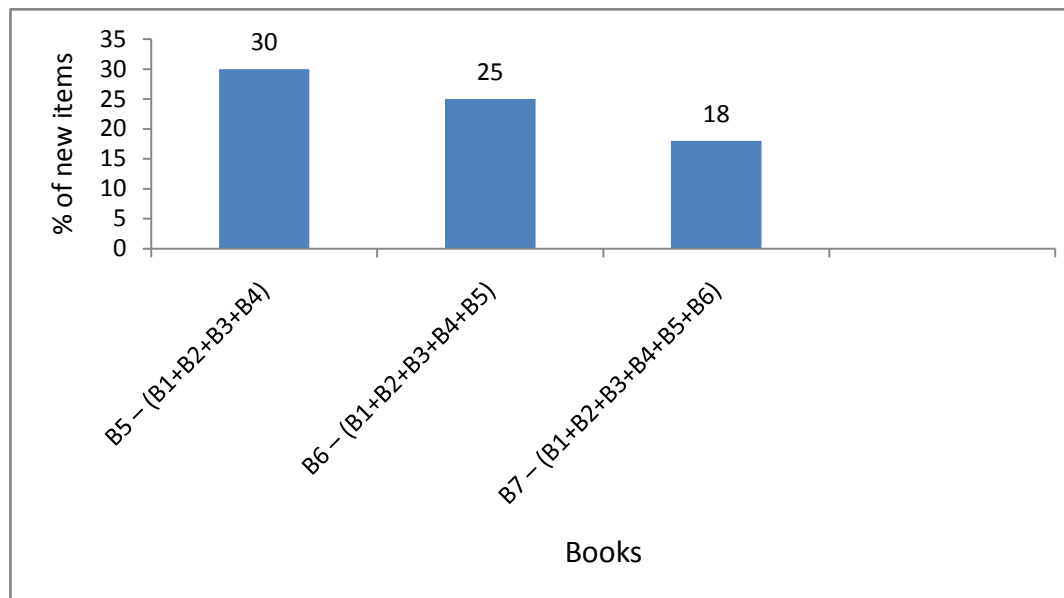
Table 13 and Histogram 6 reveal the existence of ‘systematicity’ in lexical coverage since the results display a linear increase in lexical coverage as we move from one secondary school textbook to the other. To present a clearer picture of this ‘systematicity’ of vocabulary occurrence, only the difference in the number of lexical items of the three books under consideration is displayed in Table 14 and Histogram 7 below.

Table 14

New lexical items in secondary school textbooks

Textbooks compared	Word distribution	
	%	# words
B5 – (B1+B2+B3+B4)	30	675
B6 – (B1+B2+B3+B4+B5)	25	646
B7 – (B1+B2+B3+B4+B5+B6)	18	412

Note: the percentage in column 2 is computed out of the number of headwords in the book(s) from which has been subtracted the number of headwords in the other book(s). Thus, the figure 30% is out of B5.



Histogram 7: New lexical items in secondary school textbooks: Books 5, 6, and 7 expressed in percentage of words

The table and histogram above reveal that:

- First year secondary school students have the opportunity to encounter 675 lexical new items which did not occur in the first year book and which represent a percentage of 30% of the total items occurring in the previous textbooks.
- second year students have the opportunity to encounter 646 lexical new items which did not occur in any of the previous textbooks and which represent a percentage of 25% of the total items occurring in the third year textbook.
- Third year students have the opportunity to encounter 412 lexical new items which did not occur neither in first year book nor in the second year book,

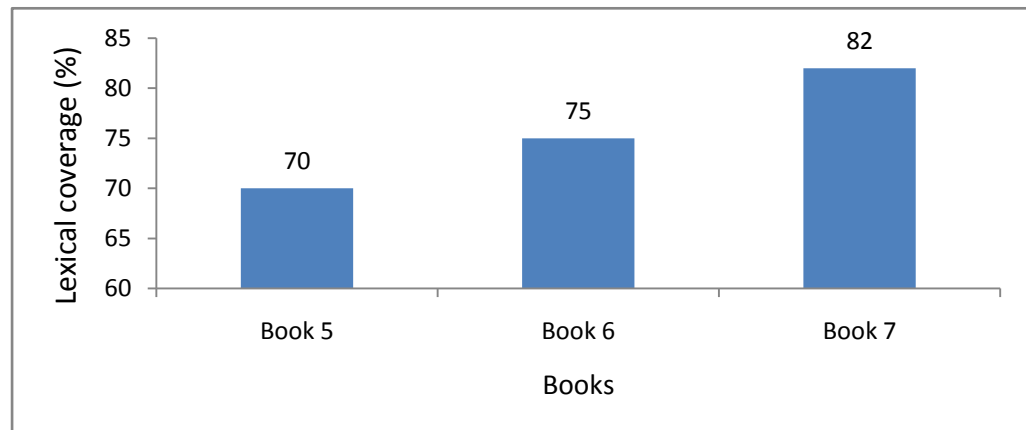
nor in the third year one and which represent a percentage of 18% of the total items occurring in all the previous textbooks.

It is evident that the number of new headwords in secondary school textbooks is in inverse variation. The higher the level the fewer headwords contained in the textbook. As we move from first to second to third year vocabulary learning opportunities decrease more and more and learners are likely to come across and have opportunity to learn fewer and fewer headwords.

From the data above we can compute lexical coverage following the same procedure adopted for the middle school textbooks. The lexical coverage of the secondary school textbooks is displayed in Table 15 and Histogram 8 below.

Table 15
Lexical coverage in secondary school textbooks

Textbook	level	Lexical coverage (%)	
		unknown	Likely to be known
Book 5	1AS	30	70
Book 6	2AS	25	75
Book 7	3AS	18	82



Histogram 8: Lexical coverage in secondary school textbooks.

Again, the results are interpreted in a way similar to that of the middle school textbooks.

8.2.2. Interpretation

The results displayed in Table 15 and Histogram 8 above suggest that in first year secondary school, lexical coverage is 70% (the learners are likely to be familiar with 70% of the lexis in the textbook), in second year lexical coverage is (75%), the learners are likely to be familiar with 75%, and in the third year the lexical coverage is 82 %.

Hence, concerning secondary school textbooks, **Hypothesis 2** and **Hypothesis 3** are rejected and **Hypothesis 1** is found to be valid. We can safely infer that all three textbooks (Book 5, Book 6, and Book 7) have a **low readability**, they are **above students' lexical coverage** and consequently, they are **at the frustrational level** (reading material is **difficult**).

However, if we were to compare the three textbooks against each other, it emerges that Book 7 is the most readable followed by Book 6. Book 5 is the least readable.

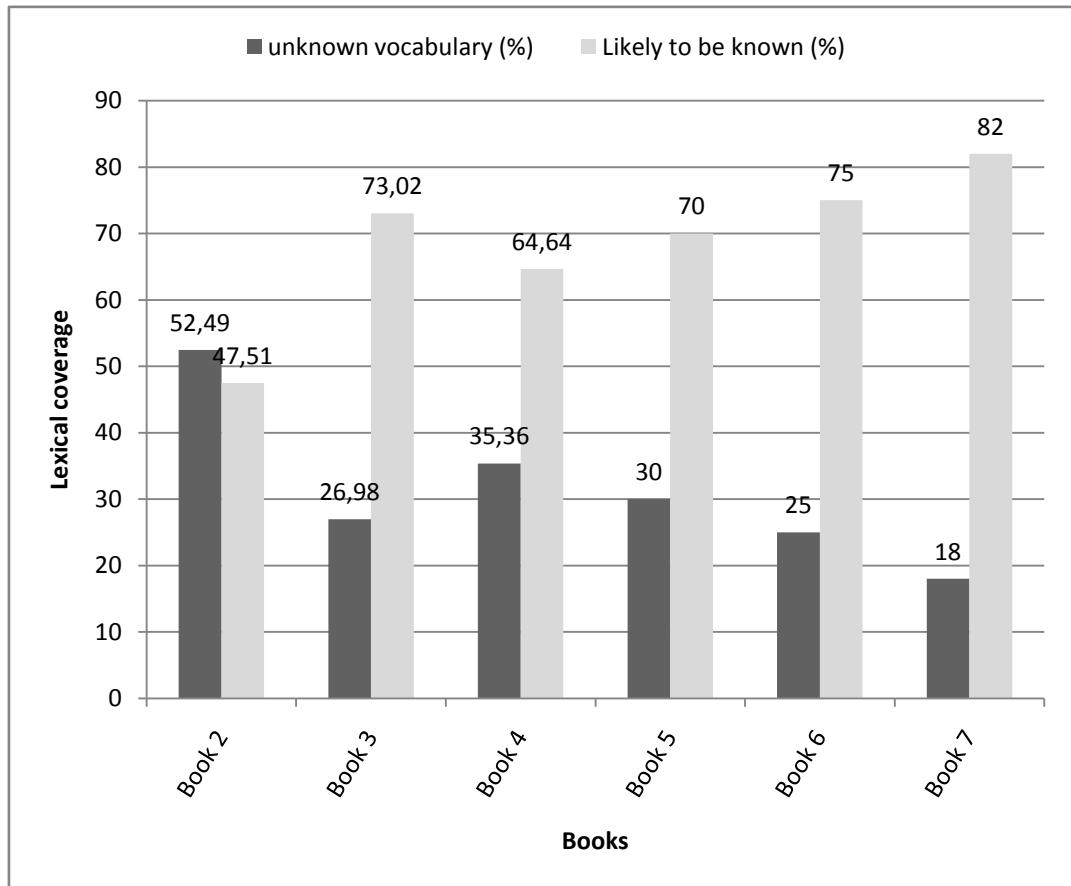
8.3. Putting it all together: Lexical coverage and readability in middle and secondary school textbooks

Armed with the outcome from the analysis of the 7 textbooks we can now put everything together to draw conclusion as regard the 7 books and answer the main research question which guided the whole study (Cf. page 13): *Are textbooks used by Algerian EFL students at, above, or below students' lexical coverage?*

The lexical coverage for the seven books put together in displayed in Table 16 and Histogram 9 below.

Table 16
Lexical coverage in all textbooks

Textbook	level	Lexical coverage (%)	
		unknown	Likely to be known
Book 2	2AM	52.49	47.51
Book 3	3AM	26.98	73.02
Book 4	4AM	35.36	64.64
Book 5	1AS	30	70
Book 6	2AS	25	75
Book 7	3AS	18	82



Histogram 9: Lexical coverage in all textbooks

results are displayed in the apple chart and graph below.

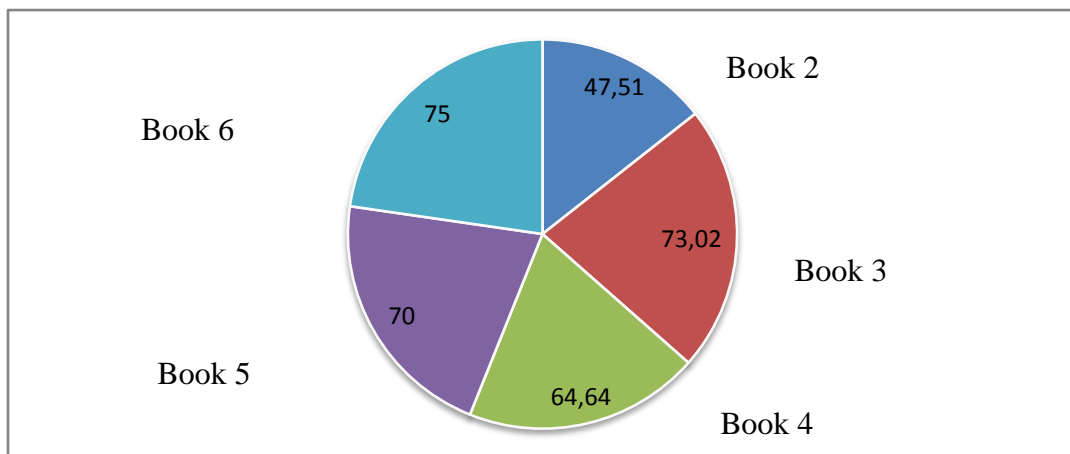
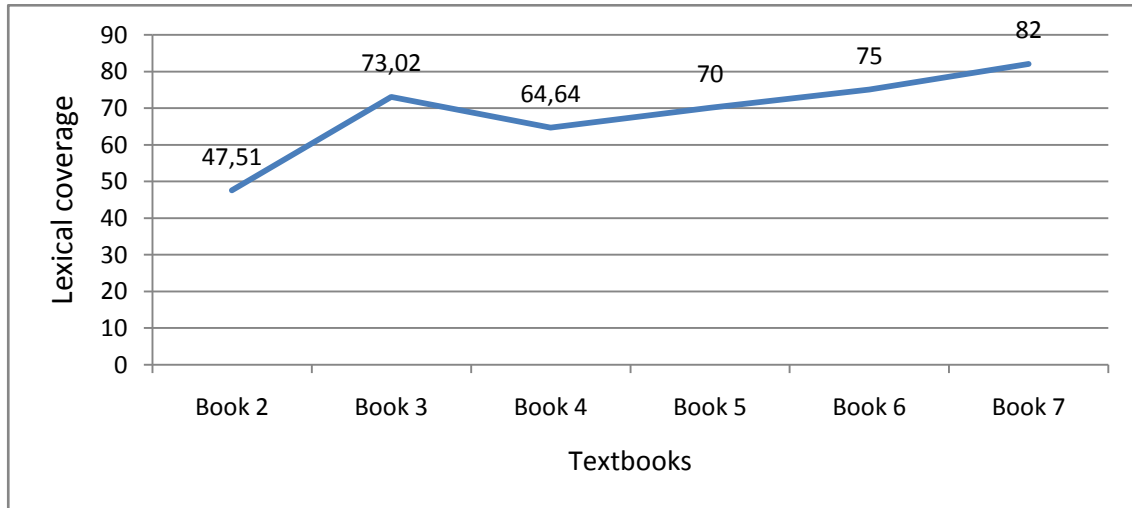


Chart 2: Lexical coverage in all textbooks



Graph 2: Lexical coverage in all textbooks

The progression of lexical coverage through the seven books is made clear by the above diagrams. We can see that the textbook with the lowest text coverage, and so which is the least readable, is that of second year middle school (Book 2). The textbook with the highest text coverage, and so which is the most readable, is that of third year secondary school (Book 7). We can also notice (as already pointed out) a regular progression of text coverage in secondary school textbooks.

Another interesting fact revealed by the above diagrams is that there is a noticeable regular increase in text coverage of about 5% from Book 4 through Book 7. If the third year middle school textbook (Book 3) had followed this rate, the progression through all the books would have been regular. That is when students move from one level to the other, the rate of vocabulary available for learning is 5% of the total items contained

in the textbook they use. Had the rate of increase been the same for Book 3, the progression of lexical coverage through the seven books would have been linear.

The lexical coverage of middle school second year (Book 2) which is 47.51% reveals that when students move from the very beginning level (Level zero) to the next one they may be overwhelmed by the number of words they have opportunity to encounter, and likely to learn.

However, we can still infer that in the ideal conditions secondary school textbooks are by far more readable than middle school textbook. The reverse would have been of much better help to learners as they would move from the most readable to the least readable as the results of the study seem to suggest.

The linear progression of text coverage in secondary school textbooks can be considered as a 'positive' point as this progression is in terms of increasing likelihood of known vocabulary. Worded differently, the percentage of likely to be known vocabulary in third year secondary school textbook (Book 7) is higher than of second year (Book 6) which in turn is higher than in first year (Book 5). Apparently there is a problem in second year middle school (Book 2) where student are likely to encounter more lexical items than in any other book, especially with regard to the book they had been using in first year middle school. Learners moving from a beginning level (1AM) find themselves in front of a flood of new lexical items they had never encountered before. This would make it a task for them and for the teacher to cope with such a lexical coverage.

Considering the results displayed in Table 16 and the following figures which summarize all the results of the data analysis as regard lexical coverage, and in response to the main research question posed we can conclude the following:

Textbooks used by Algerian EFL students are **above** students' lexical coverage.

If we want to rank them in order of lexical coverage, we would get the results displayed in Table 17 below.

Table 17
Rank order of textbooks by lexical coverage

Textbook	level	Lexical coverage (%)	
		unknown	Likely to be known
Book 7	3AS	18	82
Book 6	2AS	25	75
Book 3	3AM	26.98	73.02
Book 5	1AS	30	70
Book 4	4AM	35.36	64.64
Book 2	2AM	52.49	47.51

Unexpectedly, Book 2 (2AM) is the textbook with the least lexical coverage. The textbook with the highest lexical coverage is the one in the series, Book 7 in use in the year of the secondary school. Then comes the textbook of the preceding year (2AS) followed by Book 3 in use in third year middle school. After that we find the first year

secondary school textbook, the fourth year middle school textbook, and in the last position the second year middle school textbook.

To sum up, by and large if a student in an Algerian English class knows all the vocabulary contained in the textbook in use within an academic year, he/she would come across too many lexical items. However, as already pointed out it is well known that in any textbook learners are not expected to learn all the lexical item contained therein and research has shown that it is necessary to encounter a word in a variety of contexts a number of times, at regular intervals, in order for the learner to have a realistic chance of learning the word (see page 213). It would then be interesting to investigate the frequency of vocabulary repetition throughout all the textbooks, and also of equal importance to know whether the Algerian learners have opportunity to encounter and then are likely to learn is the ‘right’ vocabulary. By ‘right vocabulary, it is meant the vocabulary contained in the General Service List (GSL), the Academic word List (AWL), and The British National Corpus – High Frequency Words (BNC-HFW) (Cf. pages 216-219). Hence, the task now is to answer the third research question: *Do Algerian EFL textbooks provide sufficient, useful and appropriate vocabulary items vocabulary?*

It has already been mentioned that such a question can be answered by comparing the headwords contained in all the textbooks to the criterion lists, namely: the Academic Word List (AWL), the General Service List (GSL), and the British National Corpus – High Frequency Words (BNC-HFW).

8.4. Do Algerian EFL textbooks provide sufficient, useful and appropriate vocabulary items?

There are hundreds and thousands of words in the English language. Around 114,000 word families excluding proper names in Webster's Third New International Dictionary were counted in a study (Goulden, Nation and Read, 1990). This huge number is well beyond the goals of most first language learners, let alone second/foreign language students. To teach vocabulary to ESL/EFL learners, teachers should be aware of the selection of vocabulary items to be taught. However, if a textbook is chosen and closely followed in teaching, the selection of vocabulary heavily depends on the textbook lexical coverage. In order to examine a textbook in terms of appropriateness of lexical choice, an attempt was made to answer the following questions: Do Algerian EFL textbooks provide sufficient, useful and appropriate vocabulary items? The objective is to find out whether Algerian EFL textbooks provide learning opportunities of the vocabulary which is used most frequently and which they will use in their future life, particularly for an academic purpose. It is axiomatic that some words can be used in a wide variety of circumstances while others have limited use. Therefore, teaching useful vocabulary before less useful vocabulary gives learners the best reward for their learning effort. However, how do we define "usefulness"? One measure of usefulness is word frequency, that is how often the word occurs in normal use of the language. Nation (2003) provided a practical description:

the most useful vocabulary that every English language learner needs whether they use the language for listening, speaking, reading, or writing, or whether they use the language in formal and informal situations, is the most frequent 1,000 word families of English. (pp. 135-136)

Nation further explained that these 1,000 word families are so important and useful that they cover around 75 percent of the running words in academic texts and newspapers, over 80 percent of the running words in novels and about 85 percent of running words in conversation. In other words, these 1,000 words families considerably help learners' communication. If learners know these words, they will know a very large proportion of the running words in either a spoken or a written text.

The next most useful list is the second 1,000 words of English. The classic list of high-frequency words is Michael West's *General Service List* which contains 2,000 word families. Before the 2,000 high frequency words, the most useful vocabulary depends on what the learners intends to use English for, that is the goals of the learner. If a learner intends to go on to academic study in a university, then *Academic Word List* (Coxhead, 2000) is the next most useful vocabulary. AWL is a list of 570 word families that occur frequently in wide range of academic texts, satisfying learners' needs for general academic vocabulary. Hence, the question: *Do Algerian EFL textbooks provide sufficient, useful and appropriate vocabulary items?* The last part of the textbooks analysis was to examine the percentage of words included in the *General Service List (GSL)* and the *Academic Word List (AWL)*.

The headwords contained in all the textbooks under investigation were compared to these lists in order to determine lexical coverage of the textbooks as compared to these lists. The instrument used was Range and Frequency programs for Windows described in Chapter 7 (see under 7.3.1.).

All the headwords in each of the seven textbooks were made into a single file (see Appendix 5, List of all headwords in all textbook,) Their number was 4434. They were then analyzed by the computer programmes. First, the list was processed using Range and the results were crosschecked using Lextutor.

8.4.1. Results

Table 18 below displays the results yielded by RANGE and Figure 7 is a snapshot of the results displayed by Lextutor.

Table 18

Results yielded by RANGE

Processing file: C:\Documents and Settings\Admin\LocalSettings\Temp\Rar\$EX10.641\INDO1.txt
 0001000, 0002000, 0003000, 0004000,
 Number of lines: 4434
 Number of words: 4434
 Reading: C:\DOCUME~1\Admin\LOCALS~1\Temp\Rar\$EX10.641\BASEWRD1.txt
 Reading: C:\DOCUME~1\Admin\LOCALS~1\Temp\Rar\$EX10.641\BASEWRD2.txt
 Reading: C:\DOCUME~1\Admin\LOCALS~1\Temp\Rar\$EX10.641\BASEWRD3.txt

WORD LIST	TOKENS/%	TYPES/%	FAMILIES
one	849/19.16	849/19.16	801
two	854/19.27	854/19.27	832
three	472/10.65	472/10.65	449
not in the lists	2256/50.91	2256/50.91	????
Total	4431	4431	

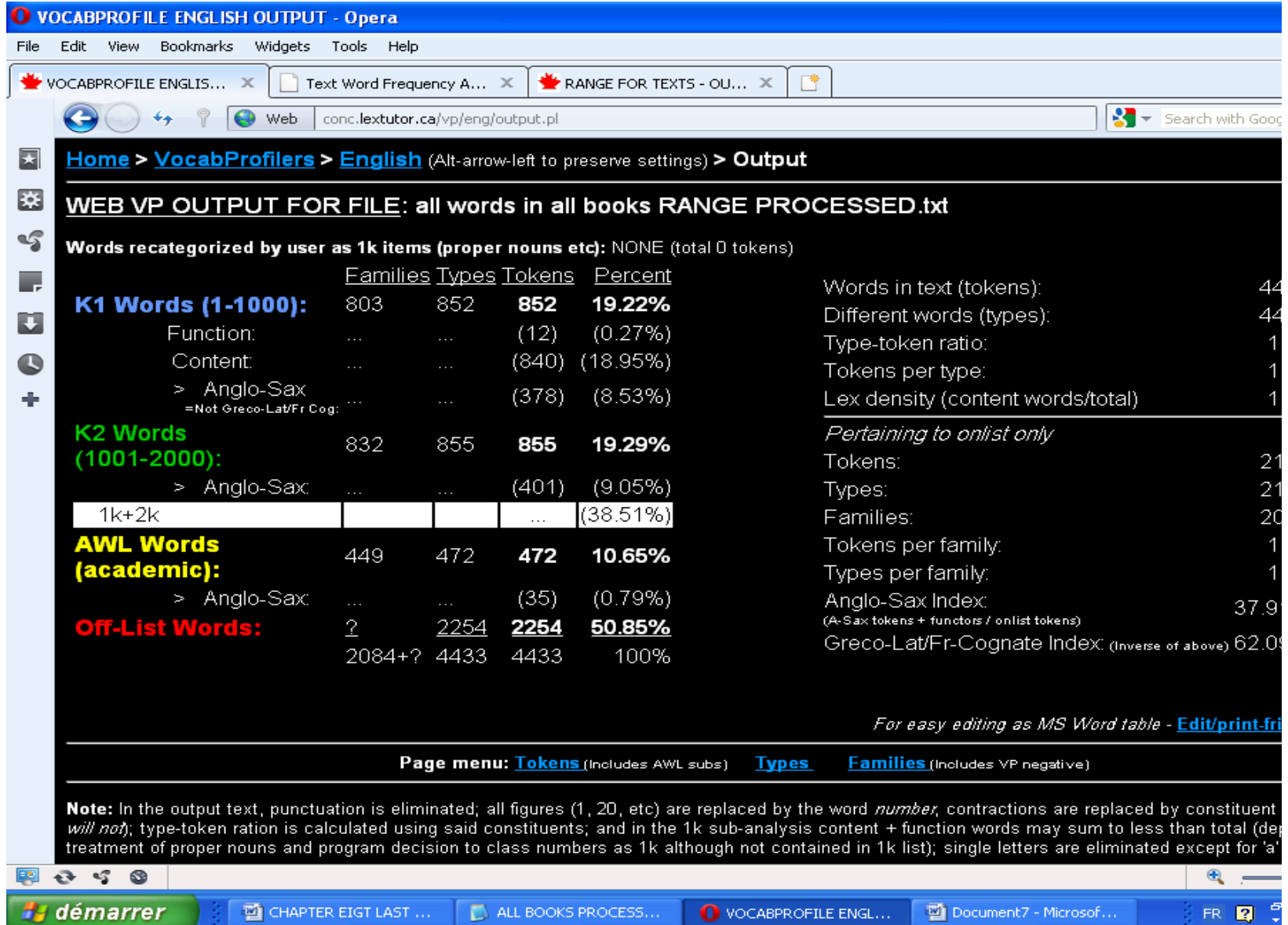


Figure 7: Snapshot of the results displayed by Lextutor

In the table above and figure:

- Word list '*one*' refers to the first 1000 most frequent word in English appearing in the General Service List (discussed in Chapter 6) and which correspond to the *K1* Level of Lextutor.
- Word list '*two*' refers to 2000 most frequent word in English appearing in *the General Service List*, and which correspond to the *K2* Level of Lextutor.
- Word list '*three*' refers to the number of words in the *Academic Word List* appearing in Figure 7 under the heading '*AWL Words (Academic)*'.
- *not in the lists* (Table 18) and *Off-List Words* in Figure 7 refer to a list of words which do not appear in any of either the *General Service List* or the *Academic Word List*.

Results from both software are displayed in Table 19.

	Results by Range		Results by Lextutor	
	# words	%	# words	%
first 1000 most frequent word /K1	849	19.16	853	19.22
first 2000 most frequent word /K2	854	19.27	855	19.29
The 3000 most frequent words/K1+K2	1701	38.43	1708	38.51
Academic /AWL Words	472	10.65	472	10.65
not in the lists/ Off-List Words	2256	50.91	2254	50.85

The number of words from the first -1000 list of most frequent words in English occurring in the seven Algerian textbooks is around 850 (see Appendix 6). The number of words from the second-1000 list of most frequent words in English occurring in the seven Algerian textbooks is also around 850 (see Appendix 7). The number of words from the Academic Word List occurring in the seven Algerian textbooks is 472 (see Appendix 8).

It appears from the above table that the results displayed by the two software used are rigorously the same. We can notice a very slight negligible discrepancy in the figures and which pertains to one hundredth, for example one result is 19.16% for one software and 19.22% for the other.

The above results can read by rounding the nearest whole number as follows:

Words in the 1000 Most Frequent Word Family = 19%

Words in the 2000 Most Frequent Word Family = 19 %)

Number of words in the Academic Word List = 11%)

Total number of words not in either list = 60%

8.4.2. Interpretation

The above results speak by themselves as they show an unquestionable proof that the textbooks under investigation are far away from required level of vocabulary.

Throughout seven year of study of English in a formal setting, Algerian EFL learners are likely to encounter only 19% of the first thousand most frequent words of English, and 19% of the second thousand of most frequent words. This adds up to and yields a lexical coverage of 38%. Algerian EFL learners are likely to encounter 38% of the 2000 most frequent words of English. Then even in the 'ideal' situation with an 'ideal' student who would learn all the 38% of words discussed above, this student would still miss 62% of the most frequent words.

As for the Academic Word List, the lexical coverage is 11%. This means that an Algerian EFL learner has opportunity / is likely to learn 11% of the words appearing in the Academic Word List. An Algerian student entering university would never have encountered and have opportunity to learn 89% of the useful vocabulary needed for academic studies.

The last column of Table 19 summarises the picture. By and large, a student attending school in the Algerian educational system has opportunity/is likely to learn only 40% of the sufficient, useful and appropriate vocabulary advocated by research in English language pedagogy. 60% of such vocabulary remain out of the learners' reach

knowing that the textbooks constitute for the overwhelming majority of students the sole source of vocabulary.

Conclusion

This chapter was devoted to the data analysis and findings. The data gathered by means of the various instrument were analysed in light of the hypothesis stated and the research questions posed. The strategy adopted was to analyse middle school textbooks to determine their lexical coverage and then the same procedure was followed for the secondary school textbooks. Finally, results from both sets of books were put together to complete the analysis of lexical coverage and readability. Furthermore, another issue, viz., the usefulness and appropriateness of the vocabulary occurring in the Algerian EFL books was investigated. The results were interpreted in the light of the hypotheses stated and the questions posed. It remains to answer the classic question: so what?

The following chapter presents an overall view of the present dissertation, exposes its findings, implications, and limitations. Finally, Recommendations for Future Research are made.

CHAPTER 9

FINDINGS, IMPLICATIONS, LIMITATIONS, AND RECOMMENDATIONS

Introduction

This concluding chapter is a ‘putting it all together’ chapter. Its aims at considering all the findings of the study in order to draw the pedagogical implications for teachers using the textbooks analyzed in the study, educational authorities, textbook writers, researchers and university English teachers. In additions the limitations of the study are stated and recommendations for future research are made. Before doing so, an overall picture of the dissertation is in line here.

9.1. The overall picture of the thesis

The impetus for the conduct of this research stemmed mainly from the researcher’s personal experience as an EFL teacher. As a field practitioner I have so many times witnessed and experienced what, to paraphrase Laufer (1997), I will call the ‘lexical plight’. In fact, one of the most recurring problems Algerian EFL teachers face daily is to know whether a particular piece of writing is likely to be comprehensible to their students. A daily concern for frontline teachers s the question: *is the level of vocabulary in this reading passage appropriate for my learners?* In the absence of reliable means to assess their students’ knowledge of vocabulary, those teachers rely inescapably on their educated guess for their decisions on the comprehensibility of a particular piece of reading material for their learners. The selection of an appropriate

reading passage is critical. If the passage chosen is inappropriate for whatever reason, the chances that learners understand that particular passage are substantially jeopardized. The problem raised here is then whether the textbooks in use for teaching and learning English in the Algerian schools are matched to learners' level and whether teachers are equipped with tools that help them gauge their textbooks difficulty level. Therefore, there was an imperative need of investigating the readability of the Algerian EFL textbooks.

There is in the literature a plethora of studies which investigated textbook readability. Similarly, many studies have investigated lexical coverage. However, to the best of the researcher's knowledge, no study has ever focussed on readability relying on lexical coverage. The present one advocates a novel approach to the study of readability. It is then a first attempt to assess readability via lexical coverage. In fact, the literature reviewed has revealed that most, if not all, researchers agree that vocabulary knowledge has a great effect on L2 reading comprehension (and hence readability). Many others feel that there is a threshold of vocabulary knowledge below which FL reader cannot achieve an adequate comprehension of written texts. Many studies have been conducted to investigate this issue. Laufer (1989b) found that readers need to know at least 95% of the words in a text for adequate comprehension of English academic texts. Hirsh and Nation (1992) determined that an un-simplified text can be comprehended when 95% of words are known. This acts as a lexical threshold below which L2 reader might not adequately comprehend the text. In 2000, Hu & Nation

concluded that 98% of the words given in a reading text for pleasure are needed for adequate unassisted comprehension. More recently, Laufer, and Ravenhorst- Kalovski (2010), Matsuoka, W. & Hirsh, D. (2010), and Laufer, and Waldman (2011) pointed this out. Building upon on such findings the present study relied on student's lexical coverage to determine the readability level of textbooks, knowing that readability depends primarily on vocabulary knowledge and that it is a function of lexical coverage.

The theoretical foundation is laid upon the long-standing empirical tradition of the readability of text by examining its linguistic characteristic. Hence, the need of review of the literature related to the investigation of text readability, text difficulty and corpus linguistics. Such a tradition holds that learning vocabulary is the fundamental step to learn a foreign language and, as already mentioned, that vocabulary knowledge is the single best predictor of reading comprehension.

Thus, the primary purpose of the study was to assess the readability of the materials used to teach Algerian learners to read in English through lexical coverage. The research was conducted to examine if the readability level of current textbooks is within the range of the students' lexical coverage. The ultimate objective was to answer the main research question: *Are textbooks used by Algerian secondary school students **at, above, or below** students' lexical coverage?* In the course of answering this main question there was a set of other questions that this research asked in order to predict the textbook readability based on lexical coverage. These were:

1. What is the lexical coverage of Algerian EFL textbooks?
2. What is the readability level of Algerian EFL textbooks?
3. Do those textbooks provide sufficient, useful and appropriate vocabulary items?

The hypotheses to be proved or disproved were:

H1: The EFL textbooks used by Algerian students are **above** the students' lexical coverage. If so, then the textbook readability is **low** (reading material is **difficult**).

H2: The EFL textbooks used by Algerian students are **at** the students' lexical coverage. If so, then the textbook readability is **medium** (reading material is of **medium difficulty**).

H3: The EFL textbooks used by Algerian students are **below** the students' lexical coverage. If so, then the textbook readability is **high** (reading material is **easy**).

As for the research methodology, in order answer the research questions, it was necessary to analyze the seven English as a Foreign Language textbooks in use in the Algerian schools. Such a task required necessarily the use of computer software. Hence it was made use of freely available computer programmes. Another computer programme was developed specially for this study with the collaboration of colleagues from the Department of Computer Science (Ferhat Abbas University, Setif).

Armed with such tools, the task of data collection and analysis began. The data collection was conducted in two steps: the first step consisted in corpus compilation and the second step consisted in the study of the corpus with software.

The study relied on four computer software for the data analysis two of which are vocabulary profilers. Vocabulary profiling is a measure of the proportions of low and high frequency vocabulary used in a written text. In addition to frequency information, a profiler designed for lexical analysis of texts often provides other information such as the presence/absence of the set of words from the input text in other specialized word lists.

A vocabulary profiler divides the words of any text into four categories by frequency: (1) the most frequent 1000 words of English (level K1), (2) the second most frequent thousand words of English, i.e. 1001 to 2000 (level K2), (3) the academic words of English (the AWL, 570 words that are frequent in academic texts across subjects), and (4) the remainder which are not found on the other lists (off-list). In other words, a vocabulary profiler measures the proportions of low and high frequency vocabulary used by a native speaker or language learner in a written text.

One of the software used was 'Range and Frequency programs for Windows based PCs' which was developed by Paul Nation and Alex Heatley (2002) of the Victoria University of Wellington and is freely downloadable. RANGE is used to compare the vocabulary of up to 32 different texts at the same time. For each word in the texts, it provides a range or distribution figure (how many texts the word occurs in),

a headword frequency figure (the total number of times the actual headword type appears in all the texts), a family frequency figure (the total number of times the word and its family members occur in all the texts), and a frequency figure for each of the texts the word occurs in. It can be used to find the coverage of a text by certain word lists, create word lists based on frequency and range, and to discover shared and unique vocabulary in several pieces of writing.

RANGE can also be used to compare a text against vocabulary lists to see what words in the text are and are not in the lists, and to see what percentage of the items in the text are covered by the lists, namely lexical coverage.

The second software used was ‘The Compleat Lexical Tutor’s also a vocabulary profiler. It is a Canadian free web-based resource developed by Tom Cobb (2008) as a web based version of Nation’s Range and Frequency programs.

The third computer programme used was ‘TextCompare software’ developed for the purposes of this study. This programme compares any two texts, say Text1 and Text2 and helps find out which words occurred in one text and not the other (Text1 minus Text2, or Text2 minus Text1), and which words occurred in both texts (the intersect).

The last programme used was Text Master (2009, available at <http://www.nontube.com/products/text-master/>), a plain text editor freely downloadable from the internet. Its chief features are statistical analysis and generating a word list.

9.1. Conclusion

The analysis of the vocabulary in the textbooks used here showed that overall vocabulary contents of the textbooks reflected a level of difficulty inappropriate for the learners using them. It revealed that learners lexical coverage is far from the level advocated by researchers in the field of teaching and learning English as a foreign language. This would certainly make reading comprehension a daunting, if not impossible, task for learners. Certainly, one cannot expect a four-skills textbook to provide extensive vocabulary practice, yet one would hope that the design of the textbook would be such that the vocabulary items it does contain are organised in a way that facilitates their learning. It should be pointed out that the analysis conducted in this research was not intended to criticize the vocabulary of the textbooks; indeed, the author assumes that most textbooks suffer from similar issues as those described above. Is it possible to write a textbook that suits all the heterogeneous level found in classrooms? The answer is definitely ‘no’. No textbook is perfect and no single one can meet all the needs and interests of each group of learners, let alone the “suitable” lexical coverage for learners. However, textbooks limitations can be reduced to a great extent, especially as regards vocabulary. As this study demonstrates, teachers can reduce these limitations by becoming familiar with the vocabulary contents of their textbooks. Teachers generally do not have enough time to devote to vocabulary-focused instruction, i.e. formally presenting vocabulary items and providing activities that are designed to practice and reinforce the target vocabulary. Furthermore, because as

stressed through this study, the number of vocabulary items in any language is so large, only a small number can be dealt with through vocabulary-focused instruction and most will either have to be acquired through exposure or not at all (Schmitt, 2000, p.3).

Consequently, these factors make it very important for teachers to be well-informed about the vocabulary in their textbooks, so that that they can make pedagogically-sound decisions regarding vocabulary instruction. The results show the teacher which and how many words need to receive additional practice beyond the textbook to meet the vocabulary learning goals of the course. A certain number of measures could be taken to such an end. These are considered below.

9.2. Pedagogical implications

I believe the findings of the study provide pedagogical significance for the teachers who are using the textbooks investigated. The procedures of this study also give insights for teachers to examine and to choose their target lexis. The findings could inform teachers the treatment of vocabulary and the limitations of the textbooks. Obviously, lexical input in textbooks is prepared by textbook writers and publishers, but not classroom teachers. However, it is the teachers who decide what to teach and how to teach. By understanding the hidden aspects of vocabulary treatments in a textbook, teachers could adapt the materials in the textbooks more effectively in the vocabulary teaching. This would also be of a great asset to select or reject the vocabulary items according to their students' level and needs more wisely. It is hoped that teachers could adopt some measures to supplement the vocabulary teaching.

Some of these measure could be:

1. Because the 2000 word level is an essential goal for second language learners, it is hoped that the textbook would enable the students to attain at least the 1000 word level by the end of the course. The task of teachers would be to remedy to textbooks pitfalls, if any, by not overlooking and by including the most frequent lexical items in the target lexis they deal with in classrooms. One of the merits of the present study is to have drawn up a list of lexical items from the most frequent words which are included in the textbooks.
2. Word frequency lists are one solution for teachers and textbook designers to select what lexical items to teach. Schmitt (2000, p.82) suggested “word frequency lists are one important linguistic tool to come out of corpus research. They can be used to great effect to improve vocabulary teaching...” Clearly, frequency lists do provide useful information for these people. When teachers are aware that frequency lists can be an extremely useful pedagogical tool, they would also be aware that the most frequent words need to be taught first. Nation (2003) precisely advised that teaching useful vocabulary before less useful vocabulary gives learners the best return for their learning effort.
3. Once teachers understand the limitation of the textbooks by knowing the frequency profile of the texts, teachers could supplement lexical input by many ways. For example, teachers can make use of the supplementary materials

-
- provided by the same textbook, or deliberately include other learning materials or activities.
4. The data from this study could be used to make glossaries to appear at the end of the textbooks as most publishers do, but not present systematically in Algerian textbooks. These glossaries would help textbooks designers and educational authorities to suggest a list of target lexis and hence unify the teaching and learning of lexis in the Algerian school. This would be of great help in the design of national end of cycle examination taken by middle and secondary school students on leaving school.
 5. The design of the textbook should be such that words appearing in AWL and GSL and already in textbooks are included in the target lexis in lessons and given due emphasis to ensure that learners know them. Furthermore, they should be included in tests and term examinations
 6. When including the aforementioned lexical items, the design of the textbook should be such that these items it does contain are recycled at regular intervals in order both to refresh the students' learning and give them needed practice of the words.
 7. Another merit of this study is to make teachers aware of the existence of vocabulary treatment software such as the ones used in this study. In addition the Textcompare software conceived for this study can be used as a quick tool to compare the lexis occurring in a textbook file/unit to another one, or to the

-
- whole book, or even all the books. The results can be used for pedagogical purposes such as the selection of the target lexis, the planning of lexis distribution over the whole school year or even the curriculum.
8. When teachers know what words are or are not covered by the textbooks, they would be better able to provide supplemental reading material. Without significant supplementation by the teacher to extend the vocabulary of the course beyond that offered in the textbook, students probably have little chance of increasing their vocabulary to even the 1000 word level.
 9. Based on the findings of the present study, the textbooks writers and EFL teachers could improve vocabulary teaching by paying closer attention to the systematicity of vocabulary selection. The number of new lexical items in each unit/file or textbook can controlled for a better presentation and sequencing of vocabulary.
 10. The high frequency words can be distributed in a well informed way throughout all the curriculum so that when learner leave the secondary school to go to the university, they would have encountered and would have had opportunity to learn at least to the items included in the General Service List and the Academic Word List.
 11. Textbook designers should build the vocabulary progression and gradation in one book on items in previous one. This study provides data for doing so.

12. The results and data of this study can further be exploited to help bridge the gap, as far as lexical coverage, readability and consequently reading comprehension, that exists on the one hand between the middle school and the secondary school and, on the other hand, between the secondary school and the university. It has already been mentioned earlier in this dissertation that secondary school teachers of English are generally unaware of the contents of the middle school textbooks and university teachers of English are even more ignorant of the issue. Consequently, a secondary school teacher can by using the list of headwords in middle school textbooks and the Textcompare tool compare any reading material intended for secondary school against that list and deduce the lexical coverage in order to be able to predict whether the comprehension likeliness of the reading material. Similarly, a university teacher of English would be able to ‘gauge’ his lectures or reading material in a similar way.

Finally, it is hoped that the study would give the teachers insights into the investigation of vocabulary treatment in a textbook series. The study is particularly informative to the teachers using one of the textbooks in the study. It is also hoped that the methodology can demonstrate a few ways to examine words in a textbook and the result can be exploited for improvement of the teaching and learning of English in Algerian schools.

9.3. Limitations of the study

A few limitations of the study are worth noting before making some recommendations for future research. First the results generated from the study will be restricted to EFL textbooks in use in Algerian schools. Second, the lexical coverage and readability findings will be limited to the texts contained in the textbooks used in this study. Next, this research was only a preliminary attempt to survey the Algerian EFL textbooks lexical coverage. Its focus on the vocabulary component by no means implies that lexical coverage and readability are exclusive factors in designing a textbook for classroom use.

Through a lexical corpus study, the researcher hopes to draw the attention of EFL teachers and textbook designers to the vocabulary component of textbooks and how important it is to control it to insure effective reading comprehension. A good command of vocabulary is essential for success at higher levels of education. The current research data may serve as a basis for in-depth research into the acquisition of vocabulary. Other parameters such as syntax and content area knowledge may be worth investigating but are beyond the focus of the current study. Last but not least, vocabulary may not be the only component of an English course, but it is a component that learners notice and that can occupy a lot of their learning time. It is a component that deserves more attention from course designers and the aim of this study has been to generate that attention.

9.4. Recommendations for Future Research

I may prove useful to further research the issues considered here through a qualitative analysis involving the actors in the learning situation: learners and teachers.

Investigating students' perception toward language textbooks in relation to lexical coverage, readability, vocabulary load and reading difficulty and eliciting teachers' opinion would certainly generate new insights that would hopefully confirm the findings of the present study. A study that would focus only on the target lexis by asking teachers which vocabulary is focused on in classrooms (though learning is idiosyncratic and incidental) would be most welcome.

It would also be interesting to examine how a textbook can be used and how books within one series may provide a pathway for language development.

Another issue that can be exploited easily by making use of the data collected in this study is to examine the the degree of repetition of words throughout all the seven textbooks. Research shows that it is necessary to encounter a word in a variety of contexts a number of times, at regular intervals, in order for the learner to have a realistic chance of learning the word (Nation, 2001, Schmitt, 2000). An analysis of the textbook can show if target vocabulary occurs frequently enough and is given enough repetitions over time to provide optimum vocabulary-learning conditions. The results can guide teachers in deciding how best to supplement the text with activities that will give learners exposure to target vocabulary that is not sufficiently presented in the textbook.

Concluding remark

The gist of the present study can be summarized in a in a single statement:

If we want to avoid frustration to our learners and help them succeed in reading, we should gives them material they understand. How do we want them to succeed in reading comprehension when we provide them with material of which they are likely to understand only 40%; how do we want them to continue listening to us or reading in such conditions, especially that a textbook has been defined as ‘a book that no one would like to read’.

BIBLIOGRAPHY

- Alderson, J. C., & Urquhart, A. H. (1983). The effect of student background discipline on comprehension: A pilot study. In A. Hughes & D. Porter (Eds.), *Current developments in language testing*. London: Academic Press.
- Alderson, J.C. & Urquhart, A.H. (Eds.), (1984). *Reading in a Foreign Language*. London: Longman.
- Alderson, J.C. (1984). Reading in a foreign language: a reading problem or a language problem? In J.C. Alderson and A. H. Urquhart (Eds.), *Reading in a Foreign Language*. London: Longman.
- Alderson, J.C. (1988, March). Testing reading comprehension skills. *Paper presented at TESOL '88 Convention*. Chicago.
- Alexander, P. A. (1997). Knowledge –seeking and self-schema: A case for the motivational dimensions of exposition. *Educational Psychologist*, 32(2), 83–94.
- Alexander, P.A., & Jetton, T.L. (2000). Learning from text: A multidimensional and developmental perspective. In M.L. Kamil, P.B. Mosenthal, P.D. Pearson, & R. Barr (Eds.), *Handbook of reading research*, Vol. 3, 285–310. Mahwah, NJ: Erlbaum.
- Alexander, P.A., Kulikowich, J.M., & Schulze, S.K. (1994). The influence of topic knowledge, domain knowledge, and interest on the comprehension of scientific exposition. *Learning and Individual Differences*, 6, 379-397.
- Al-Jarf, Reima (2007). Online instruction and Creative Writing by Saudi Freshman Students. *Asian EFL Journal*, 22, August.
- Alliende F. (1990). Readability evaluation of written materials. *Lectura Y. Vida*, 11 (S), 14-18.
- Alliende, F. (1987). Four quick procedures to determine Text Readability. *Lectura Y. Vida*, 8 (4), 12-22.

- Anderson, N.J.(1999). Exploring Second Language Reading: Issues and Strategies. Boston, MA: Heinle & Heinle.
- Anderson, R.C. and Freebody, P. (1981). Vocabulary Knowledge. In J.T. Guthrie (Ed.), *Comprehension and teaching: Research reviews*. Newark, DE: International Reading Association.
- Anderson, R.C. and Pearson, P.D. (1984). A Schema-Theoretic View of Basic Processes in Reading Comprehension. In Carrell, P.L., Devine, J. & Eskey, D.E. (Eds.) (1988). *Interactive Approaches to Second Language Reading*. Cambridge: CUP.
- Andrich, D. and Godfrey, J. R. (1979). Hierarchies in the skills of Davis' reading comprehension test. *Reading Research Quarterly*, 19, 182-200.;
- Appelt, Ken. 2006. Readability and Reading Ability. *Literacy Links* . Volume 10, No. 3, June 2006 URL: <http://www-tcall.tamu.edu/newsletr/jun06/jun06j.html> . [Accessed July 2006].
- Aslanian, Y. (1985). Investigating the reading problems of English as a second language students: an alternative. *English Language Testing Journal* 39(1), 20-27.
- Astika, G. G. (1993). Analytical assessment of foreign students' writing. *RELC Journal*, 24, 61-72.
- Bachman, L. F. (2000). Modern language testing at the turn of the century: assuring that what we count counts. *Language Testing* 17, 1-42
- Baddeley, A. (1990). *Human memory*. London: Lawrence Erlbaum Associates.
- Baker, L., & Brown, A.L. (1984). Metacognitive skills and reading. In P.D. Pearson, R. Barr, M. Kamil, & P. Mosenthal (Eds.), *Handbook of reading research: Volume I* (pp. 353-394). White Plains, NY: Longman.
- Barcroft, J. (2004). Promoting second language vocabulary acquisition during reading: Some pivotal issues and eight instructional techniques. *The Southern Journal of Linguistics*, 26(2), 94-114.

- Barcroft, J. (2004). Second language vocabulary acquisition: A lexical input processing approach. *Foreign Language Annals*, 37(2), 200-208.
- Barnett, M.A. (1989). *More than meets the eye: Foreign language learner reading: Theory and practice*. Englewood Cliffs, NJ: Prentice Hall Regents.
- Barrett, M.D. (1982) Distinguishing between prototypes: the early acquisition of the meaning of object names (313-334). In S.A. Kuczaj II (ed). *Language Development*. Lawrence Erlbaum Associates, New Jersey:.
- Bartlett, F.C. (1932). *Remembering: A Study in Experimental and Social Psychology*. Cambridge University Press.
- Beck, I.L. & McKeown, M.G. (1991). Conditions of vocabulary acquisition. In P.D. Pearson (Ed.), *Handbook on Reading Research, Vol. 2*, (pp.789-814). New York: Longman
- Beck, I.L., Perfetti, C.A. and McKeown, M.G. (1982) The effects of long-term vocabulary instruction on lexical access and reading comprehension. *Journal of Educational Psychology* 74, 506-521.
- Beglar, D. (1995) Review of New Ways in Teaching Vocabulary by I.S.P. Nation (ed). *The Language Teacher* 19, 2: 50-51.
- Beglar, D. (2009). A Rasch-based validation of the Vocabulary Size Test. *Language Testing*, 27, 101–118.
- Belouahem, R. (2008). *The Suitability of the First Year Secondary School Coursebook 'At The Crossroads' to the Algerian Teaching Environment*. Doctoral dissertation. Mentouri University, Constantine.
- Bensoussan, M. & Laufer, B. (1984). Lexical guessing in context in EFL reading comprehension. *Journal of Research in Reading*, 7, 5-32.
- Bernhardt, E. (1991). *Reading Development in a Second Language: Theoretical, Empirical, and Classroom Perspectives*. New Jersey: Ablex.

- Bernhardt, E. B. (1984). Toward an Information-Processing Perspective in Foreign Language Reading. *Modern Language Journal* 68, 322–31.
- Bernhardt, E. B. (1985). Developmental Processes in Second Language Reading: An Update. *ACTFL Annual Meeting*. New York.
- Bernhardt, E. B. (1986). Proficient Texts or Proficient Readers? *Association of Departments of Foreign Languages Bulletin* 18, no. 1 (September 1986): 25-28.
- Bernhardt, E. B., & Kamil, M. L. (1995). Interpreting relationship between L1 and L2 reading: Consolidating the linguistic threshold and the linguistic interdependence hypothesis. *Applied Linguistics*, 16, 15-34.
- Billuroğlu, A. & Neufeld, S. (2005). *The Bare Necessities in Lexis: a new perspective on vocabulary profiling*. Retrieved September 2007 from http://lextutor.ca/vp/BNL_Rationale.doc
- Blake, M. E., & Majors, P. L. (1995). Recycled words: Holistic instruction for LEP students. *Journal of Adolescent & Adult Literacy*, 39(2), 132-137.
- Bloom, B. S. (ed.). (1956). *Taxonomy of Educational Objectives, the classification of educational goals – Handbook I: Cognitive Domain*. New York: McKay.
- Bloom, K. C., & Shuell, T. J. (1981). Effects of massed and distributed practice on the learning and retention of second-language vocabulary. *Journal of Educational Research*, 74, 245–248.
- Bormuth, J. R. (1966). Readability: A new approach. *Reading Research Quarterly* 1, 79-132.
- Bormuth, J. R. (1969). *Development of readability of analysis*. Chicago: The University of Chicago.
- Bormuth, J. R. (1971). *Development of standards of readability: Toward a rational criterion of passage performance*. Chicago: The University of Chicago.

- Bossers, B. (1991). On Thresholds, Ceilings and Short-Circuits: The Relation between L1 Reading, L2 Reading and L2 Knowledge. In J. H. Hulstijn and J.F. Matters (Eds.), *Reading in two Languages: AILA review*. Vol. 8, 45-60.
- Britton, B. K; Guelgoez, S. (1991). Using Kintsch's computational model to improve instructional text: Effects of repairing inference calls on recall and cognitive structures. *Journal of Educational Psychology*, 83, 329-345.
- Britton, B.K. & Black, J.B. (Eds.), (1985). *Understanding expository text: A theoretical and practical handbook for analysing explanatory text*. Hillsdale NJ:Lawrence Erlbaum Associates, pp315-362.
- Britton, Bruce K. (1994). Understanding Expository Text. In M.A. Gernsbacher (Ed.), *Handbook of Psycholinguistics* (pp.641-674). San Diego : Academic Press.
- Brown, H. D. (2001). *Teaching by Principles* (2nd ed). Pearson Education Company.
- Brown, J. D. (1992). What text characteristics predict human performance on cloze test items? In *Proceedings of the 3rd conference on second language research in Japan, Tokyo. November 16, 1991* (pp. 1-26). Niigata: Language Programs of the International University of Japan.
- Brown, J. D. (1998). An EFL readability index. *JALT Journal*, 20(2).
- Brown, J.D. (1997). An EFL readability index. *University of Hawaii Working Papers in ESL* 15(2), 85-119.
- Bruce, B. and Rubin, A. (1988). Readability formulas: matching tool and task. In: Davison and Green (Eds.), *Linguistic complexity and text comprehension: Readability issues reconsidered*. Hillsdale, NJ: Lawrence Erlbaum Associates, , pp 5-22.
- Bruce, B., A Rubin, and K. Starr. (1980). Why readability formulas fail. *Reading Education Report No. 28*. Champaign, IL: University of Illinois at Urbana-Champaign.

- Bügel, K. & Buunk, B. P. (1996). Sex differences in foreign language text comprehension: The role of interests and prior knowledge. *Modern Language Journal*, 80(1), 15-31.
- Burton, G.E. (1991). The readability of consumer-oriented bank brochures: an empirical investigation. *Business & Society* 30(1):21-26.
- Campbell, G. and Weir, G. R. S. (2006). Matching readers to texts with the Strathclyde readability measure. In: *ICT in the Analysis, Teaching and Learning of Languages*, Preprints of the ICTATLL Workshop 2006, Glasgow, UK.
- Carrel, P. (1991). Second Language reading: Reading ability or language proficiency? *Applied Linguistics*, 12, 159-179.
- Carrell, P, Devine, J. and Eskey, D.E. (eds). (1988). *Interactive Approaches to Second Language Reading*. Cambridge: CUP.
- Carrell, P. (1984a). Schema theory and ESL reading: Classroom implications and applications. *Modern Language Journal*, 68, 332-43.
- Carrell, P. (1984b). Evidence of a formal schema in second language comprehension. *Language Learning*, 34, 87-112.
- Carroll, J.B., Davies, P. and Richman, B. (1971) . *The American Heritage Word Frequency Book*. Houghton Mifflin, Boston American Heritage, New York.
- Carter, R. (1987). *Vocabulary*. London: Allen and Unwin.
- Carver, R. P. (1977). Toward a theory of reading and comprehension and rauding. *Reading Research Quarterly*, 1, 9-89.
- Carver, R.P. (1994). Percentage of unknown vocabulary words in text as a function of the relative difficulty of the text: implications for instruction. *Journal of Reading Behavior* 26, 4: 413-437.
- Chall, J. & Dale, E. (1995). *Readability revisited: The New Dale-Chall Readability Formula*. Cambridge, Massachusetts: Brookline Books.

- Chall, J. (1958). *Readability: An appraisal of research and applications*. Columbus, Ohio: Ohio State University Press.
- Chall, J. S. (1988). The beginning years. In B. L. Zakaluk, and S. J. Samuels (Eds.), *Readability: Its Past, Present, and Future* (pp. 2-13). Newark, DE: International Reading Association.
- Chall, J. S., and S. S. Conrad. (1991). *Should textbooks challenge students? The case for easier or harder textbooks*. New York: Teachers College Press.
- Chall, J. S., Bissex, G. L., Conard, S. S., & Harris-Sharples, S. (1996). *Qualitative assessment of text difficulty: A practical guide for teachers and writers*. Cambridge, MA: Brookline Books.
- Chall, J., & Dale, E. (1948). A formula for predicting readability. *Educational Research Bulletin*, 27, 11-20.
- Chall, J.S. (1987) Two vocabularies for reading: recognition and meaning. In M.G. McKeown & M.E. Curtis (Eds.), *The nature of vocabulary acquisition* (165-184). Hillsdale, NJ: Erlbaum.
- Chambliss, M., & Calfee, R. C. (1998). *Textbooks for learning: Nurturing children's minds*. Cambridge, MA: Basil Blackwell.
- Chastain, K. (1988). *Developing Second Language Skills*. (3rd ed.). Chicago: Harcourt Brace Jovanovich.
- Child, D. (1973). *Psychology and the Teacher*. London, Holt Rinehart & Winston.
- Clarke, M.A. (1979). Reading in Spanish and English: Evidence from adult ESL students. *Language Learning*, 29, 121-150.
- Clarke, M.A. (1980). The short circuit hypothesis of ESL reading--or when language competence interferes with reading performance. *Modern language Journal*, 64, 203-209
- Coady, J. & Huckin, T. (Eds.), (1997). *Second Language Vocabulary Acquisition*. C.U.P.,

- Coady, J. (1979) A psycholinguistic model of the ESL reader. In R. Mackay et al (eds). *Reading in a Second Language* (pp. 5-12). Newbury House, Rowley, Massachusetts.
- Coady, J. (1997) L2 acquisition: a synthesis of the resarch. In J. Coady, & T. Huckin, (Eds.), *Second Language. Vocabulary Acquisition* (pp. 273-290) . Cambridge: Cambridge University.
- Cobb, T. (2000). *Compleat Lexical Tutor* [computer software]. <http://132.208.224.131/>
- Cohen, J. (1975). Effect of content area on cloze test performance. *Journal of Reading*, (December) 247-250.
- Coleman, M., and Liau, T. L. (1975). A computer readability formula designed for machine scoring. *Journal of Applied Psychology*, 60, 283-284.
- Connaster, B. F. 1999. "Last rites for readability formulas in technical communication." *Journal of technical writing and communication*. 29, no. 3:271-287.
- Cook, G. (1989). *Discourse in language teaching: A scheme for teacher education*. Oxford: Oxford University Press.
- Cooper, M. (1984). Linguistic competence of practised and unpractised non-native readers of English. In: Alderson, J.C. & Urquhart, A.H. C *Reading in a Foreign Language* (pp122-138). London:Longman.
- Coxhead, A. (1998). *An Academic Word List*. Occasional Publication No.18 LALS, Victoria University of Wellington, New Zealand.
- Coxhead, A. (2000). A new academic word list. *TESOL Quarterly*, 34(2), 213-238.
- Crawford, W0 J., King, C.E., Brophy, J. E., and Everston, C. M. (1975). *Error rates and question difficulty related to elementary children's learning*. Paper presented at the Annual Meeting of the American Educational Research Association. Washington, D.C.
- Crossley, S. A., M. M. Louwerse, P. M. McCarthy & D. S. McNamara (2007). A linguistic analysis of simplified and authentic texts. *Modern Language Journal* 91(2), 15–30.

- Crossley, S.A. 2006. *A computational Approach to assessing Language reading texts*. Unpublished PhD dissertation University of Memphis. Retrieved from www.jstor.org/.
- Crossley, S.A. and McNamara, D.S. (2008). Assessing L2 reading texts at the intermediate level: An approximate replication of Crossley, Louwse, McCarthy & McNamara (2007). *Language Teaching* 41(3), 409-429.
- Cummins, J. (1976). The influence of bilingualism on cognitive growth: a synthesis of research findings and explanatory hypothesis. *Working Papers on Bilingualism*, 9,1-43.
- Cummins, J. (1980) The crosslingual dimensions of language proficiency: implications for bilingual education and the optimal age issue. *TESOL Quarterly* 14(2), 175-187.
- Cummins, J. (1980)0 Age on arrival and immigrant second language learning in Canada: a reassessment. *Applied Linguistics* 11(2), 132-149.
- Cziko, G.A. (1978) Differences in first- and second-language reading: the use of syntactic, semantic and discourse constraints. *Canadian Modern Language Review* 34, 473-489.
- Cziko, G.A. 1980. Language competence and reading strategies: a comparison of first and second-language oral reading errors. *Language Learning*, 30, 101-116.
- Davis, F. (1944). Fundamental factors of comprehension in reading. *Psychometrika*, 9,185-197.
- Davis, F. (1968). Research in comprehension in reading. *Reading Research Quarterly*, 3, 499-545.
- Davis, F.B. (1972) Psychometric research on comprehension in reading. *Reading Research Quarterly* 7: 628-678.
- Davison, A. & Kantor, R.N. On the failure of readability formulas to define readable text: A case study from adaptations. *Reading Research Quarterly*, 17(2), 187-209, 1982.
- Day, R. R. (1994). Selecting a passage for the EFL reading class. *Forum*, 32(1).

- Day, R. R. and Bamford, J. (1998). *Extensive reading in the second language classroom*. Cambridge: Cambridge University Press.
- de Beaugrande, R., and W. U. Dressler . (1981). *Introduction to text linguistics*. London: Longman.
- Beck & McKeown 1999
- Dempster, F. N. (1987). Effects of variable encoding and spaced presentation on vocabulary learning. *Journal of Educational Psychology*, 79, 162–170.
- Devine, J. (1988b). The relationship between general language competence and second language reading proficiency: Implications for teaching. In P. Carrell, J. Devine, & D.E. Eskey (Eds.), *Interactive Approaches to Second Language Reading* (pp. 260-77). Cambridge: Cambridge University Press.
- Dewey, J. (1938). *Education and experience*. New York: Macmillan.
- Donald C. S. (1993). *Editing Technical Writing*. Oxford University Press.
- DuBay. W. H. (2004). *The Principles of Readability*. Costa Mesa, CA: Impact Information.
- Ellis, N.C. (1994). Vocabulary acquisition: the implicit ins and outs of explicit cognitive mediation. In N.C. Ellis (ed), *Implicit and Explicit Learning of Languages* (211-282). Academic Press, London.
- Ellis, N.C. (1995). Vocabulary acquisition: psychological perspectives and pedagogical implications. *The Language Teacher* 19(2), 12-16.
- Engels, L.K. (1968). The fallacy of word counts. *IRAL* 6(3), 213-231.
- Equations in Scientific Proofs: Effects on Comprehension
- Erlbaum. Greenfield, G. (1999). *Classic readability formulas in an EFL context: Are they valid for Japanese speakers?* Unpublished doctoral dissertation, Temple University, Philadelphia, PA, United States. (University Microfilms No. 99–38670).

- Eskey, D. E. and Grabe, W. (1988). "Interactive Models for Second Language Reading: Perspectives on Instruction," Carrell, P. L., Devine, J and Eskey. D. E. (eds). *Interactive Approaches to Second Language Reading*. Cambridge: CUP.
- Eskey, D.E. (1973). A model program for teaching advanced reading to students of English as a foreign language. *Language Learning*, 23, 169-184.
- Eskey, D.E. (1988) "Holding in the Bottom: an Interactive Approach to the Language Problems of Second Language Readers," in Carrell, P.L., Devine, J. and Eskey, D.E. (Eds.), *Interactive Approaches to Second Language Reading*. Cambridge: CUP.
- Farr, R. (1969). *Reading: What can be measured?* Newark, DE: International Reading Association.
- Faucett, L., & Maki, I. (1932). *A study of English-word values statistically determined from the latest extensive word counts*. Tokyo: Matsumura Sanshodo.
- Feldmann, U., & Stemmer, B. (1987). Think aloud a retrospective: different languages and different learner's approaches? In C. French, & G.Kasper, (Eds.), *Introspection in second language research* (pp. 251-67). Clevedon: Multilingual Matters.
- Finocchiaro, Mary. (1958). *Teaching English as a Second Language in Elementary and Secondary Schools*. New York, Harper and Row.
- Fisher, D.F., & Shebilske, W.L. (1985). There is more that meets the eye than the eye mind assumption. In R. Groner, G.W. McConkie, & C. Menz (Eds.), *Eye movements and human information processing*. Amsterdam: Elsevier.
- Flesch, R. F. (1943). *Marks of readable style: A study in adult education*. New York: Bureau of Publications, Teachers College, Columbia University.
- Flesch, R. F. (1955). *Why Johnny Can't Read—And What You Can Do About It* - New York: NY: Harper & Row.
- Flesch, R. F. (1964). *The ABC of style: A guide to plain English* . New York: Harper & Row.

- Flesch, Rudolf. (1979). *How to Write Plain English: A Book for Lawyers and Consumers*. New York: Harper and Row.
- Francis, W.N. and Kucera, H. (1982). *Frequency Analysis of English Usage* Houghton Mifflin Company, Boston.
- Freebody, P. (1981). Effects on text comprehension of differing propositions and locations of difficult vocabulary. *Journal of Reading Behavior* 15(3), 19-39.
- Fry, E. (1964). Judging readability of books. *Teacher Education*, 5, 34-9.
- Fry, E. (1977). Fry's readability graph. *Journal of Reading*, 20, 242-52.
- Fry, E. B. 1989. Reading formulas: Maligned but valid. *Journal of Reading*, 32, 292-297.
- Ganderton, R. (1998). New strategies for a new medium? Observing L2 reading on the World Wide Web. *ON-CALL*, 2. <http://www.cltr.uq.edu.au/oncall/gander122>.
- Gardner, D. (2008). Vocabulary recycling in children's authentic reading materials: A corpusbased investigation of narrow reading. *Reading in a Foreign Language*, 20, 92-122.
- Gardner, I.C. (1978). Reading comprehension. In *The teaching of comprehension (ETIC Occasional Paper* (pp. 65-81). London: The British Council, English Teaching Information Centre.
- Gardner, I.C. (1978). Reading comprehension. In *The teaching of comprehension (ETIC Occasional Paper* (pp. 65-81). London: The British Council, English Teaching Information Centre.
- Geva, E. (1983). Facilitating reading comprehension through flowcharting. *Reading Research Quarterly*, 18, 384-341.
- Ghahraki, S. and F. Sharifian. (2005). The relationship between overall reading comprehension and determination of fact/opinion in L2. *The Reading Matrix*, 5(1), Retrieved from: www.readingmatrix.com
- Gilliland, J. 1972. *Readability*. London: Hodder and Stoughton.

- Glenn, C.G. (1978). The role of episodic structure and of story length in children's recall of simple stories. *Journal of Verbal Learning and Verbal Behaviour*, 17(2), 229-247.
- Goldman, S. R., & Rakestraw, J. A. (2000). Structural aspects of constructing meaning from text. In M. L. Kamil, P. B. Mosenthal, P. D. Pearson, & R. Barr. *Handbook of reading research Vol III*, (pp. 311-355). Mahwah, N. J.: Erlbaum.
- Goodman, K. S. (1975) "The Reading Process," in Carrell, P.L., Devine, J. and Eskey, D.E. (eds) (1988). *Interactive Approaches to Second Language Reading*. Cambridge: CUP.
- Goodman, K. S. (1965). A linguistic study of cues and miscues in reading. *Elementary English*, 42, 639-643.
- Goodman, K. S. (1967). Reading: A psychological guessing game. *Journal of the Reading Specialist*, 6, 35-126.
- Goodman, K. S. (1970). Reading a Psycholinguistic Guessing Game. In H. Singer and R. B. Ruddell (Eds.), *Theoretical Models and Processes of Reading*. Newark, Delaware: International Reading Association.
- Goodman, K. S. (1971). Psycholinguistic universals in the reading process. In P. Pimsleur & T. Quinn (Eds.), *The psychology of second language learning* (pp. 135-142). Cambridge, MA: Cambridge University.
- Goodman, K. S. (1973). *Theoretically based studies of patterns of miscues in oral reading performance*. Detroit: Wayne State University.
- Goodman, K. S. (1975). Miscues: Windows on the reading process. In K.S. Goodman (Ed.) *Miscue analysis, applications to reading instruction* (pp. 3-14). Educational Resources Information Clearinghouse, Urbana, IL.
- Goodman, K. S. (1982). What is universal about the reading process. In F.V. Gollasch (Ed.), *Language and literacy: the selected writings of Kenneth S. Goodman, Volume 1: Process, theory, research* (pp. 70-75). Boston: Routledge and Kegan Paul LTD.

- Goodman, K. S. (1986). *What's whole in whole language*. Portsmouth, NH: Heinemann Educational Books.
- Goodman, K. S. (1988). The reading process. in Carrell, P.L., Devine, J. and Eskey, D.E. (Eds.). (*Interactive Approaches to Second Language Reading* pp. 11-21). Cambridge: CUP.
- Goodman, K. S. (1994). Reading, writing, and written texts: A transactional sociopsycholinguistic view. In R.B. Ruddell, M.R. Ruddell, & H. Singer (Eds.), *Theoretical models and processes of reading*, 4th edition. Newark, DE: International Reading Association.
- Goodman, K. S. (1996). *On Reading: A common-sense look at the nature of Language and the science of reading*. Portsmouth, NH: Heinemann.
- Goodman, K. S., & Goodman, Y. (1994). To err is human: Learning about language processes by analyzing miscues. In R. B. Ruddell, M. R. Ruddell, & H. Singer (Eds.), *Theoretical models and processes of reading* (104-123). Newark, DE: International Reading Association.
- Goodman, K.S., & Goodman, Y.M. (1980). Learning to read is natural. In L.B. Resnick & P.A. Weaver (Eds.), *Theory and practice of early reading* (Vol. 1, pp. 137-154). Hillsdale, NJ: Erlbaum.
- Gordan, C. (1992). The Role of prior knowledge in narrative and expository text. *The document Reproduction Service*.
- Gough, P. B., Hoover, W. A., & Peterson, C. (1996). A simple view of reading. In C. Cornoldi & J. Oakhill (Eds.), *Reading comprehension difficulties: Processes and intervention* (pp 1-14). Mahwah, NJ: Lawrence Erlbaum Associates.
- Gough, P.B. (1972). One second of reading. In J.F. Kavanaugh & I.G. Mattingly (Eds.), *Language by ear and by eye*. Cambridge, MA: MIT Press.
- Goulden, R., Nation P, & Read J. (1990). How large can a receptive vocabulary be? *Applied Linguistics*, 11(4), p. 341-363.

- Grabe, W. (1988) "Reassessing the Term 'Interactive'", in Carrell, P.L., Devine, J. and Eskey, D.E. (eds) (1988). *Interactive Approaches to Second Language Reading*. Cambridge: CUP.
- Grabe, W. (1991) Current developments in second language reading research. *TESOL Quarterly* 25, 3: 375-406.
- Grabe, W. (2002). Foundations for L2 Reading Instruction. [online]. *The Language Teacher* . Retrieved from: <http://www.jalt-publications.org/tlt/articles/2002/07/grabe> [Accessed 4 April 2006]
- Grabe, W., & Stoller, F. (2002). *Teaching and researching reading: Applied linguistics in action*. New York: Longman.
- Graves, M. F., Ryder, R., Slater, W., & Calfee, R. (2001). The Relationship between word frequency and reading vocabulary using six metrics of frequency. *Journal of Educational Research*, 81(2), 81-90.
- Graves, M.F. & Graves, B.B. (2003). *Scaffolding reading experiences: Designs for student success* (2nd ed.). Norwood, MA: Christopher-Gordon.
- Graves, M.F. (1986). Vocabulary learning and instruction. *Review of Research in Education* 13, 49-89.
- Gray, W. S. and Leary, B. (1935). *What makes a book readable?* Chicago: University of Chicago Press.
- Greene, R. L. (1992). *Human memory: Paradigms and paradoxes*. Hillsdale, NJ: Lawrence Erlbaum.
- Greenfield, G. (1999). *Classic readability formulas in an EFL context: Are they valid for Japanese speakers?* Unpublished doctoral dissertation, Temple University, Philadelphia, PA, United States. (University Microfilms No. 99-38670). Grellet 1981;
- Greenfield, G. (2003). The Miyazaki EFL Readability Index. *Comparative Culture* 9, 41-49. Miyazaki, Japan: Miyazaki International College.

- Greenfield, J. (2004). Readability formulas for EFL. *JALT Journal*, 26, 5–24.
- Gross, P.P. & Sadowski, K. (1985). For Index - A, Readability formula program for microcomputers. *Journal of Reading*, 28(7), 614-618.
- Gunning, R. 1952. *The technique of clear writing*, New York: McGraw-Hill.
- Halliday, M. A. K., & Hasan, R. (1976). *Cohesion in English*. London: Longman.
- Hamsik, M. J. (1984). *Reading, readability, and the ESL reader*. Unpublished doctoral dissertation, The Florida State University.
- Harley, B. (1996) Vocabulary learning and teaching in a second language. *Canadian Modern Language Journal* 53(1) 3-12.
- Harris, T. L., & Hodges, R. E. (Eds.), (1981). *A dictionary of reading and related terms*. Newark, DE: International Reading Association.
- Harrison, C. (1980). *Readability in the Classroom*. Cambridge, England: Cambridge University.
- Harris-Sharples, S. (1983). *A study of the "watch" between student reading ability and textbook difficulty during classroom instruction*. Doctoral dissertation: Harvard graduate school of education. Cambridge, MA.
- Haynes, M. (1989). *Individual differences in Chinese readers of English Orthography and reading*. Doctoral dissertation, Michigan State University.
- Haynes, M. (1993). Patterns and perils of guessing in second language reading. In Huckin, T. Haynes, M. & Coady, J. (Eds.), *Second Language Reading and Vocabulary Acquisition* (pp. 46-64). Norwood, NJ: Ablex.
- Hazenberg, S. and Hulstijn, J.H. (1996). Defining a minimal receptive second-language vocabulary for non-native university students: an empirical investigation. *Applied Linguistics* 17(2), 145-163.
- Henriksen, B. (1999). Three dimensions of vocabulary development. *Studies in Second Language Acquisition* 21, 303–317.

- Hirsh, D. (1993). *The vocabulary demands and vocabulary learning opportunities in short novels*. Unpublished M.A. thesis. Victoria University of Wellington, New Zealand.
- Hirsh, D. and P. Nation. (1992). What vocabulary size is needed to read unsimplified texts for pleasure? *Reading in a Foreign Language* 8(2), 689-696.
- Hockey, S. (1988). *Micro-OCP*. Oxford University Press, Oxford.
- Holley, F.M. (1973) A study of vocabulary learning in context: the effect of new-word density in German reading materials. *Foreign Language Annals* 6, 339-347.
- Horn, E. (1926). A basic writing vocabulary, 10,000 words most commonly used in writing. College of Education, University of Iowa.
- Hosenfeld, C. (1979). Cindy: A learner in today's foreign language classroom. In W. C. Born, (Ed.). *The foreign language learner in today's classroom environment*. Northeast Conference Reports (ERIC Document Reproduction Service No. ED 185 834).
- Hu, M. and Nation, I.S.P. (2000). Vocabulary density and reading comprehension. *Reading in a Foreign Language* 13(1), 403-430.
- Huang, C. (1999). *The effects of vocabulary knowledge and prior knowledge on reading comprehension of EFL students in Taiwan*. (Doctoral Dissertation). Ohio University.
- Huckin, T. and Coady, J. (1999) Incidental vocabulary acquisition in a second language: a review. *Studies in Second Language Acquisition* 21, 181-193.
- Huckin, T. Haynes, M. & Coady J. (Eds.), (1993). *Second Language Reading and Vocabulary Acquisition*. Norwood, NJ: Ablex.
- Huckin, T., & Bloch, J. (1993). Strategies for inferring word meaning in context: A cognitive model. In T. Huckin, M. Haynes, & J. Coady (Eds.), *Second Language Reading and Vocabulary Acquisition* (pp. 153–178). Norwood, NJ: Ablex.
- Huey, E.B. (1968). *The psychology and pedagogy of reading*. Cambridge, MA: MIT Press. (Originally published 1908).

- Hughes, A. (1989). Synthesis of research on teaching writing. *Educational Leadership*, 44(8), 71-82.
- Hunt, A. & Belgar, D. (2005). A framework for developing EFL reading vocabulary. *Reading in a Foreign Language*, 17(1), 24.
- Hurlock, E. B. (1972). *Child Development*. Tokyo, McGrahill kogakusha Ltd.
- Hwang K. and Nation, P. (1989). Reducing the vocabulary load and encouraging vocabulary learning through reading newspapers. *Reading in a Foreign Language* 6(1), 323-335.
- Hwang, K. (1989). *Reading newspapers for the improvement of vocabulary and reading skills*. Unpublished M.A. thesis. Victoria University of Wellington, New Zealand.
- Hwang, K., & Nation, I. S. P. (1989). Reducing the vocabulary load and encouraging vocabulary learning through reading newspapers. *Reading in a Foreign Language*, 6, 323–335.
- Irwin, J. & Davis, C.A. (1980). Assessing readability: The checklist approach. *Journal of Reading*, 24, 124-30.
- Johnson, D. and Pearson, P.D. (1984). *Teaching Reading Vocabulary*. Holt, Rinehart & Winston, New York.
- Johnson, P. (1981) Effects on reading comprehension of language complexity and cultural background of a text. *TESOL Quarterly* 15(2), 69-181.
- Johnson, P. (1998). *What exactly are comprehension skills and how do I teach them?* Oxford: UKRA, Blackwell.
- Johnson, W. (1946). *People in Quandaries*. New York: Harpers.
- Johnston, P. H. (1983). *Reading comprehension assessment: A cognitive basis*. Newark, Delaware: International Reading Association.
- Jones, F.R. (1997). Learning an alien lexicon: a teach-yourself case study. *Second Language Research*, 11(2), 95-111.

- Jordan, R. R. (1997). *English for Academic Purposes- A Guide and Resource Book for Teachers*. Cambridge University Press, Cambridge, UK.
- Judd, C.H., & Buswell, G.T. (1922). *Silent reading: A study of the various types*. Chicago, IL: University of Chicago Press.
- Just, M. A. & P. A. Carpenter (1980). A theory of reading: From eye fixations to comprehension. *Psychological Review* 87, 329–354.
- Kern, R. G. (1988). Foreign language reading: Linguistics, cognitive, and affective factors which influence comprehension. In Fryer, T. B. and Medley, F. W., (Eds.), *New Challenges and Opportunities. Dimension: Languages '87* (49-61), Columbia, SC: Southern Conference on Language Teaching.
- Kern, R. G. (1989). Second language reading strategy instruction: Its effects on comprehension and word inference ability. *The Modern Language Journal* 73, 135-49.
- Kern, R. G. (1992). Teaching second language texts: Schematic interaction, affective response and the Directed-Thinking activity. *Canadian Modern Language Review* 48(2), 307-25.
- Kern, R. G. (1994). The role of mental translation in second language reading. *Studies in Second Language Acquisition* 16(4), 441-61.
- Kintsch, W. (1974). *The representation of meaning in memory*. New York: John Wiley and Sons.
- Kintsch, W. and E. Vipond. (1979). Reading comprehension and readability in educational practice and psychological theory. In *Perspectives on memory research*. Nilsson, L. G. (ed.). Hillsdale, NJ: Erlbaum.
- Kintsch, W. & T. A. van Dijk (1978). Toward a model of text comprehension and production. *Psychological Review* 85, 363–394
- Kitson, H. D. (1921). *The mind of the buyer*. New York: Macmillan.

- Klare, G. (1984). Readability . In P.D. Pearson (Ed.), *Handbook of reading research*, (pp. 681-744). New York: Longman.
- Klare, G. R. (1952). Measures of the readability of written communication: An evaluation. *Journal of Educational Psychology*, 43(7), 385-399.
- Klare, G. R. (1963). *The measurement of readability*. Ames, Iowa: Iowa State University Press.
- Klare, G. R. (1974-75). Assessing readability. *Reading Research Quarterly*, 10, 62-102.
- Klare, G. R. (1976). A second look at the validity of the readability formulas. *Journal of Reading Behaviour*, 8, 159-152.
- Klare, G. R. (1980). *How to write readable English*. London: Hutchinson.
- LaBerge, D. and Samuels, S.J. (1974). Toward a theory of automatic information processing in reading. *Cognitive Psychology*, 6, 293-323.
- Lakehal-Ayat – Benmati, K. (2008). *Is The Algerian Educational System Weakening? An Investigation of the High School Curricula and Their Adequacy with the University Curricula*. Doctoral dissertation. Mentouri University, Constantine.
- Landauer, T. K., Foltz, P. W., & Laham, D. (1998). Introduction to Latent Semantic Analysis. *Discourse Processes*, 25, 259-284.
- Lapkin, S. and Swain, M. (1977). The Use of English and French Cloze Tests in a Bilingual Program Evaluation: Validity and Error Analysis, *Language Learning* 27, 279-314.
- Laufer B, and Waldman T. (2011). Verb-noun Collocations in Second Language Writing: A Corpus Analysis of Learners' English'. *Language Learning*.
- Laufer, B. & Bensoussan, M. (1982) Meaning is in the eye of the beholder. *English Teaching Forum* 20(2), 10-13.

- Laufer, B. & Ravenhorst- Kalovski. G. C. (2010). Lexical threshold revisited: lexical text coverage, learners' vocabulary size and reading comprehension. *Reading in a Foreign Language* 22, 15-30.
- Laufer, B. (1981) A problem in vocabulary learning - synophones. *ELT Journal* 35(3), 294-300.
- Laufer, B. (1985) . *Vocabulary acquisition in a second language: the hypothesis of `synforms' (similar lexical forms)*. PhD, University of Edinburgh.
- Laufer, B. (1986). Possible changes in attitude towards vocabulary acquisition research. *IRAL* 24(1), 69-75.
- Laufer, B. (1987a). Teaching vocabulary: the lexical perspective of reading comprehension. *English Teachers' Journal* 35, 58-67.
- Laufer, B. (1987b). A case for vocabulary in EAP reading comprehension materials. In A.M.Cornu et al. (eds). *Beads or Bracelet? How do we approach ESP?* Oxford University Press, Oxford: 284-291.
- Laufer, B. (1988). The concept of `synforms' (similar lexical forms) in vocabulary acquisition. *Language and Education* 2(2), 113-132
- Laufer, B. (1989a). A factor of difficulty in vocabulary learning: deceptive transparency. *AILA Review* 6, 10-20.
- Laufer, B. (1989b). What percentage of text-lexis is essential for comprehension? In C. Lauren and M. Nordman (eds). *Special Language: From Humans Thinking to Thinking Machines*. Clevedon: Multilingual Matters.
- Laufer, B. (1990).Sequence' and `Order' in the development of L2 lexis: some evidence from lexical confusions. *Applied Linguistics*, 11(3), 281-296.
- Laufer, B. (1992). Corpus-based versus lexicographer examples in comprehension and production of new words. *EURALEX '92 – Proceedings*,71-76.

- Laufer, B. (1992). How much lexis is necessary for reading comprehension? In L. Arnaud & H. Bejoint (Eds.), *Vocabulary and applied linguistics* (pp. 126-312). London: Macmillan.
- Laufer, B. (1992a). Reading in a foreign language: how does L2 lexical knowledge interact with the reader's general academic ability? *Journal of Research in Reading* 15(2), 95-103.
- Laufer, B. (1992b). They think they know them, but they don't. *English - A World Language* 2, 1: 79-85
- Laufer, B. (1992c). How much lexis is necessary for reading comprehension? In L. Arnaud & H. Bejoint (Eds.), *Vocabulary and applied linguistics* (pp. 126-312). London: Macmillan
- Laufer, B. (1994). The lexical profile of second language writing: does it change over time? *RELC Journal* 25(2): 21-33.
- Laufer, B. (1997). The lexical plight in second language reading: words you don't know, words you think you know and words you can't guess. In Coady, J. & Huckin, T. (Eds.), *Second Language Vocabulary Acquisition*, pp 20-34. C.U.P.,
- Laufer, B. (2001). Reading, word-focused activities and incidental vocabulary acquisition in a second language. *Prospect* 16(3), 44-54.
- Laufer, B. (2003) Vocabulary acquisition in a second language: Do learners really acquire most vocabulary by reading? Some empirical evidence. *Canadian Modern Language Review* 59(4), 567-587.
- Laufer, B. (2009). Second language vocabulary acquisition from language input and from form focused activities. *Language Teaching*, 42, 341–354.
- Laufer, B. and Hill, M. (2000). What lexical information do L2 learners select in a CALL dictionary and how does it affect word retention? *Language Learning and Technology* 3(2). 58-76. (<http://llt.msu.edu/>)

- Laufer, B. and Nation, P. (1995). Vocabulary size and use: lexical richness in L2 written production. *Applied Linguistics*, 16, 307-322.
- Laufer, B. and Nation, P. (1999). A vocabulary-size test of controlled productive ability. *Language Testing* 16, 33-51.
- Laufer, B. and Sim, D.D. (1985). Measuring and explaining the reading threshold needed for English for academic purposes texts. *Foreign Language Annals* 18(5), 405-411.
- Laufer, B. and Waldman, T. (2011). Verb-noun collocations in second language writing: a corpus analysis of learners' English. *Language Learning*, 61(2), 647–672. DOI: 10.1111/j.1467-9922.2010.00621.x
- Laufer, B.& Roitblat-Rozovski, B. (2011). Incidental Vocabulary Acquisition: the Effects of Task Type, Word Occurrence and Their Combination. *Language Teaching Research*. doi: 10.1177/1362168811412019
- Lewis, M. (1997). Pedagogical implications of the lexical approach. In Coady and Huckin: 255-270.
- Liesveld, C.M. 1988. *The influence of interest, vocabulary difficulty and print size on the situation-specific reading self-efficacy of high achieving and low achieving high school readers*. (Doctoral dissertation). The University of Nebraska – Lincoln.
- Lin, Z. 2002. *Reading Comprehension and a Computerized Analytical Model: Thoretical Possibility, Technical Feasibility and Intrinsic Limitations*. (Doctoral dissertation). Retrieved from: <http://digthesis.ballarat.edu.au/adt/uploads/approved/adt-ADT20060424.122206/public/01front.pdf> .
- Linderholm, T., Everson, M. G., van den Broek, P., Mischinski, M., Crittenden, A., & Samuels, J. (2000). Effects of causal text revisions on more- and less-skilled readers' comprehension of easy and difficult texts. *Cognition & Instruction*, 18, 525-556.
- Liontas, J.I. (2002). Transactional Idiom Analysis: Theory and Practice. *Journal of Language and Linguistics*, 1(1), 17-53.

- Lipson, M. (1995). The effect of semantic mapping instruction on prose comprehension of below-level college readers. *Reading Research and Instruction*, 34, 367-378.
- Litz, David. (2005). [Textbook Evaluation and ELT Management: A South Korean Case Study](http://www.asian-efl-journal.com/Litz_thesis.pdf). Retrieved from: www.asian-efl-journal.com/Litz_thesis.pdf
- Lively, B. A. and S. L. Pressey. (1923). A method for measuring the 'vocabulary burden' of textbooks. *Educational Administration and Supervision*, 9, 389-398.
- Lunzer, E., Waite, M. & Dolan, T. (1979). Comprehension and comprehension tests. In Lunzer, E. & Gardner, K (Eds.), *The effective use of Reading*. London: Heinemann Educational.
- MacGinitie, W. and R. Tretiak. 1971. Sentence depth measures as predictors of reading difficulty. *Reading Research Quarterly* 6, 364-376.
- MacWhinney, B. (1995). *The CHILDES Project: Tools for Analyzing Talk* (2nd ed). Erlbaum, Hillsdale, New Jersey.
- MacWhinney, B. and Snow, C. E. (1990). The Child Language Data Exchange System: an Update. *Journal of Child Language*, 17, 457-72.
- Mallikarjun, B. (2002). An Introduction to the Selection, Gradation, and Presentation of Words in First, Second, and Foreign Language Teaching. *Language in India*. Vol. 2. Retrieved from: [www. http://www.languageinindia.com/nov2002/vocabulary.html](http://www.languageinindia.com/nov2002/vocabulary.html)
- Mandler, J.M & Johnson, M.S. (1977). Remembrance of things parsed: story structure and recall. *Cognitive Psychology*, 9, 111-151.
- Mandler, J.M. (1984). *Stories, scripts and scenes: Aspects of schema theory*. Hillsdale, NJ: Erlbaum
- Matsuoka, W. & Hirsh, D. (2010). Vocabulary learning through reading: Does an ELT course book provide good opportunities? *Reading in a Foreign Language* 22(1).
- Maxwell, M. (1997). *Improving student learning skills*. Clearwater, FL: H & H.
- McCarthy, M. (1990). *Vocabulary*. Oxford: Oxford University Press.

- McConkie, G. W. and Rayner, K. (1976). Asymmetry of the perceptual span in reading. *Bulletin of the Psychonomic Society*, 8, 365-368.
- McKeown, M. G., Beck, I. L., Omanson, R. C. & Perfetti, C. A., (1983). The effects of long-term vocabulary instruction on reading comprehension: A replication. *Journal of Reading Behavior*, 15, 3-18.
- McKeown, M.G. (1985) The acquisition of word meaning from context by children of high and low ability. *Reading Research Quarterly* 20, 482-496.
- McLaughlin, G. H. (1968). Proposals for British readability measures. *Third international reading symposium eds.* Brown and Downing. London: Cassell.
- McLaughlin, H. (1969). Smog grading: a new readability formula. *Journal of Reading* 12, 639-646.
- McNamara, D. S., Kintsch, E., Butler-Songer, N., & Kintsch, W. (1996). Are good texts always better? Interactions of text coherence, background knowledge, and levels of understanding in learning from text. *Cognition and Instruction*, 14, 1–43.
- Meade, C.D. & Smith, C.F. (1991). Readability formulas: cautions and criteria. *Patient Education & Counseling* 17, 153-158.
- Meara, P. & Jones, G. (1990). Eurocentres Vocabulary Size Test, Version E1.1/K10. Zurich: Eurocentres Learning Service.
- Meara, P. (1993). Tintin and the world service: A look at lexical environments. *IATEFL Annual Conference Report*, 32-37.
- Meara, P. (1996). The dimensions of lexical competence. In G. Brown, K. Malmkjaer and J. Williams, (Eds.), *Performance and Competence in Second Language Acquisition*: 35-53. Cambridge: Cambridge University Press.
- Mezynski, K. (1983). Issues concerning the acquisition of knowledge: Effects of vocabulary training on reading comprehension. *Review of Educational Research*, 53 (2), 253-279.

- Miller, J. F. and Chapman, R. (1993). *SALT: Systematic Analysis of Language Transcripts, Version 3.0*. University Park Press, Baltimore.
- Mirhassani, A., & Khosravi, A. A. (2002). Issues on reading comprehension. *Roshd Quarterly, 16*, 41-53.
- Miura, T. (2005). Interview with Paul Nation: The past, present, and future of second language vocabulary acquisition. *The Language Teacher, 29*(7), 11-14.
- Munby, J. (1979). Teaching intensive reading skills. In R. Mackay, B. Barknan, & R.R. Jordan (Eds.), *Reading in a second language* (pp. 142-159). Rowley, MA: Newbury House.
- Muth, K. D. (Ed.). (1989). *Children's comprehension of narrative and expository text: Research into Practice*. Newark, DE: International Reading Association.
- Nagy, W. E., Herman, P. A., & Anderson, R. C. (1985). Learning words from context. *Reading Research Quarterly, 20*(2), 233-253.
- Nagy, W., Anderson, R., Schommer, M., Scott, J.A., & Stallman, A. (1989). Morphological families in the internal lexicon. *Reading Research Quarterly, 24*, 262-281.
- Nation, I. S. P. (1994). *New ways in teaching vocabulary*. Alexandria, VA: TESOL, Inc.
- Nation, I. S. P. (2001). *Learning vocabulary in another language*. Cambridge: Cambridge University Press.
- Nation, I. S. P. (2006). How large a vocabulary is needed for reading and listening? *The Canadian Modern Language Review, 63*, 59-82.
- Nation, I. S. P. (2009). *Teaching vocabulary: Strategies and techniques*. Cambridge: Cambridge University Press.
- Nation, I. S. P., & Wang, M. (1999). Graded readers and vocabulary. *Reading in a Foreign Language, 12*, 1999.

- Nation, I.S.P. & Heatley, A. (2002). *RANGE* [Computer software]. Retrieved from <http://www.vuw.ac.nz/lals/staff/paul-nation/nation.aspx>
- Nation, I.S.P. (1990). *Teaching and Learning Vocabulary*. New York: Heinle and Heinle.
- Nation, I.S.P. and Hwang K. (1995). Where would general service vocabulary stop and special purposes vocabulary begin? *System* 23, 1: 35-41.
- Nation, I.S.P., & Waring, R. (1997). Vocabulary size, text coverage and word lists. In N. Schmitt & M. McCarthy (Eds.), *Vocabulary: Description, acquisition and pedagogy* (pp. 6–19). Cambridge, Cambridge University Press.
- National Reading Panel. (2000). *Teaching Children to read: An evidence-based assessment of the scientific research literature on reading and its implications for reading instruction*. National Institute of Child Health and Human Development, Washington, D.C.
- Nichols, C.N. (2007). *The Effects of Three Methods of Introducing Elementary Students Vocabulary: Traditonal, Friendly Definitions and Parsing*. (Doctoral Dissertation). University of Pittsburgh.
- Nilagupta, Sirirat. (1975). *The Relationship Between Text Readability and Student Reading Level and its Effect on College Achievement*. (Doctoral Dissertation). The Florida State University.
- Nisbet, J.D. (1960). Frequency counts and their uses. *Educational Research* 3, 51-64.
- Nuttall, C. (1982). *Teaching reading skills in a foreign language*. London: Heinemann Educational.
- Nuttall, Ch. (1996): *Teaching reading skills in a foreign language*. Macmillan-Heineman.
- Ogden C. K. (1932). *The Basic Vocabulary: A Statistical Analysis*. London, Routledge & Kegan Paul.
- Ozuru, Y., Dempsey, K., Sayroo, J., & McNamara, D. S. (2005). Effects of text cohesion on comprehension of biology texts. *Proceedings of the 27th Annual Meeting of the Cognitive Science Society* (pp. 1696-1701). Hillsdale, NJ: Erlbaum.

- Palmer, H. (1930). *Interim Report on Vocabulary Selection submitted to the Seventh Annual Conference of English Teachers under the auspices of the Institute for Research in English Teaching*. Tokyo: IRET.
- Paulson, E. & Goodman, K. (1999). *Influential studies in eye movement research*. Retrieved from <http://www.readingonline.org>.
- Pearson, P. D. (Ed.). (1984). *Handbook of Reading Research*. New York: Longman.
- Pearson, P.D., & Stephens, D. (1994). Learning about literacy: A 30-year journey. In R.B. Ruddell, M.R. Ruddell, & H. Singer (Eds.), *Theoretical models and Processes of Reading* (pp. 33-68). Newark, DE: International Reading Association.
- Pillsbury, W.B. (1897). A Study in Apperception. *American Journal of Psychology*, 8, 315-393.
- Pressley, M. (1998). *Reading instruction that works*. New York: Guilford.
- Hayden, M. B. , Prichard. J. J. (2008). Assessing the readability of freeware end-user licensing Agreements. Retrieved from: . *Issues in Information System* 9(1). Retrieved from: www.ric.edu/faculty/mhayden/
- Pritchard, R. (1990). The effects of cultural schemata on reading processing strategies. *Reading Research Quarterly*, 25(4), 273-95, 1990.
- Qian, D. (1999). Assessing the roles of depth and breadth of vocabulary knowledge in reading comprehension. *Canadian Modern Language Review* 56(2), 282-307.
- Qian, D. (2002). Investigating the relationship between vocabulary knowledge and academic reading performance: an assessment perspective. *Language Learning* 52(3), 513-536.
- Rabin, A. T. (1988). Determining difficulty levels of text written in languages other than English., In L. Zakaluk and S. J. Samuels (Eds.), *Readability: Its past, present, and future*. Newark, DE: International Reading Association.
- RAND Reading Study Group. (2002). *Reading for understanding: Toward an R&D program in reading comprehension*. Santa Monica, CA: RAND.

- Read, J. (2000) *Assessing Vocabulary*. Cambridge: Cambridge University Press.
- Redish, J. C. 2000. Readability formulas have even more limitations than Klare discusses. *ACM Journal of Computer documentation* 24(3), 132-137.
- Reicher, G.M. (1967). Perceptual Recognition as a Function of Meaningfulness of Stimulus Material. *Journal of Experimental Psychology* 81, 274-280.
- Richards, J.C. (1974). Word lists: problems and prospects. *RELC Journal* 5(2), 69-84.
- Richards, J.C. (1976). The role of vocabulary teaching. *TESOL Quarterly* 10(1),77-89.
- Richeck, M., List, A., & Learner, J. (1983). *Reading Problems, diagnosis, and remediation*. Prentice-Hall, Inc.
- Robinson, A.H. (1963) A study of techniques of word identification. *Reading Teacher* 16, 238-242.
- Roller, C. M. (1990). The interaction between knowledge and structure variables in the processing of expository prose. *Reading Research Quarterly*, 25, 79–89.
- Roller, C.M. (1988). Transfer of cognitive academic competence and L2 reading in a rural Zimbabwean primary school. *TESOL Quarterly*, 22, 303-318.
- Rosenblatt, L. (1994), The transactional theory of reading and writing. In R.B. Ruddell, M.R. Ruddell, & H. Singer (eds.), *Theoretical models and processes of reading, 4th edition*. Newark, DE: International Reading Association.
- Rosenshine, B.V. (1980). Skill hierarchies in reading comprehension. In R.J. Spiro, B.C. Bruce, & W.F. Frewer (Eds.), *Theoretical issues in reading comprehension* (pp. 535-554). Hillsdale, NJ: Lawrence Erlbaum.
- Royer, J. M., & Carlo, M. S. (199). Transfer of comprehension skills from native to second language. *Journal of Reading*, 34, 450-455.
- Rubin, A. (1985). How useful are readability formulas? In J. Osborn, P. T. Wilson, & R. C. Anderson (Eds.), *Reading education: Foundations for a literate America* (pp. 61–77). Lexington, MA: Lexington Books.

- Ruddell, M.R. (1994). Vocabulary knowledge and comprehension: a comprehension-process view of complex literacy relationships. In R.B. Ruddell, M.R. Ruddell and H. Singer (eds), *Theoretical Models and Processes of Reading, 4th edition* (pp. 414-447). International Reading Association, Newark, Del.
- Ruddell, R.B., Ruddell, M.R., & Singer, H., (Eds.). (1994). *Theoretical models and processes of reading, 4th edition*. Newark, DE: International Reading Association.
- Rumelhart, D. E. (1977a). Toward an interactive model of reading. In S. Dornic (ed). Paper presented at the *Attention and Performance VI*. Hillsdale, New Jersey: Lawrence Erlbaum Associates.
- Rumelhart, D. E. (1977b). *Introduction to Human Information Processing*. New York: Wiley.
- Rumelhart, D.E. (1975). Notes on a schema for stories. In D.G. Bobrow & A.M. Collins (Eds.) *Representation and understanding*. New York: Academic Press.
- Rumelhart, D.E. (1980). Schemata: the building blocks of cognition. In Spiro, R.J., Bruce, B. & Brewer, W. (Eds.) *Theoretical Issues in Reading Comprehension*. Hillsdale, N.J.: Erlbaum.
- Rumelhart, D. E. (1984). Understanding understanding. In J. Flood (Ed.) *Understanding Reading Comprehension* (pp. 1-20). Newark, Delaware: International Reading Association.
- Samuels, S.J., & Kamil, M.L. (1984). Models of the reading process. In P.D. Pearson, R. Barr, M.L. Kamil, & P. Mosenthal (Eds.), *Handbook of reading research* (pp. 185–224). New York: Longman.
- Sanders, M.S. & McCormick, E. J. (1993). Human Error, Accidents, and Safety. In: Sanders, M.S. McCormick, E. J. *Human Factors in Engineering and Design*. 7th ed. New York: McGraw-Hill, 1993. chap. 20, p. 655 - 695.
- Sanders, T. J. M., & Noordman, L. G. M. (2000). The role of coherence relations and their linguistic markers in text processing. *Discourse Processes*, 29(1), 37-60.

- Saragi, T., Nation, I. S. P., & Meister, G. F. (1978). Vocabulary learning and reading. *System*, 6, 72–78.
- Sawyer, M. H. (1991). A review of the research in revising instructional text. *Journal of Reading Behavior*, 23, 307-333.
- Schmitt, N. (1999) The relationship between TOEFL vocabulary items and meaning, association, collocation and word class knowledge. *Language Testing* 16(2), 189-216.
- Schmitt, N. (2000) . *Vocabulary in Language Teaching*. Cambridge University Press, Cambridge.
- Schmitt, N. and M. McCarthy (eds.). (1997). *Vocabulary: Description, Acquisition and Pedagogy*. Cambridge: Cambridge University Press.
- Schmitt, N. and P. Meara. (1997). Researching vocabulary through a word knowledge framework: word associations and verbal suffixes. *SSLA*, 19, 17-36.
- Schmitt, N. & Schmitt, D. & Clapham, C. (2001). Developing and exploring the behaviour of two new versions of the Vocabulary Levels Test. *Language Testing* 18, 55-88.
- Schoonen, R., Hulstijn, J, & Bossers, B. (1998). Metacognitive and language-specific knowledge in native and foreign language reading comprehension: An Empirical study among Dutch students in Grades 6, 8, and 10. *Language Learning*, 48(1), 71-106.
- Selzer, Jack. (1981). Readability is a four-letter word. *Journal of Business Communication* 18(4), 23-34.
- Sherman, A. L. (1893). *Analytics of literature: A manual for the objective study of English prose and poetry*. Boston: Ginn and Co.
- Shokrpour N. S. (2005). Comparison of Three Methods of Assessing Difficulty. *Asian EFL Journal*. Retrieved from: www.asian-efl-journal.com/pta_august_05_ns.php

- Sinclair, J.M and A. Renouf (1988). A lexical syllabus for language teaching. In R. Carter and M. McCarthy (eds.) *Vocabulary and Language Teaching*. Longman: London. 140-160.
- Singer. H. & R. B. Ruddell (Eds.). (1976). *Theoretical models and process of reading*. Newark, DE: International Reading Association.
- Singh, J. (2000). The readability of HIV/AIDS education materials. *AIDS Education and Prevention* 12(3), 214-224.
- Singhal, M. (1998). A comparison of L1 and L2 Reading: Cultural differences and schema. *The Internet TESL Journal*, IV(10), 1-6. Retrieved from <http://www.aitech.ac.jp/~iteslj/Articles/Singhal-ReadingL1L2.html>.
- Smith, F. (1971). *Understanding reading: A psycholinguistic analysis of reading and learning to read*. New York: HOLT, Rinehart and Winston, Inc.
- Smith, F. (1973). *Psycholinguistics and reading*. New York: Holt, Rinehart & Winston.
- Smith, F. (1978). *Understanding reading (2nd ed.)*. New York: Holt, Rinehart, and Winston
- Smith, F. (1985). *Reading without nonsense (2nd ed.)*. New York: Teachers College Press.
- Smith, H P. and Dechant, E.V. (1961). *Psychology in Teaching Reading*. Englewood Cliffs: Prentice-Hall, Inc.
- Smith, S. (1998). *Reviewer's checklist for health information materials*. Retrieved from: <http://www.PrenatalEd.com/evaluate.html>.
- Snow, C. (2002): Reading for understanding: Toward a R&D programme in reading comprehension. *Rand Research Brief*. Retrieved from: www.rand.org/multi/achievement
- Squires, D. A., Huitt, W. G., & Segars. J. K. (1983). *Effective schools and classrooms*. Alexandria, VA: Association for Supervisor and Curricular Development.

- Stahl, S.A. (1990). Beyond the instrumentalist hypothesis: some relationships between word meanings and comprehension. *Technical report N°. 505 of the Center for the Study of Reading*, University of Illinois at Urbana-Champaign.
- Stahl, S.A. and Fairbanks, M.M. (1986) The effects of vocabulary instruction: a model-based meta-analysis. *Review of Educational Research* 56, 1: 72-110
- Stanovich, Keith E. (1980). Toward an Interactive- Compensatory Model of Individual Differences in the Development of Reading Fluency, *Reading Research Quarterly*, 16, 32-71.
- Steffensen, M. S. (1987). The effects of context and culture on children's 12 reading: A review. In J. Devine, P. L. Carrell, & D. E. Eskey (Eds.), *Research in reading in English as a second language* (pp. 43–54). Washington, DC: TESOL.
- Stenner, A. J. & Burdick, D. S. (1997). The objective measurement of reading comprehension in response to technical questions raised by the California department of education technical study group. Durham, NC: MetaMetrics, Inc.
- Stenner, A.J. (2004). Measuring reading comprehension with the Lexile Framework. Durham, NC: Metametrics, Inc.
- Stephens, C. (2000). *All About Readability*. Retrieved from:
<http://plainlanguage.com/newreadability.html>
- Sternberg, R.J. (1987). Most Vocabulary is learned from context. In M.McKeown & M. Curtis (Eds.), *The Nature of Vocabulary Acquisition* (pp. 89-106). Hillsdale, NJ: Erlbaum
- Stockwell, R. & Minkova, D. (2001). *English words: History and structure*. New York: Cambridge University Press.
- Strother, J. & Ulijn, J. (1987) Does syntactic rewriting affect English for Science and Technology (EST) text comprehension? In: Devine, J., Carrell, P. & Eskey, D (Eds.) *Research in reading in English as a second language*. Washington, DC: TESOL, 1987, pp.89-101.

- Swaffar, J.K., Arens, K., & Byrnes, H.(1991). *Reading for meaning: An integrated approach to language learning*. Englewood Cliffs, NJ: Prentice Hall.
- Taillefer, G. E. (1996). L2 Reading Ability: Further Insight into the Short-circuit Hypothesis. *Modern Language Journal*. 80. 4. 461–477.
- Taylor, B. and Hiebert, E. (1994). Early literacy interventions: Aims and issues. In E. Hiebert and B. Taylor, *Getting reading right from the start*. Boston: Allyn and Bacon.
- Taylor, D. S. (1985). *Teaching reading for comprehension in the context of English as a second or foreign language*. Retrieved 16 January, 2008 from <http://education.leeds.ac.uk/~dst/inted/read.htm>
- Taylor, W. 1953. Cloze procedure: A new tool for measuring readability. *Journalism Quarterly* 30, 415-433.
- Thorndike, E.L. (1917). Reading as reasoning. *Journal of Educational Psychology*, 8, 323-332.
- Thorndike, E.L. (1921). *The Teacher's Word Book*. New York: Teacher's College, Columbia University.
- Thorndike, E.L. (1932) . *Teacher's Word Book of 20,000 Words*. Teachers College, Columbia.
- Thorndike, E.L. and Lorge, I. (1944) . *The Teacher's Word Book of 30,000 Words*. Teachers College, Columbia University.
- Thorndike, R. L. (1973). *Education: A first book*. New York: Arno Press.
- Torki, S. (2001). *A Study of the Relationship Between Teachers' Intention and Learners' Attention: The Case of Vocabulary Teaching and Learning In an Algerian EFL Class*. Unpublished Magister dissertation. University of Algiers.
- Torki, S. (2011). Teachers' Intention vs. Learners' Attention: Do Learners Attend to What Teachers Want Them to Attend to in an EFL Vocabulary Class? *Asian EFL Journal*, 12(2), 337-362.

- Trabasso, T. (1981). On the making of inferences during reading and their assessment. In J. T. Guthrie (Ed.), *Comprehension and teaching: Research reviews*. (pp. 56-76). Newark, Delaware: International Reading Association.
- Ulijn, J. M. & Salager-Meyer, F. (1998). The professional reader and the text: Insights from L2 research. *Journal of Research in Reading*, 21(2), 79-95.
- Ulijn, J.M. & Kempen, G.A.M. (1976). The role of the first language in second language reading comprehension: Some experimental evidence. In G. Nickel (Ed.). *Proceedings of the 4th International Congress of Applied Linguistics, vol 1*. Hochschulverlag, Stuttgart, 1976, pp. 495-507.
- Ulijn, J.M. & Strother J. B. (1990). The effect of syntactic simplification on reading EST texts as L1 and L2. *Journal of research in reading* 1,38-54.
- Underwood, N. R. and McConkie, G. W. (1985). Perceptual span for letter distinctions during reading. *Reading Research Quarterly*, 20, 153-162.
- Urquhart, S. and Weir, C. (1998). Reading in a second language: Process, product and practice. Longman.
- Vogel, M. and Washburne C. (1928). An objective method of determining grade placement of children's reading material. *Elementary school journal*,28, 373-381.
- Wang, Yan. (2011). The Effect of Text Length on Students' Reading Comprehension. *Theory and Practice in Language Studies*, Vol. 1, No. 9, pp. 1160-1166. doi:10.4304/tpls.1.9.1160-1166.
- Waring, R. and Nation, P. (2004). Second language reading and incidental vocabulary learning. *Angles on the English-speaking World* 4, 97-110.
- Weaver, P. A. (1994). Improving Reading Comprehension: Effects of Sentence Organization Instruction. *Reading Research Quarterly*, 15, 129-146.
- Webb, S. (2007). The effects of repetition on vocabulary knowledge. *Applied Linguistics*, 28, 46-65.

- Weber, R.M. (1984): Reading: United States. *Annual review of applied linguistics* 4, (p.111-123).
- Weir, G. R. S. and Ritchie, C. (2006) Estimating readability with the Strathclyde readability measure. In: *ICT in the Analysis, Teaching and Learning of Languages, reprints of the ICTATLL Workshop 2006*. Glasgow, UK.
- West, M. (1953) . *A General Service List of English Words*. Longman, London: Green & Co.
- Widdowson, H.G. (1978). *Teaching language as communication*. Oxford: Oxford University Press.
- Widdowson, H.G. (1983). *Learning purpose and language use*. Oxford: Oxford University Press.
- Willows, D. M., Borwick, D. & Hayvren, (1981). *Reading research: Advances in theory and practice*, 2nd edition. London: Academic Press.
- Wimmer, D. Roger & Dominick, R. Joseph., (2005). *Mass Media Research: An Introduction (8th ed.)*. CA: Thomson Wadsworth.
- Wright, B. D. and Stenner, A. J. (1998). *Readability and Reading Ability*. Paper presented to the Australian Council on Education Research
- Xue, G. & Nation, P. (1984). A university word list. *Language Learning & Communication* 32, 215-219.
- Yorio, C. (1971) Some sources of reading problems for foreign language learners. *Language Learning*. 21(1), 107-115, 1971.
- Zabrucky, K. M., & Moore, D. (1999). Influences of text genre on adults' monitoring of understanding and recall. *Educational Gerontology*, 25, 691 - 710.
- Zakaluk, B.L. & Samuels, S.J. (1988). Toward a new approach to predicting text comprehensibility. In: Zakaluk, B.L. & Samuels, S.J. (Eds.) *Readability: Its Past, Present and Future* (pp121-143. Newark, DE: International Reading Association.

Zipf, G. (1949) . *Human Behavior and the Principle of Least Effort: An Introduction to Human Ecology*. New York: Hafner.

Relevant websites

<http://www.lex tutor.ca>

<http://www.nontube.com/products/text-master>

<http://www.victoria.ac.nz/lals/staff/paul-nation.aspx>

APPENDICES

APPENDIX I

Academic Word List

(Source: <http://www.victoria.ac.nz/lals/resources/academicwordlist/sublists.aspx>)

Each word in italics is the most frequently occurring member of the word family in the Academic Corpus. For example, *analysis* is the most common form of the word family analyse. British and American spelling is included in the word families, so contextualise and contextualize are both included in the family *context*.

Sublist 1 contains the most common words in the AWL.

Sublist 2 contains the next most common words, and so on.

There are 60 families in each sublist, except for sublist 10 which has 30.

Sublist 1 of the Academic Word List

<i>analyse</i>	reassessing	consisted	uncontextualized	redefining
analysed	reassessment	consistency	contract	undefined
analyser	unassessed	consistent	contracted	derive
analysers	assume	consistently	contracting	derivation
analyses	assumed	consisting	contractor	derivations
analysing	assumes	consists	contractors	derivative
<i>analysis</i>	assuming	inconsistencies	contracts	derivatives
analyst	assumption	inconsistency	create	derived
analysts	assumptions	inconsistent	created	derives
analytic	authority	constitute	creates	deriving
analytical	authoritative	constituencies	creating	distribute
analytically	authorities	constituency	creation	distributed
analyze	available	constituent	creations	distributing
analyzed	availability	constituents	creative	<i>distribution</i>
analyzes	unavailable	constituted	creatively	distributional
analyzing	benefit	constitutes	creativity	distributions
approach	beneficial	constituting	creator	distributive
approachable	beneficiary	constitution	creators	distributor
approached	beneficiaries	constitutions	recreate	distributors
approaches	benefited	<i>constitutional</i>	recreated	redistribute
approaching	benefiting	constitutionally	recreates	redistributed
unapproachable	benefits	constitutive	recreating	redistributes
area	concept	unconstitutional	data	redistributing
areas	conception	context	define	redistribution
assess	concepts	contexts	definable	economy
assessable	conceptual	contextual	defined	economic
assessed	conceptualisation	contextualise	defines	economical
assesses	conceptualise	contextualised	defining	economically
assessing	conceptualised	contextualising	definition	economics
assessment	conceptualises	uncontextualised	definitions	economies
assessments	conceptualising	contextualize	redefine	economist
reassess	conceptually	contextualized	redefined	economists
reassessed	consist	contextualizing	redefines	uneconomical

environment	financier	interpretation	majorities	researched
environmental	financiers	interpretations	majority	researcher
environmentalist	financing	interpretative	method	researchers
environmentalists	formula	interpreted	methodical	researches
environmentally	formulae	interpreting	methodological	researching
environments	formulas	interpretive	methodologies	respond
establish	formulate	interprets	methodology	responded
disestablish	formulated	misinterpret	methods	respondent
disestablished	formulating	misinterpretation	occur	respondents
disestablishes	formulation	misinterpretations	occurred	responding
disestablishing	formulations	misinterpreted	occurrence	responds
disestablishment	reformulate	misinterpreting	occurrences	<i>response</i>
<i>established</i>	reformulated	misinterprets	occurring	responses
establishes	reformulating	reinterpret	occurs	responsive
establishing	reformulation	reinterpreted	reoccur	responsiveness
establishment	reformulations	reinterprets	reoccurred	unresponsive
establishments	function	reinterpreting	reoccurring	role
estimate	functional	reinterpretation	reoccurs	roles
estimated	functionally	reinterpretations	percent	section
estimates	functioned	involve	percentage	sectioned
estimating	functioning	involved	percentages	sectioning
estimation	functions	involvement	period	sections
estimations	identify	involves	periodic	sector
over-estimate	identifiable	involving	periodical	sectors
overestimate	identification	uninvolved	periodically	significant
overestimated	<i>identified</i>	issue	periodicals	insignificant
overestimates	identifies	issued	periods	insignificantly
overestimating	identifying	<i>issues</i>	policy	significance
underestimate	identities	issuing	policies	significantly
underestimated	identity	labour	principle	signified
underestimates	unidentifiable	labor	principled	signifies
underestimating	income	labored	principles	signify
evident	incomes	labors	unprincipled	signifying
evidenced	indicate	laboured	proceed	similar
<i>evidence</i>	indicated	labouring	procedural	dissimilar
evidential	indicates	labours	<i>procedure</i>	similarities
evidently	indicating	legal	procedures	similarity
export	indication	illegal	proceeded	similarly
exported	indications	illegality	proceeding	source
exporter	indicative	illegally	proceedings	sourced
exporters	indicator	legality	proceeds	sources
exporting	indicators	legally	process	sourcing
exports	individual	legislate	processed	specific
factor	individualised	legislated	processes	specifically
factored	individuality	legislates	processing	specification
factoring	individualism	legislating	require	specifications
<i>factors</i>	individualist	<i>legislation</i>	<i>required</i>	specificity
finance	individualists	legislative	requirement	specifics
financed	individualistic	legislator	requirements	structure
finances	individually	legislators	requires	restructure
<i>financial</i>	individuals	legislature	requiring	restructured
financially	interpret	major	research	restructures

restructuring	structuring	theories	variability	variant
structural	unstructured	theorist	variable	variants
structurally	theory	theorists	variables	variation
structured	theoretical	invariable	variably	variations
structures	theoretically	invariably	variance	
vary				
varied				
varies				
varying				

Sublist 3 of Academic Word List

alternative	constrains	corporate	dominates	illustration
alternatively	constraint	corporates	dominating	illustrations
alternatives	constraints	corporation	domination	illustrative
circumstance	unconstrained	corporations	emphasis	immigrate
circumstances	contribute	correspond	emphasise	immigrant
comment	contributed	corresponded	emphasised	immigrants
commentaries	contributes	correspondence	emphasising	immigrated
commentary	contributing	corresponding	emphasize	immigrates
commentator	contribution	correspondingly	emphasized	immigrating
commentators	contributions	corresponds	emphasizes	immigration
commented	contributor	criteria	emphasizing	imply
commenting	contributors	criterion	emphatic	implied
<i>comments</i>	convene	deduce	emphatically	<i>implies</i>
compensate	<i>convention</i>	deduced	ensure	implying
compensated	convenes	deduces	ensured	initial
compensates	convened	deducing	ensures	initially
compensating	convening	deduction	ensuring	instance
compensation	conventional	deductions	exclude	instances
compensations	conventionally	demonstrate	excluded	interact
compensatory	conventions	demonstrable	excludes	interacted
component	unconventional	demonstrably	excluding	interacting
componentry	coordinate	demonstrated	exclusion	interaction
components	coordinated	demonstrates	exclusionary	interactions
consent	coordinates	demonstrating	exclusionist	interactive
consensus	coordinating	demonstration	exclusions	interactively
consented	coordination	demonstrations	exclusive	interacts
consenting	coordinator	demonstrative	exclusively	justify
consents	coordinators	demonstratively	framework	justifiable
considerable	co-ordinate	demonstrator	frameworks	justifiably
considerably	co-ordinated	demonstrators	fund	justification
constant	co-ordinates	document	funded	justifications
constancy	co-ordinating	documentation	funder	justified
constantly	co-ordination	documented	funders	justifies
constants	co-ordinator	documenting	funding	justifying
inconstancy	co-ordinators	documents	funds	unjustified
inconstantly	core	dominate	illustrate	layer
constrain	cores	dominance	illustrated	layered
constrained	coring	dominant	illustrates	layering
constraining	cored	dominated	illustrating	layers

link	negates	publish	reliant	specifiable
linkage	negating	published	relied	specified
linkages	negatively	publisher	relies	specifies
linked	negatives	publishers	relying	specifying
linking	outcome	publishes	unreliable	unspecified
links	<i>outcomes</i>	publishing	remove	sufficient
locate	partner	unpublished	removable	sufficiency
located	partners	react	removal	insufficient
locating	partnership	reacted	removals	insufficiently
location	partnerships	reacts	removed	sufficiently
locations	philosophy	reacting	removes	task
relocate	philosopher	reaction	removing	tasks
relocated	philosophers	reactionaries	scheme	technical
relocates	philosophical	reactionary	schematic	technically
relocating	philosophically	reactions	schematically	technique
relocation	philosophies	reactive	schemed	<i>techniques</i>
maximise	philosophise	reactivate	schemes	<i>technology</i>
max	philosophised	reactivation	scheming	technological
maximised	philosophises	reactor	sequence	technologically
maximises	philosophising	reactors	sequenced	valid
maximising	philosophize	register	sequences	invalidate
maximisation	philosophized	deregister	sequencing	invalidity
maximize	philosophizes	deregistered	sequential	validate
maximized	philosophizing	deregistering	sequentially	validated
maximizes	physical	deregisters	sex	validating
maximizing	physically	deregistration	sexes	validation
maximization	proportion	registered	sexism	validity
maximum	disproportion	registering	sexual	validly
minor	disproportionated	registers	sexuality	volume
<i>minorities</i>	isproportionately	registration	sexually	volumes
minority	proportional	rely	shift	vol
minors	proportionally	reliability	shifted	
negate	proportionate	reliable	shifting	
negative	proportionately	reliably	shifts	
negated	proportions	reliance	specify	

Sublist 4 of the Academic Word List

access	accessible	adequacy	inadequate	apparent
accessed	accessing	adequately	inadequately	apparently
accesses	inaccessible	inadequacies	annual	approximate
accessibility	adequate	inadequacy	annually	approximated

approximately	contrast	hence	investigated	parameters
approximates	contrasted	hypothesis	investigates	phase
approximating	contrasting	hypotheses	investigating	phased
approximation	contrastive	hypothesise	investigation	phases
approximations	contrasts	hypothesised	investigations	phasing
attitude	cycle	hypothesises	investigative	predict
attitudes	cycled	hypothesising	investigator	predictability
attribute	cycles	hypothesize	investigators	predictable
attributable	cyclic	hypothesized	job	predictably
attributed	cyclical	hypothesizes	jobs	predicted
attributes	cycling	hypothesizing	label	predicting
attributing	debate	hypothetical	labeled	prediction
attribution	debatable	hypothetically	labeling	predictions
civil	debated	implement	labelled	predicts
code	debates	implementation	labelling	unpredictability
coded	debating	implemented	labels	unpredictable
codes	despite	implementing	mechanism	principal
coding	dimension	implements	mechanisms	principally
commit	dimensional	implicate	obvious	prior
commitment	dimensions	implicated	obviously	professional
commitments	multidimensional	implicates	occupy	professionally
commits	domestic	implicating	occupancy	professionals
committed	domestically	implication	occupant	professionalism
committing	domesticate	implications	occupants	project
communicate	domesticated	impose	occupation	projected
communicable	domesticating	<i>imposed</i>	occupational	projecting
communicated	domestics	imposes	occupations	projection
communicates	emerge	imposing	occupied	projections
communicating	<i>emerged</i>	imposition	occupier	projects
communication	emergence	integrate	occupiers	promote
communications	emergent	integrated	occupies	promoted
communicative	emerges	integrates	occupying	promoter
ommunicatively	emerging	integrating	option	promoters
uncommunicative	error	<i>integration</i>	optional	promotes
concentrate	erroneous	internal	options	promoting
concentrated	erroneously	internalise	output	promotion
concentrates	errors	internalised	outputs	promotions
concentrating	ethnic	internalises	overall	regime
concentration	ethnicity	internalising	parallel	regimes
confer	goal	internalize	paralleled	resolve
conference	goals	internalized	paralleled	resolution
conferences	grant	internalizes	paralleling	resolved
conferred	granted	internalizing	parallels	resolves
conferring	granting	internally	unparalleled	resolving
confers	grants	investigate	parameter	unresolved

retain	statistician	stressing	summaries	summarized
retained	statisticians	unstressed	summarise	summarizes
retaining	statistical	subsequent	summarised	summarizing
retainer	statistically	subsequently	summarises	undertake
retainers	statistics	sum	summarising	undertaken
retains	status	summation	summarisation	undertakes
retention	stress	summed	summarisations	undertaking
retentive	stressed	summing	summarization	undertook
series	stresses	sums	summarizations	
statistic	stressful	summary	summarize	

Sublist 5 of the Academic Word List

academy	amend	consult	redraft	expand
academia	amended	consultancy	redrafted	expanded
<i>academic</i>	amending	consultant	redrafting	expanding
academically	amendment	consultants	redrafts	expands
academics	amendments	consultation	enable	expansion
academies	amends	consultations	enabled	expansionism
adjust	aware	consultative	enables	expansive
adjusted	awareness	consulted	enabling	expose
adjusting	unaware	consults	energy	exposed
<i>adjustment</i>	capacity	consulting	energetic	exposes
adjustments	capacities	contact	energetically	exposing
adjusts	incapacitate	contactable	energies	exposure
readjust	incapacitated	contacted	enforce	exposures
readjusted	challenge	contacting	enforced	external
readjusting	challenged	contacts	enforcement	externalisation
readjustment	challenger	decline	enforces	externalise
readjustments	challengers	declined	enforcing	externalised
readjusts	challenges	declines	entity	externalises
alter	challenging	declining	entities	externalising
alterable	clause	discrete	equivalent	externality
alteration	clauses	discretely	equivalence	externalization
alterations	compound	discretion	evolve	externalize
altered	compounded	discretionary	evolution	externalized
altering	compounding	indiscrete	evolved	externalizes
alternate	compounds	indiscretion	evolving	externalizing
alternating	conflict	draft	evolves	externally
alters	conflicted	drafted	evolutionary	facilitate
unalterable	conflicting	drafting	evolutionist	facilitated
unaltered	conflicts	drafts	evolutionists	facilitates

facilities	logical	precisely	stylized
facilitating	logically	precision	stylizes
facilitation	logician	prime	stylizing
facilitator	logicians	primacy	substitute
facilitators	margin	psychology	substituted
facility	marginal	psychological	substitutes
fundamental	marginally	psychologically	substituting
fundamentally	margins	psychologist	<i>substitution</i>
generate	medical	psychologists	sustain
generated	medically	pursue	sustainable
generates	mental	pursued	sustainability
generating	mentality	pursues	sustained
generation	mentally	pursuing	sustaining
generations	modify	pursuit	sustains
image	modification	pursuits	sustenance
imagery	modifications	ratio	unsustainable
images	modified	ratios	symbol
liberal	modifies	reject	symbolic
liberalise	modifying	rejected	symbolically
liberalism	unmodified	rejecting	symbolise
liberalisation	monitor	rejection	symbolises
liberalised	monitored	rejects	symbolised
liberalises	<i>monitoring</i>	rejections	symbolising
liberalising	monitors	revenue	symbolism
liberalization	unmonitored	revenues	symbolize
liberalize	network	stable	symbolized
liberalized	networked	instability	symbolizes
liberalizes	networking	stabilisation	symbolizing
liberalizing	networks	stabilise	symbols
liberate	notion	stabilised	target
liberated	notions	stabilises	targeted
liberates	objective	stabilising	targeting
liberation	objectively	stabilization	targets
liberations	objectivity	stabilize	transit
liberating	orient	stabilized	transited
liberator	orientate	stabilizes	transiting
liberators	orientated	stabilizing	transition
liberally	orientates	stability	transitional
liberals	orientation	unstable	transitions
licence	orientating	style	transitory
licences	oriented	styled	transits
license	orienting	styles	trend
licensed	orients	styling	trends
licensing	reorient	stylish	version
licenses	reorientation	stylise	versions
unlicensed	perspective	stylised	welfare
logic	perspectives	stylises	whereas
illogical	precise	styling	
illogically	imprecise	stylize	

Sublist 6 of the Academic Word List

abstract	unattached	displayed	flexible	inputs
abstraction	author	displaying	flexibility	instruct
abstractions	authored	displays	inflexible	instruction
abstractly	authoring	diverse	inflexibility	instructed
abstracts	authors	diversely	furthermore	instructing
accurate	authorship	diversification	gender	instructions
accuracy	bond	diversified	genders	instructive
accurately	bonded	diversifies	ignorant	instructor
inaccuracy	bonding	diversify	ignorance	instructors
inaccuracies	bonds	diversifying	ignore	instructs
inaccurate	brief	diversity	ignored	intelligent
acknowledge	brevity	domain	ignores	intelligence
acknowledged	briefed	domains	ignoring	intelligently
acknowledges	briefing	edit	incentive	unintelligent
acknowledging	briefly	edited	incentives	interval
acknowledgement	briefs	editing	incidence	intervals
acknowledgemens	capable	<i>edition</i>	incident	lecture
aggregate	capabilities	editions	incidentally	lectured
aggregated	capability	editor	incidents	lecturer
aggregates	incapable	editorial	incorporate	lecturers
aggregating	cite	editorials	incorporated	lectures
aggregation	citation	editors	incorporates	lecturing
allocate	citations	edits	incorporating	migrate
allocated	cited	enhance	incorporation	migrant
allocates	citing	enhanced	index	migrants
allocating	cites	enhancement	indexed	migrated
allocation	cooperate	enhances	indexes	migrates
allocations	cooperated	enhancing	indexing	migrating
assign	cooperates	estate	inhibit	migration
assigned	cooperating	estates	inhibited	migrations
assigning	cooperation	exceed	inhibiting	migratory
assignment	cooperative	exceeded	inhibition	minimum
assignments	cooperatively	exceeding	inhibitions	ministry
assigns	co-operate	exceeds	inhibits	ministered
reassign	co-operated	expert	initiate	ministering
reassigned	co-operates	expertise	initiated	ministerial
reassigning	co-operation	expertly	initiates	ministries
reassigns	co-operative	experts	initiating	motive
unassigned	co-operatively	explicit	initiation	motivate
attach	discriminate	explicitly	initiations	motivated
attached	discriminated	federal	initiative	motivates
attaches	discriminates	federation	initiatives	motivating
attaching	discriminating	federations	initiator	motivation
attachment	discrimination	fee	initiators	motivations
attachments	display	fees	input	motives

unmotivated	reveal	utilise
neutral	revealed	utilisation
neutralisation	revealing	utilised
neutralise	reveals	utilises
neutralised	revelation	utilising
neutralises	revelations	utiliser
neutralising		utilisers
neutrality	scope	utility
neutralization	subsidy	utilities
neutralize	subsidiary	utilization
neutralized	subsides	utilize
neutralizes	subsidise	utilized
neutralizing	subsidised	utilizes
nevertheless	subsidises	utilizing
overseas	subsidising	
precede	subsidize	transform
preceded	subsidized	transformation
precedence	subsidizes	transformations
precedent	subsidizing	transformed
precedes	tape	transforming
<i>preceding</i>	taped	transforms
unprecedented	tapes	transport
presume	taping	transportation
presumably	trace	transported
presumed	traceable	transporter
presumes	traced	transporters
presuming	traces	transporting
presumption	tracing	transports
presumptions		underlie
presumptuous		underlay
rational		underlies
irrational		underlying
rationalisation		
rationalisations		
rationalise		
rationalised		
rationalises		
rationalising		
rationalism		
rationality		
rationalization		
rationalizations		
rationalize		
rationalized		
rationalizes		
rationally		
recover		
recoverable		
recovered		
recovering		
recovers		
recovery		

Sublist 7 of the Academic Word List

adapt	conversion	empiricism	inferring	prioritizes
adaptability	conversions	equip	infers	prioritizing
adaptable	converted	equipment	innovate	prohibit
adaptation	convertible	equipped	innovation	prohibited
adaptations	converting	equipping	innovated	prohibiting
adapted	converts	equips	innovates	prohibition
adapting	couple	extract	innovating	prohibitions
adaptive	coupled	extracted	innovations	prohibitive
adapts	coupling	extracting	innovative	prohibits
adult	couples	extraction	innovator	publication
adulthood	decade	extracts	innovators	publications
adults	<i>decades</i>	file	insert	quote
advocate	definite	filed	inserted	<i>quotation</i>
advocacy	definitely	files	inserting	quotations
advocated	definitive	filing	insertion	quoted
advocates	indefinite	finite	inserts	quotes
advocating	indefinitely	infinite	intervene	quoting
aid	deny	infinitely	intervened	release
aided	deniable	foundation	intervenes	released
aiding	denial	foundations	intervening	releases
aids	denials	globe	intervention	releasing
unaided	denied	<i>global</i>	interventions	reverse
channel	denies	globally	isolate	reversal
channelled	denying	globalisation	<i>isolated</i>	reversed
channelling	undeniable	globalization	isolates	reverses
channels	differentiate	grade	isolating	reversible
chemical	differentiated	graded	isolation	reversing
chemically	differentiates	grades	isolationism	reversals
chemicals	differentiating	grading	media	irreversible
classic	differentiation	guarantee	mode	simulate
classical	dispose	guaranteed	modes	simulated
classics	disposable	guaranteeing	paradigm	simulates
comprehensive	disposal	guarantees	paradigms	simulating
comprehensively	disposed	hierarchy	phenomenon	simulation
comprise	disposes	hierarchical	phenomena	sole
comprised	disposing	hierarchies	phenomenal	solely
comprises	dynamic	identical	priority	somewhat
comprising	dynamically	identically	priorities	submit
confirm	dynamics	ideology	prioritisation	submission
confirmation	eliminate	ideological	prioritise	submissions
confirmed	eliminated	ideologically	prioritised	submits
confirming	eliminates	ideologies	prioritises	submitted
confirms	eliminating	infer	prioritising	submitting
contrary	elimination	inference	prioritization	successor
contrarily	empirical	inferences	prioritize	succession
convert	empirically	inferred	prioritized	successions

successive
successively
successors

survive

survival
survived
survives
surviving
survivor
survivors *thesis*
theses

topic

topical
topics

transmit

transmission
transmissions
transmitted
transmitting
transmits

ultimate

ultimately

unique

uniquely
uniqueness

visible

visibility
visibly
invisible
invisibility

voluntary

voluntarily
volunteer
volunteering
volunteered
volunteers

Sublist 8 of the Academic Word List

abandon	automation	contradicting	dramatisations	infrastructure
abandoned	bias	contradiction	dramatist	infrastructures
abandoning	biased	contradictions	dramatists	inspect
abandonment	biases	contradictory	dramatization	inspected
abandons	biasing	contradicts	dramatizations	inspecting
accompany	unbiased	crucial	dramatize	inspection
<i>accompanied</i>	chart	crucially	dramatized	inspections
accompanies	charted	currency	dramatizes	inspector
accompaniment	charting	currencies	dramatizing	inspectors
accompanying	charts	denote	eventual	inspects
unaccompanied	uncharted	denotation	eventuality	intense
accumulate	clarify	denotations	<i>eventually</i>	intensely
accumulated	clarification	denoted	exhibit	intenseness
accumulating	clarified	denotes	exhibited	intensification
accumulation	clarifies	denoting	exhibiting	intensified
accumulates	clarifying	detect	exhibition	intensifies
ambiguous	clarity	detectable	exhibitions	intensify
ambiguities	commodity	detected	exhibits	intensifying
ambiguity	commodities	detecting	exploit	intension
unambiguous	complement	detection	exploitation	intensity
unambiguously	complementary	detective	exploited	intensive
append	complemented	detectives	exploiting	intensively
appendix	complementing	detector	exploits	manipulate
appended	complements	detectors	fluctuate	manipulated
appends	conform	detects	fluctuated	manipulates
appending	conformable	deviate	fluctuates	manipulating
appendices	conformability	deviated	fluctuating	manipulation
appendixes	conformance	deviates	fluctuation	manipulations
appreciate	conformation	deviating	fluctuations	manipulative
appreciable	conformed	deviation	guideline	minimise
appreciably	conforming	deviations	guidelines	minimised
appreciated	conformist	displace	highlight	minimises
appreciates	conformists	displaced	highlighted	minimising
appreciating	conformity	displacement	highlighting	minimize
appreciation	conforms	displaces	highlights	minimized
unappreciated	nonconformist	displacing	implicit	minimizes
arbitrary	nonconformists	drama	implicitly	minimizing
arbitrariness	nonconformity	dramas	induce	nuclear
arbitrarily	non-conformist	dramatic	induced	offset
automate	non-conformists	dramatically	induces	offsets
automatic	non-conformity	dramatise	inducing	offsetting
automated	contemporary	dramatised	induction	paragraph
automates	contemporaries	dramatising	inevitable	paragraphing
automating	contradict	dramatises	inevitability	paragraphs
automatically	contradicted	dramatisation	inevitably	plus

pluses	randomly	revision	terminate	vehicles
practitioner	randomness	revisions	terminal	via
practitioners	reinforce	schedule	terminals	virtual
predominant	reinforced	reschedule	terminated	virtually
predominance	reinforcement	rescheduled	terminates	visual
predominantly	reinforcements	reschedules	terminating	visualise
predominate	reinforces	rescheduling	termination	visualised
predominated	reinforcing	scheduled	terminations	visualising
predominates	restore	schedules	theme	visualisation
predominating	restoration	scheduling	themes	visualize
prospect	restored	unscheduled	thematic	visualized
prospective	restores	tense	thematically	visualizing
prospects	restoring	tension	thereby	visualization
radical	revise	tensely	uniform	visually
radically	revised	tenser	uniformity	widespread
radicals	revises	tensest	uniformly	
random	revising	tensions	vehicle	

Sublist 9 of the Academic Word List

	attains	recommences	devotion	formats
accommodate	unattainable	recommended	devotions	found
accommodated	behalf	recommencing	diminish	<i>founded</i>
accommodates	bulk	compatible	<i>diminished</i>	founder
accommodating	bulky	compatibility	diminishes	founders
<i>accommodation</i>	cease	incompatibility	diminishing	founding
analogy	ceased	<i>incompatible</i>	diminution	unfounded
analogies	ceaseless	concurrent	undiminished	inherent
<i>analogous</i>	<i>ceases</i>	concurrently	distort	inherently
anticipate	ceasing	confine	<i>distorted</i>	insight
<i>anticipated</i>	coherent	<i>confined</i>	distorting	insightful
anticipates	<i>coherence</i>	confines	distortion	<i>insights</i>
anticipating	coherently	confining	distortions	integral
anticipation	incoherent	unconfined	distorts	intermediate
unanticipated	incoherently	controversy	<i>duration</i>	manual
assure	coincide	controversies	erode	manually
<i>assurance</i>	coincided	controversial	eroded	manuals
assurances	coincides	controversially	erodes	mature
assured	coinciding	uncontroversial	eroding	immature
assuredly	coincidence	converse	<i>erosion</i>	immaturity
assures	coincidences	<i>conversely</i>	ethic	maturation
assuring	coincident	device	<i>ethical</i>	maturational
attain	coincidental	devices	ethically	matured
attainable	commence	devote	ethics	matures
<i>attained</i>	commenced	<i>devoted</i>	unethical	maturing
attaining	commences	devotedly	format	maturity
attainment	commencement	devotes	formatted	mediate
attainments	commencing	devoting	formatting	mediated

mediates	restraining	trigger	violate
mediating	restrains	triggered	violated
<i>mediation</i>	restraint	triggering	violates
medium	<i>restraints</i>	triggers	violating
military	unrestrained	unify	<i>violation</i>
minimal	revolution	unification	violations
minimalisation	revolutionary	<i>unified</i>	vision
minimalise	revolutionaries	unifies	visions
minimalises	revolutionise	unifying	
minimalised	revolutionised		
minimalising	revolutionises		
minimalist	revolutionising		
minimalists	revolutionist		
minimalistic	revolutionists		
minimalization	revolutionize		
minimalize	revolutionized		
minimalized	revolutionizes		
minimalizes	revolutionizing		
minimalizing	revolutions		
minimally	rigid		
mutual	rigidities		
mutually	rigidity		
norm	rigidly		
<i>norms</i>	route		
overlap	routed		
overlapped	routes		
overlapping	routing		
overlaps	scenario		
passive	scenarios		
passively	sphere		
passivity	spheres		
<i>portion</i>	spherical		
portions	spherically		
preliminary	subordinate		
preliminaries	subordinates		
protocol	subordination		
protocols	supplement		
qualitative	<i>supplementary</i>		
qualitatively	supplemented		
refine	supplementing		
refined	supplements		
refinement	suspend		
refinements	<i>suspended</i>		
refines	suspending		
refining	suspends		
relax	suspension		
relaxation	team		
<i>relaxed</i>	teamed		
relaxes	teaming		
relaxing	teams		
restrain	temporary		
restrained	temporarily		

Sublist 10 of the Academic Word List

<i>adjacent</i>	compilation	convincing	incline	<i>nonetheless</i>
<i>albeit</i>	compilations	convincingly	<i>inclination</i>	<i>notwithstanding</i>
<i>assemble</i>	compiled	unconvinced	inclinations	<i>odd</i>
assembled	compiles	<i>depress</i>	inclined	odds
assembles	compiling	depressed	inclines	<i>ongoing</i>
assemblies	<i>conceive</i>	depresses	inclining	<i>panel</i>
assembling	conceivable	depressing	<i>integrity</i>	panelled
assembly	conceivably	<i>depression</i>	<i>intrinsic</i>	panelling
<i>collapse</i>	conceived	<i>encounter</i>	intrinsically	panels
collapsed	conceives	encountered	<i>invoke</i>	<i>persist</i>
collapses	conceiving	encountering	<i>invoked</i>	persisted
collapsible	inconceivable	encounters	invokes	persistence
collapsing	inconceivably	<i>enormous</i>	invoking	persistent
<i>colleague</i>	<i>convince</i>	enormity	<i>levy</i>	persistently
colleagues	convinced	enormously	levies	persisting
<i>compile</i>	convinces	<i>forthcoming</i>	<i>likewise</i>	persists
<i>pose</i>				
posed				
poses				
posing				
<i>reluctance</i>				
reluctant				
reluctantly				
<i>so-called</i>				
<i>straightforward</i>				
<i>undergo</i>				
undergoes				
undergoing				
undergone				
underwent				
<i>whereby</i>				

APPENDIX II

The General Service List

(Source: www.lex tutor.ca/freq/lists_download/)

The first 1000 words of the General Service List

a	appear	between	chance	cover	dog	example
able	apply	beyond	change	creature	door	excellent
about	appoint	big	character	cross	double	except
above	arise	bill	charge	crowd	doubt	exchange
accept	arm	bird	chief	cry	down	exercise
accord	army	bit	child	current	draw	exist
account	around	black	choose	custom	dream	expect
across	arrive	blood	church	cut	dress	expense
act	art	blow	circle	dance	drive	experience
actual	article	blue	city	danger	drop	experiment
add	as	board	claim	dare	dry	explain
address	ask	boat	class	dark	due	express
admit	associatio	body	clean	date	during	extend
adopt	n	book	clear	daughter	each	eye
advance	at	born	clock	day	ear	face
advantag	attack	both	close	dead	early	fact
e	attempt	box	club	deal	earth	factory
affair	average	boy	coast	decide	east	fail
afford	away	branch	cold	declare	easy	fair
after	back	bread	college	deep	eat	faith
again	bad	breadth	colour	degree	edge	fall
against	ball	break	come	deliver	effect	familiar
age	bank	bridge	command	demand	effort	family
ago	bar	bright	common	describe	either	famous
agree	base	bring	company	desert	else	farm
air	battle	brother	compare	desire	employ	fashion
all	be	build	complete	destroy	end	fast
allow	bear	burn	concern	detail	enemy	father
almost	beauty	business	condition	determine	English	favourite
alone	because	but	connect	develop	enjoy	favour
along	become	buy	consider	die	enough	fear
already	bed	by	contain	difference	enter	feed
also	before	call	content	difficult	entire	feel
although	begin	can	continue	direct	equal	fellow
always	behind	capital	control	discover	escape	few
among	being	car	corner	disease	even	field
amount	believe	care	cost	distance	evening	figure
ancient	belong	carry	could	distinguis	event	fill
and	below	case	council	h	ever	find
animal	beneath	catch	count	district	every	fine
another	beside	cause	country	divide	everywher	finger
answer	best	centre	course	do	e	finish
any	better	certain	court	doctor	evil	fire

first	happy	its	manners	need	paint	proper
fish	hard	join	many	neighbour	paper	propose
fit	hardly	judge	march	neither	part	protect
fix	have	just	mark	never	party	prove
floor	he	keep	market	new	pass	provide
flower	head	kill	marry	next	past	public
follow	health	kind	mass	night	pay	pull
food	hear	know	material	no	peace	purpose
for	heat	lack	matter	none	people	put
force	heaven	lady	may	nor	perfect	quality
foreign	heavy	land	me	north	perhaps	quarter
forget	help	language	mean	not	permanent	quiet
form	her	large	measure	note	permit	quite
former	here	last	meet	nothing	person	race
forth	hide	late	member	notice	picture	raise
fortune	high	law	memory	now	piece	rank
forward	hill	lay	mention	nowhere	place	rate
free	his	lead	mere	number	plan	rather
friend	history	learn	metal	object	plant	reach
from	hold	least	middle	observe	play	read
front	home	leave	might	occasion	please	ready
full	honour	left	mile	of	point	real
further	hope	length	mind	off	political	reason
future	horse	less	mine	offer	poor	receive
gain	hot	let	minister	office	popular	recent
game	hour	letter	minute	often	population	recognize
garden	house	level	miss	oil	position	record
gate	how	library	modern	old	possess	red
gather	hullo	lie	moment	once	possible	reduce
general	human	life	money	one	post	refuse
gentle	hurrah	lift	month	only	pound	regard
get	husband	light	more	open	poverty	regular
give	I	like	moreover	operation	power	relation
glad	idea	likely	morning	opinion	practical	religion
glass	if	limit	most	opportuni	prepare	remain
go	impossibl	line	mother	ty	present	remark
God	e	listen	motor	or	preserve	remember
gold	in	little	mountain	order	press	repeat
good	inch	live	mouth	ordinary	pretty	reply
great	include	local	move	organize	prevent	report
green	increase	long	much	other	price	represent
ground	indeed	look	music	otherwise	print	respect
group	independe	lose	must	ought	private	rest
grow	nt	lost	my	ounce	problem	result
guard	influence	lot	name	our	produce	return
habit	instead	love	narrow	out of	product	rich
half	intend	low	nation	out	production	ride
hall	interest	machine	native	over	program	right
hand	into	main	nature	owe	programme	ring
handle	introduce	make	near	own	progress	rise
hang	iron	man	necessary	page	promise	river
happen	it	manner	neck	pain	proof	road

roll	shall	soon	strike	therefore	up	who
room	shape	sort	strong	these	upon	whose
rough	share	sound	struggle	they	use	why
round	shave	south	study	thing	usual	wide
rule	she	space	subject	think	valley	wife
ruler	shine	speak	substance	this	value	wild
run	shoe	special	succeed	though	various	will
rush	shoot	speed	such	thought	very	win
safe	shore	spend	sudden	through	view	wind
sail	short	spirit	suffer	throw	village	window
same	should	spite	suggest	thus	visit	wing
save	shoulder	spot	summer	time	voice	winter
saw	show	spread	sun	to	vote	wise
say	side	spring	supply	today	wait	wish
scale	sight	square	support	together	walk	with
scarce	sign	stage	suppose	too	wall	within
scene	silence	stand	sure	top	want	without
school	silver	standard	surface	total	war	woman
science	simple	start	surprise	touch	warn	wonder
sea	since	state	surround	toward/ s	waste	wood
season	single	station	sweet	town	watch	word
seat	sir	stay	system	trade	water	work
second	sister	steal	table	train	wave	world
secret	sit	steel	take	travel	way	worse
secretary	situation	step	talk	tree	we	worth
see	size	stick	taste	trouble	weak	would
seem	skill	still	teach	trust	wear	write
seize	sky	stock	tear	truth	week	wrong
sell	sleep	stone	tell	try	welcome	year
send	slight	stop	term	turn	well	yellow
sense	slow	store	terrible	type	west	yes
separate	small	storm	test	under		yesterday
serious	smile	story	than	understan	what	yet
serve	so	straight	that	d	when	you
set	society	strange	the	union	where	young
settle	soft	stream	their	unite	whether	
several	soil	street	them	university	which	
shadow	some	strength	then	unless	while	
shake	son	stretch	there	until	white	

The second 1000 words of the General Service List

abroad	advertise	alike	annoy	arrange	attend
absence	advice	alive	anxiety	arrest	attract
absolutely	afraid	aloud	apart	arrow	audience
accident	afternoon	altogether	apologize	artificial	aunt
accuse	agent	ambition	applaud	ash	autumn
accustom	agriculture	amongst	apple	ashamed	avenue
ache	ahead	amuse	approve	aside	avoid
admire	aim	anger	arch	asleep	awake
adventure	aeroplane	angle	argue	astonish	awkward

axe	breakfast	civilize	creep	ditch	far
baby	breath	clay	crime	dive	farther
bag	bribe	clerk	critic	dollar	fat
baggage	brick	clever	crop	donkey	fate
bake	broad	cliff	crown	dot	fault
balance	brown	climb	cruel	dozen	feast
band	brush	cloth	crush	drag	feather
barber	bucket	cloud	cultivate	drawer	female
bare	bunch	coal	cup	drink	fence
bargain	bundle	coarse	cupboard	drown	fever
barrel	burst	coat	cure	drum	fierce
basin	bury	coffee	curious	duck	fight
basket	bus	coin	curl	dull	film
bath	bush	collar	curse	dust	firm
bay	busy	collect	curtain	duty	flag
beak	butter	colony	curve	eager	flame
beam	button	comb	cushion	earn	flash
bean	cage	combine	damage	earnest	flat
beard	cake	comfort	damp	ease	flavour
beast	calculate	commerce	deaf	educate	flesh
beat	calm	committee	dear	efficient	float
beg	camera	companion	debt	egg	flood
behave	camp	compete	decay	elastic	flour
bell	canal	complain	deceive	elder	flow
belt	cap	complicated	decrease	elect	fly
bend	cape	compose	deed	electricity	fold
berry	captain	confess	deer	elephant	fond
bicycle	card	confidence	defeat	empire	fool
bind	carriage	confuse	defend	empty	foot
birth	cart	congratulate	delay	enclose	forbid
bite	castle	conquer	delicate	encourage	forest
bitter	cat	conscience	delight	engine	forgive
blade	cattle	conscious	department	entertain	fork
blame	caution	convenience	depend	envelope	formal
bless	cave	conversation	descend	envy	frame
blind	cent	cook	deserve	especial	freeze
block	century	cool	desk	essence	frequent
boast	ceremony	copper	despair	exact	fresh
boil	chain	copy	devil	examination	fright
bold	chair	cork	diamond	excess	fruit
bone	chalk	corn	dictionary	excite	fry
border	charm	correct	dig	excuse	fun
borrow	cheap	cottage	dinner	explode	funeral
bottle	cheat	cotton	dip	explore	fur
bottom	check	cough	dirt	extra	furnish
bound	cheer	courage	disappoint	extraordinary	gallon
boundary	cheese	cousin	discipline	extreme	gap
bow	cheque	cow	discuss	fade	garage
bowl	chest	coward	disgust	faint	gas
brain	chicken	crack	dish	false	gay
brass	chimney	crash	dismiss	fan	generous
brave	Christmas	cream	disturb	fancy	girl

glory	hurry	ladder	mercy	pack	pot
goat	hurt	lake	merry	pad	pour
govern	hut	lamp	message	pair	powder
grace	ice	latter	mild	pale	practice
gradual	ideal	laugh	milk	pan	praise
grain	idle	lazy	mill	parcel	pray
grammar	ill	leaf	mis-	pardon	preach
grand	imagine	lean	miserable	parent	precious
grass	imitate	leather	mistake	park	prefer
grateful	immediate	leg	mix	particular	prejudice
grave	immense	lend	model	passage	president
grease	important	lessen	moderate	paste	pretend
greed	improve	lesson	modest	path	pride
greet	indoors	liberty	monkey	patient	priest
grey	industry	lid	moon	patriotic	prison
grind	inform	limb	moral	pattern	prize
guess	ink	lip	motion	pause	probable
guest	in-law	liquid	mouse	paw	procession
guide	inn	list	mud	pearl	profession
guilty	inquire	literature	multiply	peculiar	profit
gun	insect	load	murder	pen	prompt
hair	inside	loaf	mystery	pencil	pronounce
hammer	instant	loan	nail	penny	property
handkerchief	instrument	lock	neat	per	proud
harbour	insult	lodging	needle	perform	pump
harm	insure	log	neglect	persuade	punctual
harvest	interfere	lonely	nephew	pet	punish
haste	international	loose	nest	photograph	pupil
hat	interrupt	lord	net	pick	pure
hate	invent	loud	nice	pig	purple
hay	invite	loyal	niece	pigeon	push
heal	inward/s	luck	noble	pile	puzzle
heap	island	lump	noise	pin	qualify
heart	jaw	lunch	nonsense	pinch	quantity
height	jealous	lung	noon	pink	quarrel
hesitate	jewel	mad	nose	pint	quart
hinder	joint	mail	noun	pipe	queen
hire	joke	male	nuisance	pity	question
hit	journey	manage	nurse	plain	quick
hole	joy	manufacture	nut	plaster	rabbit
holiday	juice	map	oar	plate	radio
hollow	jump	master	obey	plenty	rail
holy	key	mat	ocean	plough	rain
honest	kick	match	offend	plural	rake
hook	king	meal	omit	pocket	rapid
horizon	kiss	meanwhile	onto	poet	rare
hospital	kitchen	meat	opposite	poison	rat
host	knee	mechanic	orange	police	raw
hotel	kneel	medicine	organ	polish	ray
humble	knife	melt	origin	polite	razor
hunger	knock	mend	ornament	pool	recommend
hunt	knot	merchant	overcome	postpone	refer

reflect	sake	slip	stuff	tin	vowel
refresh	salary	slope	stupid	tip	voyage
regret	sale	smell	suck	tire	wage/s
rejoice	salt	smoke	sugar	title	waist
relieve	sample	smooth	suit	tobacco	wake
remedy	sand	snake	supper	toe	wander
remind	satisfy	snow	suspect	tomorrow	warm
rent	sauce	soap	swallow	ton	wash
repair	saucer	socks	swear	tongue	wax
replace	scatter	soldier	sweat	tonight	wealth
reproduce	scent	solemn	sweep	tool	weapon
republic	scissors	solid	swell	tooth	weather
reputation	scold	solve	swim	tough	weave
request	scorn	sore	swing	tour	weed
rescue	scrape	sorry	sword	towel	weigh
reserve	scratch	soul	sympathy	tower	wet
resign	screen	soup	tail	toy	wheat
resist	screw	sour	tailor	track	wheel
responsible	search	sow	tall	translate	whip
restaurant	seed	spade	tame	trap	whisper
retire	seldom	spare	tap	tray	whistle
revenge	sentence	spell	tax	treasure	whole
review	severe	spill	taxi	treat	wicked
reward	sew	spin	tea	tremble	widow
ribbon	shade	spit	telegraph	trial	wine
rice	shallow	splendid	telephone	tribe	wipe
rid	shame	split	temper	trick	wire
ripe	sharp	spoil	temperature	trip	witness
risk	sheep	spoon	temple	true	wool
rival	sheet	sport	tempt	trunk	worm
roar	shelf	staff	tend	tube	worry
roast	shell	stain	tender	tune	worship
rob	shelter	stairs	tent	twist	wound
rock	shield	stamp	thank/s	ugly	wrap
rod	shilling	star	theatre	umbrella	wreck
roof	ship	steady	thick	uncle	wrist
root	shirt	steam	thief	unit	yard
rope	shock	steep	thin	unity	yield
rot	shop	steer	thirst	universe	zero
row	shout	stem	thorn	upper	
royal	shower	stiff	thorough	upright	
rub	shut	sting	thread	upset	
rubber	sick	stir	threaten	upwards	
rubbish	signal	stockings	throat	urge	
rude	silk	stomach	thumb	vain	
rug	sincere	stove	thunder	veil	
ruin	sing	strap	ticket	verb	
rust	sink	straw	tide	verse	
sacred	skin	strict	tidy	vessel	
sacrifice	skirt	string	tie	victory	
sad	slave	strip	tight	violent	
saddle	slide	stripe	till	virtue	

APPENDIX III

List of function words ignored in the analysis

1. A	39. Between	77. Except	115. Him
2. About	40. Beyond	78. Few	116. Himself
3. Above	41. Billion	79. Fewer	117. His
4. Across	42. Billionth	80. Fifteen	118. Home
5. After	43. Both	81. Fifteenth	119. How
6. Against	44. But	82. Fifth	120. However
7. Albeit	45. By	83. Fiftieth	121. Hundred
8. All	46. Can	84. Fifty	122. Hundredth
9. Along	47. Can't	85. First	123. I
10. Although	48. Certain	86. Five	124. I'd
11. Am	49. Could	87. For	125. I'll
12. Among	50. Couldn't	88. Forth	126. I'm
13. Amongst	51. Despite	89. Fortieth	127. I've
14. An	52. Did	90. Forty	128. If
15. And	53. Didn't	91. Four	129. In
16. Another	54. Do	92. Fourteen	130. Into
17. Any	55. Does	93. Fourteenth	131. Is
18. Anybody	56. Doesn't	94. Fourth	132. Isn't
19. Anyone	57. Doing	95. From	133. It
20. Anything	58. Don't	96. Get	134. It's
21. Are	59. Done	97. Gets	135. Its
22. Aren't	60. Down	98. Getting	136. Itself
23. Around	61. During	99. Got	137. Last
24. As	62. Each	100. Had	138. Less
25. Aside	63. Eight	101. Hadn't	139. Like
26. At	64. Eighteen	102. Half	140. Many
27. Away	65. Eighteenth	103. Has	141. May
28. Back	66. Eighth	104. Hasn't	142. Me
29. Be	67. Eightieth	105. Have	143. Might
30. Because	68. Eighty	106. Haven't	144. Million
31. Been	69. Either	107. Having	145. Millionth
32. Before	70. Eleven	108. He	146. Mine
33. Behind	71. Eleventh	109. He'd	147. Minus
34. Being	72. Enough	110. He'll	148. More
35. Below	73. Every	111. He's	149. Most
36. Beneath	74. Everybody	112. Her	150. Much
37. Beside	75. Everyone	113. Hers	151. Must
38. Besides	76. Everything	114. Herself	152. Mustn't

153. My	197. Seventy	241. This	285. Which
154. Myself	198. Several	242. Those	286. While
155. Near	199. Shall	243. Though	287. Whither
156. Neither	200. Shan't	244. Thousand	288. Who
157. Next	201. She	245. Thousandth	289. Whoever
158. Nine	202. She'd	246. Three	290. Whom
159. Nineteen	203. She'll	247. Thrice	291. Whose
160. Nineteenth	204. She's	248. Through	292. Why
161. Ninetieth	205. Should	249. Throughout	293. Will
162. Ninety	206. Shouldn't	250. Till	294. With
163. Ninth	207. Since	251. To	295. Within
164. No	208. Six	252. Towards	296. Without
165. Nobody	209. Sixteen	253. Twelfth	297. Won't
166. None	210. Sixteenth	254. Twelve	298. Would
167. Noone	211. Sixth	255. Twentieth	299. Wouldn't
168. Nor	212. Sixtieth	256. Twenty	300. You
169. Not	213. Sixty	257. Twice	301. You'd
170. Nothing	214. Some	258. Two	302. You'll
171. Notwithstanding	215. Somebody	259. Under	303. You're
172. Of	216. Someone	260. Underneath	304. You've
173. Off	217. Something	261. Unless	305. Your
174. On	218. Still	262. Unlike	306. Yours
175. Once	219. Such	263. Until	307. Yourself
176. One	220. Ten	264. Up	308. Yourselves
177. Onto	221. Tenth	265. Upon	309. Zero
178. Or	222. Than	266. Us	
179. Other	223. That	267. Via	
180. Others	224. That's	268. Was	
181. Ought	225. The	269. Wasn't	
182. Oughtn't	226. Their	270. We	
183. Our	227. Theirs	271. We'd	
184. Ours	228. Them	272. We'll	
185. Ourselves	229. Themselves	273. We're	
186. Out	230. There	274. We've	
187. Over	231. These	275. Were	
188. Per	232. They	276. Weren't	
189. Plus	233. They'd	277. What	
190. Round	234. They'll	278. Whatever	
191. Second	235. They're	279. When	
192. Seven	236. Third	280. Whence	
193. Seventeen	237. Thirteen	281. Whenever	
194. Seventeenth	238. Thirteenth	282. Where	
195. Seventh	239. Thirtieth	283. Whereas	
196. Seventieth	240. Thirty	284. Wherever	

APPENDIX IV

List of words in the Stop-list

a	abner	adair	adverb	ahmedabad	alameda
aa	aboud	adam	aea	ahmeds	alamein
aaa	about	adamafo	aegean	ahsan	alamos
aachen	above	adams	aenarion	aiadmk	alan
aaron	abra	adamson	aeneas	aib	alandi
ababa	abraham	adarsh	aerdenhout	aida	alasdair
abatuno	abrahams	adc	aeroflot	aidan	alaska
abba	abrams	adcock	aes	aidid	alastair
abbado	abrol	addams	aeschylus	aig	alba
abbas	absalom	addanki	aethelbald	aijaz	albadou
abberley	abu	addankidibba	affaire	aiken	alban
abbeydale	abul	addington	afgan	aikin	albans
abbott	abyssinia	addis	afganistan	aileen	albany
abbs	abyssinian	addiscombe	afhq	ainsley	albeck
abby	abyssinians	addison	afolabi	ainslie	albeit
abd	acade	ade	afp	ainsworth	albert
abdallah	academicians	adeane	africa	aintree	alberta
abdel	academie	adelaide	afrose	airdrie	alberto
abdelhak	acapulco	adele	after	aire	albi
abdelhamid	acas	aden	afzal	aires	albie
abdelkrim	acc	adenauer	aga	airtours	albinski
abdessamad	accardo	aderholds	agadir	aissa	albion
abdul	accra	adida	against	aitc	albrecht
abdulkarem	accrington	adidas	agamemnon	aitchison	albrights
abdulkerim	acemi	adie	agamkuan	aitken	albudi
abdullah	acet	adjective	agassi	aix	albuquerque
abe	achaeans	adjunct	agatha	aj	alcatel
abel	acharya	adkin	agence	ajax	alcoa
abelard	achebe	adlai	agg	ajdabiya	alcock
abell	acheson	adler	aggie	ajit	alcuin
aberconway	achille	adn	agincourt	ajmer	aldeburgh
abercrombie	achim	adnan	agnelli	akand	alden
aberdeen	acioli	adolf	agnes	akbar	alderley
aberdeenshire	acker	adolfo	agnew	akers	aldermaston
abergavenny	ackford	adolph	agonistes	akhil	aldershot
aberley	acking	adolphe	agr	akinsha	alderson
aberystwyth	acklam	adorno	agra	akram	aldgate
abi	ackner	adria	agrippa	al	aldington
abie	ackroyd	adrian	ahilar	alabama	aldo
abigail	acp	adriatic	ahlna	aladdin	aldous
abimael	across	adrienne	ahmad	alagiriswami	aldrich
abingdon	acton	adsetts	ahmann	alagoas	aldrige
abkhazia	ada	adt	ahmed	alain	aldrington

aldus	alksnis	altman	amis	andover	ankara
aldwych	all	alton	amman	andr	ann
alec	alla	altrincham	among	andre	anna
alejandro	allah	alun	amongst	andrea	annabel
aleksandr	allan	alusik	amor	andreas	annabelle
alem	allang	alva	amory	andrei	annadale
alemany	allegretti	alvarez	amos	andreotti	annadorai
alessandro	allen	alvey	amost	andres	annamorena
alex	allenby	alvi	ampeg	andrew	annan
alexander	allende	alvie	amrita	andrewes	annanagar
alexandra	allensbach	alvin	amritraj	andrews	annapolis
alexandre	allerdale	alwyn	amritsar	andrus	annas
alexandria	allerton	alyssia	amsterdam	andy	annaud
alexei	allison	alzheimer	amstrad	anfield	anne
alexis	allitt	am	amy	ang	anneka
alf	allman	amadeus	an	angalo	annes
alfa	alloa	amado	ana	angazi	annesley
alfie	allsop	amanda	anabelle	ange	annette
alfonso	allsopp	amaranth	anacleto	angela	annie
alford	allwright	amarnath	anaheim	angeles	anniston
alfred	alma	amazon	anand	angeli	anno
alfredo	almeida	amazonia	anandan	angelica	another
algarve	alnwick	ambedkar	anantapur	angelico	ansaldo
algeria	alomar	ambie	anastasia	angelina	ansell
algerian	along	ambleside	anatole	angelini	anselm
algernon	aloysius	ambo	anatolia	angell	anshuman
algiers	alp	ambrose	anatoly	angelo	ansi
algoma	alpharetta	amdahl	anc	angevin	ansley
algy	alphonse	amelia	anchorsholme	angharad	anson
ali	alresford	amelie	and	angie	ansonia
alia	alsace	amer	andalusia	anglasi	anstey
alianor	alsop	america	andalusian	anglesey	antall
alice	alsthom	american	ande	anglia	anthea
alicia	alston	americas	anderlini	anglian	anthony
alida	alt	americus	anders	angolese	anti
aligarh	alta	amersfoort	andersen	anguilla	antibes
alija	altaf	amersham	anderson	angus	antigua
alina	altdorf	amery	andersonstow	angy	antilles
aline	althaus	ames	n	anielli	antioch
alipur	althorp	amiens	andersson	anil	antiqua
alison	although	amies	anderton	anis	antoine
alistair	althusser	amiga	andhra	anita	antoinette
alix	altick	amin	andorra	anjou	anton

antone	aquina	arf	arnhem	ashe	aswan
antonia	aquinas	argentines	arnie	ashefelter	at
antonio	aquino	arger	arno	ashenden	ata
antony	aquitaine	argonne	arnold	asher	atal
antrim	ar	argos	around	ashfield	atari
antropov	ara	arguello	arouny	ashford	atas
antulay	arab	argus	arp	ashington	athelstan
antwerp	arabella	argyle	arrack	ashley	athena
anup	arabic	argyll	arran	ashman	athene
anwar	arac	ari	arrhythmia	ashmolean	athenee
answers	arafat	ariadne	arrington	ashmore	athenian
any	aragon	ariane	arroll	ashok	athenians
anya	araminta	ariel	arrowsmith	ashoka	athens
anybody	aramoana	aries	arsenal	ashton	atherton
anyone	aran	arif	arsenio	ashworth	athletico
anything	arata	aright	arshinkoff	asia	atholl
anzio	arauco	aristarchus	artai	asian	atkin
aoun	arazi	aristide	arte	aside	atkins
apana	arblaster	aristo	artemis	asilomar	atkinson
aparicio	arbroath	aristos	arthian	aslam	atlanta
apb	arbuckle	aristotle	arthur	asm	atlantic
apbhramsa	arbuthnot	ariz	arthurokun	asmara	atleast
apennines	archaos	arizona	artois	aspi	atlee
aphrodite	archibald	arjun	arun	asprey	atrice
apirana	archibaldcraw	arkansas	arundel	asquith	attenborough
apis	ford	arke	arvey	assab	attercliffe
apl	archie	arkwright	arvidas	assad	attica
aplin	archimedes	arle	arwen	assam	attila
aplysia	arcot	arlecdon	arya	assarlindback	attilio
apollinaire	arcy	arles	as	assenmacher	attleborough
apolline	ard	arlington	asa	asshe	atlee
apollo	arden	arlott	asante	assisi	attributive
app	ardennes	armacost	asb	assn	attwood
appa	ardiles	armageddon	asc	ast	atty
appen	ardis	armagh	ascension	astaire	atul
appleby	ardor	armagnac	ascot	asters	atwater
appleton	ardrossan	armand	asda	astley	atwells
appliques	ards	armani	asdic	aston	atwood
april	are	armisteads	ashbourne	astor	aubrey
apsley	aren	armitage	ashburton	astoria	auchinleck
apu	arens	armond	ashby	astra	auckland
aqib	arequipa	armstrong	ashcroft	astrid	auden
aquarius	ares	arn	ashdown	astros	audi

audley	averell	b	bahadur	balearic	bandh
audrey	averns	baa	bahamas	balenciaga	bandi
audubon	avery	baas	bahi	baleshwar	bandish
auerbach	avesco	bab	bahia	balfour	bandra
augsburg	aviemore	baba	bahraich	balgangadhar	bandung
august	avignon	babangida	bahrain	balham	banera
augusta	aviion	babar	bahrein	bali	banerjee
auguste	avila	babasaheb	baie	balinese	banff
augustincourn	aviv	babbage	baig	balkin	banfield
ot	avon	babcock	bajinath	ballabh	bangalore
augustine	avonmouth	babel	baikal	ballantyne	bangemann
augusto	avril	baberton	bailey	ballard	bangkok
augustus	avtar	babette	baileys	ballater	bangla
ault	awamutu	bablake	baillie	ballenger	banglasaheb
aung	away	babulal	bainbridge	ballesteros	bangor
aurae	axel	babur	baines	balletic	bann
aurangabad	axminster	babylon	bair	ballia	banon
aurangzeb	axp	babylonian	baird	balliol	bansi
aurelius	aya	bacall	bajoria	ballycastle	banskothi
aurore	aycliffe	baccalaureate	bajpe	ballyclare	banstead
auschwitz	ayer	bacchus	bajsons	ballymena	bao
auspex	aykroyd	bach	bakatin	ballymoney	baptiste
auspice	aylesbury	bachan	bakersfield	balmoral	batu
austell	ayling	back	bakewell	balmukund	baqr
austen	aylwen	baconsfield	bakht	baloghy	bara
austin	aylwin	badal	bakhtin	balor	baraclough
australasia	ayodhya	badalamenti	bakker	baltics	barak
australia	ayr	baden	baktu	baltimore	barbadian
australian	ayres	bader	baku	baltimorean	barbados
austria	ayrshire	badie	bala	balwant	barbara
automaker	ayrton	badshah	balanchine	balzac	barbaracraig
autonomen	ayton	badwar	balbinder	bamber	barbarawalters
autopacific	azad	bae	balbir	bambi	barbarossa
auvergne	azadi	baer	balboa	bamford	barbican
auxerre	azam	bagan	balcombe	bampton	barbie
auxiliary	azamgarh	bagaya	balcon	banbridge	barbitone
ava	azamgarhs	bagby	baldersdale	banbury	barbour
avadhoot	aziz	bagehot	baldev	banco	barbra
avala	azof	bagenal	baldock	bancroft	barcelona
avanigadda	azoff	baggeridge	baldrige	banda	barchester
avant	azore	bagh	baldrige	bandalore	barclay
ave	azt	baghdad	baldry	bandaranaike	barclays
avebury	azur	bagicha	baldwin	bandeira	bardesley

bardolph	bartholomew	batku	beano	bedelia	belgrano
bardot	bartle	batman	bearden	bedford	belgrave
bardsley	bartleman	batra	beardens	bedfordshire	belgravia
barford	bartlett	batt	beardown	bedi	belikov
bari	bartley	battersea	beardsley	bedley	belinda
baringer	bartok	baudelaire	bearsden	bedworth	belize
barker	barton	baudouin	beas	bedwyr	bella
barkley	bartram	bauer	beasant	beeb	bellamy
barlaston	barwick	baum	beasley	beebe	belleville
barleycroft	baryshnikov	baumann	beatle	beechem	bellevue
barling	basanti	bauscher	beatles	beechems	belloc
barlow	baseballight	bavaria	beaton	beeching	bellwood
barnabas	basel	baxley	beatrice	beecroft	belmarsh
barnaby	basher	baxter	beatrice	been	belmont
barnard	bashir	bayer	beattie	beerbohm	beloff
barnardo	basildon	bayern	beatty	beershorn	belorussia
barnardos	basinger	bayezid	beauchamp	beesemyers	below
barnes	basingstoke	bayfield	beaufort	beeston	belsen
barnet	baskerville	bayi	beaumont	beethoven	belthangdy
barnett	baskinrobbins	bayle	beauvoir	before	belur
barney	basle	bayley	beaverbrook	begum	belvedere
barnsley	basquiat	baylor	beaverton	behan	belvidere
barnstaple	basra	bayonne	beavis	behari	belville
barnum	bassanio	bayswater	beazley	behbehani	belvoir
baroda	bassenthwaite	bb	bebb	behind	beman
barr	bassett	bbc	bebington	behner	ben
barra	bassi	bbc1	bec	beijing	benazir
barrat	bassis	bbc2	because	being	benbecula
barratt	bastable	bc	beccaria	beinn	bendle
barrault	basten	bcci	beccles	beirut	beneath
barre	basti	bcg	becher	beith	benedetti
barrett	bastikar	bcp	bechtel	beitiks	benedict
barrie	bastille	bcr	beckenham	bejart	benedicta
barrington	basu	bcrs	becker	bekaa	benesch
barro	bata	bd	becket	bel	benetton
barron	bataganj	be	beckett	belafonte	bengal
barry	batchelder	bea	beckford	belanger	bengali
barrymore	bateman	beachy	beckwith	belchi	benghazi
barstow	bates	beaconsfield	becky	beldon	beni
bart	bateson	beaglehole	bedale	belet	benidorm
barth	bathgate	beale	beddington	belfast	benin
bartha	bathsheba	beame	beddow	belgium	benington
barthes	bathurst	beamer	bede	belgrade	benito

benjamin	bermuda	bethesda	bhattacharjee	bijlmermeer	birla
benn	bern	bethlehem	bhau	biju	birmingham
bennerley	bernadette	bethnal	bhavan	bilap	biro
bennet	bernadine	betjeman	bhavani	bilbao	birr
bennett	bernard	betsy	bhavnagar	billi	birse
bennie	bernardine	bett	bhayya	billie	birt
benny	bernardo	bette	bhengra	billiken	birur
benoit	berndt	betts	bhide	billikens	biscay
benson	berne	betty	bhir	billingham	biscayne
benstede	bernet	between	bhisham	billings	bischoff
bentham	bernhard	beuno	bhojan	billingsgate	bishen
bentley	bernhardt	bev	bhola	billingsley	bishopsgate
benton	bernice	bevan	bhopal	billington	bishopthorpe
bents	bernie	beveridge	bhowmick	billion	bismarck
benvinguts	bernini	beverley	bhumihars	billionth	bispham
benz	bernstein	beverly	bhupinder	billy	bissau
beowulf	beronio	bevin	bhushan	biloxi	bisset
bera	berowne	bewick	bhushayya	biltmore	bissett
berdichev	berra	bewicke	bhutan	bimbo	biswanath
beregovoy	berrellez	bexar	bhutto	bimini	biya
berenice	berrie	bexhill	bhylls	bindley	biz
beresford	bert	bexley	bianco	bing	bizerte
berg	bertelot	bey	biarritz	bingham	bjoerling
bergen	berteros	beyeler	bibb	bingles	bjorn
berger	bertha	beyond	bibby	bingley	bjp
bergerac	bertie	bf	bibhutupur	binks	blaby
bergg	bertoia	bfs	bicc	binley	blackadder
bergman	berton	bgs	bicester	binns	blackbeard
bergonzi	bertorelli	bh	bickenhill	binny	blackburn
berich	bertram	bhabha	bickerstaff	binyon	blackett
berk	bertrand	bhagat	bicknell	biondi	blackfriars
berkeley	berwick	bhagwat	biden	bippus	blackheath
berkhamsted	berwickshire	bhai	bidesh	birchgrey	blackie
berkley	beryl	bhakra	bidwell	birchwood	blackman
berks	bes	bhalee	biennale	birendra	blackmoor
berkshire	beside	bhalindra	bietnar	birgit	blackmore
berkswell	besides	bhanot	biffen	birk	blackpool
berlin	bess	bhanu	bigg	birkbeck	blackstone
berliner	besset	bharat	biggin	birkby	blackwall
berliners	bessey	bharatiya	biggles	birkdale	blackwell
berlioz	bessie	bharattya	biggs	birkenhead	blackwood
berlusconi	beth	bhaskar	biha	birkett	bladesover
bermondsey	bethel	bhaskaran	bihar	birkhall	bladesovers

blaenau	bloomsbury	bognor	boone	boueuse	boyne
blagg	blott	bogota	boonton	bougainville	boz
blaiklock	bloxham	bogside	boothby	bouge	bp
blaine	blufton	boh	boothroyd	boulestin	bpx
blair	blum	bohr	bootle	boulnor	br
blairhall	blumberg	bohrras	borde	boulogne	brabant
blaisdell	blume	bois	bordeau	boulton	brabham
blaise	blumenthal	bojan	bordeaux	bourcier	brac
blakburn	blundell	bokaro	borden	bourdieu	bracewell
blake	blunkett	boldwood	bordon	bourg	brackenwood
blakelock	bly	bolet	borehamwood	bourgogne	brackley
blakey	blyth	boleyn	borg	bourguiba	bracknell
blanca	blythe	bolger	borges	bourn	brad
blanchard	blyton	bolinger	boris	bourne	bradbury
blanche	bm	bolker	borkdale	ournemouth	braddock
blanchflower	bma	bologna	borland	bourton	bradfield
blanco	bmc	bolognese	borlase	boutel	bradford
blandford	bmh	bolshhevik	borneo	bouton	bradley
blanton	bmw	bolshoi	boro	boutros	bradman
blasingame	bnfl	bolsover	borough	bovary	bradnocks
blaxton	bnp	bolsterstone	boroughbridge	bovis	bradshaw
bld	boalt	bolton	borrough	bovril	bradstreet
bldg	boardman	boltongate	borrowdale	bowater	bradwell
ble	bob	bombay	bosanquet	bowden	brady
bleanch	bobbie	bonaparte	bosch	bowditch	braemar
bleasdale	bobie	bonar	bose	bowdoin	braer
blease	bobo	bonard	boskin	bowe	bragad
bleckley	bocce	bondgate	bosnia	bowen	bragg
blenheim	bochco	bonetti	bosphorus	bowes	brahmachari
blenkinsop	boddington	bonham	bossman	bowie	brahms
bletchley	boddy	boniface	bostock	bowles	braille
bligh	boden	bonington	boston	bowness	braintree
blissett	bodganov	bonn	bostron	bowood	braithwaite
blix	bodie	bonnard	boswell	bowring	braj
bloch	bodin	bonner	bosworth	bowthorn	braji
bloemfontein	bodmin	bonnie	both	bowthorpe	brajmohan
blofeld	bodo	bono	botha	bowyer	bramall
blogg	boehmer	bontempo	botham	boyant	brame
bloggs	boeing	booch	bothwell	boyce	bramhall
blois	bogard	boocock	botswana	boyd	bramley
blondel	bogart	booker	bott	boyer	brampton
bloomfield	bogdan	bookman	bottomley	boyes	branagh
bloomingdale	bogdanovich	bookwalter	boucher	boyle	brancepeth

brandeis	brentano	brimmer	brom	bruton	buechele
brandenburg	brentford	brinda	bromborough	brutus	buell
brando	brenton	brindley	bromley	bryan	buenos
brandon	brentwood	brinker	brompton	bryant	buffett
brandt	brereton	brioni	bromsgrove	bryce	buffy
braniff	brescia	brisbane	bromwich	bryn	buganda
brann	brest	bristol	bronson	bryony	bugatti
brannon	bretherick	bristow	bront	bryson	bughes
branson	breton	britain	bronte	bs	bugis
brantum	brett	britannia	bronx	bsa	buick
branwell	bretton	british	brooke	bsb	builth
braque	breuer	briton	brookes	bsdi	bukharin
braques	brevard	britt	brookland	bskyb	bula
brasch	brewster	britta	brooklands	bsp	bulandra
brasenose	brezhnev	brittain	brooklyn	btec	bulawayo
brasilia	brezinski	brittan	brookside	btr	bulgari
brassey	bri	brittany	broome	buaford	bulgaria
brassi	brian	britten	broomfield	bubenik	bulgarian
bratislava	briant	britton	brosnan	buc	bulger
braun	brice	brixham	brotherton	bucclench	bullcrap
braunschweig	brideshead	brixton	brough	buchan	buller
braunspeth	bridgeman	brize	broughton	buchanan	bullinger
braunton	bridgend	brm	broward	buchanans	bulloch
brava	bridget	brno	browne	bucharest	bunbury
brayton	bridgewater	broadbanks	brownlow	buchheister	bunce
brazil	bridgnorth	broadbent	brownson	buckenham	bund
brazilian	bridgwater	broadeast	broxbourne	buckhead	bundesliga
brearley	bridhe	broadgreen	bruce	buckie	bundespost
breasley	bridlington	broadhurst	bruces	buckingham	bundestag
brechin	brien	broadmoor	bruckner	buckinghamsh	bundy
brecht	brierley	broadstairs	brufton	ire	bungay
breckland	brierly	broca	bruges	buckland	bunn
breckon	brierydale	brock	brugge	buckley	bunter
brecon	brigg	brocklebank	brummie	buckmaster	bunty
bregman	briggs	brocklin	brundle	buckthorn	bunyan
brelin	brigham	brod	brunei	bucky	bupa
bremen	brighthouse	broderick	brunel	budapest	burbank
bremner	brightman	brodie	brunner	budd	burckhardt
brenda	brighton	brodsky	bruno	buddha	burdett
brendan	brigid	broeg	brunswick	buddie	burdon
brennan	brigitte	broek	bruntisfield	budgen	burford
brenner	brijesh	broglio	brunton	budgette	burg
brent	brill	brokaw	brussels	budweiser	burge

burgess	buthelezi	cairngorms	cambridgeshir	canterbury	carlyles
burghersh	butlin	cairo	e	cantona	carman
burghley	butlins	caithness	cambs	cantonese	carmarthen
burgoyne	butori	caius	cambuslang	canute	carmel
burgtheater	butterfield	calais	camden	canvey	carmen
buri	buttermere	calatin	camelot	capel	carmichael
burke	butterworth	calcutta	cameron	capern	carmine
burkes	buxar	calder	cameroon	capitol	carn
burkina	buxted	calderwood	cameroun	capm	carnaby
burl	buxton	caldon	camers	capone	carnarvon
burleigh	bv	caldwell	camiel	capps	carnegey
burley	by	caleb	camilla	capri	carnegie
burlington	byelarus	caledor	camille	capriati	caro
burma	byelorussia	calgary	camilo	capricorn	carole
burmah	byer	caliban	cammell	cara	carolina
burmese	bygraves	caliber	campagnoli	caracas	caroline
burnden	byrd	calibra	campbell	carberry	carolyn
burne	byrne	calicut	campbells	carbones	carolyne
burnell	byron	california	campden	cardiff	caron
burnes	byzantium	caligula	campese	cardoso	carovei
burneside	bzw	callaghan	campion	carella	carpathians
burnet	c	callahan	campo	carew	carpentier
burnett	caa	callan	campobello	carey	carr
burney	caborn	callanish	campos	cargill	carradine
burnham	cabot	callas	campra	caribbean	carraway
burnley	cabra	callejas	camra	caricom	carre
burnside	cadbury	callen	camus	carl	carreon
burrell	cade	callum	can	carla	carrera
burroughs	cadfael	calne	canaan	carleton	carreras
bursa	cadillac	calpers	canacona	carlett	carri
burt	cadiz	caltech	canada	carlin	carrick
burton	cadogan	calude	canadian	carling	carrickfergus
burtons	caen	calum	canaletto	carlisle	carrie
burun	caernarfon	calumet	canberra	carlo	carril
burundi	caernarvon	calvary	candice	carlos	carrington
busa	caerphilly	calvert	candida	carlotta	carroll
busacher	caesar	calvin	candide	carlow	carron
busby	cagliari	calwell	candiotti	carlsberg	carrow
busch	cagney	camberley	cann	carlson	carruthers
bushell	cahill	camberwell	cannes	carlton	carshalton
bushey	cain	camborne	cannock	carluke	carson
but	caine	cambridge	canoga	carly	carthage
bute	cairngorm		canseco	carlyle	cartier

cartland	cathay	ceausescu	chalker	charanjit	chavara
cartwright	catherine	ceci	challan	chardon	chavigny
carty	cathie	cecil	challapalli	charente	chaytor
caruso	cathy	cecilia	challis	charing	chd
carvalho	catlins	cecily	challoner	charlayne	che
carville	cato	cedergren	chalmers	charlemagne	cheadle
cary	catriona	cedille	chaman	charlene	cheam
cas	catterick	cedric	chamier	charleroi	cheapside
casablanca	caucasus	cedvet	chamonix	charles	checkland
casals	cauffman	cegb	chamorro	charleston	cheddleton
casassa	caulfield	ceil	champaign	charlestown	chedworth
casella	cauvery	celebes	chan	charlesworth	cheele
casey	cavaliere	celeste	chanana	charley	cheesefoot
cashin	cavan	celia	chancy	charlie	cheetham
caspar	cavanagh	celsius	chand	charlotte	chekhov
caspian	cavell	cemal	chander	charlton	chelmsford
cass	cavendish	cencede	chandi	charlwood	chelsea
cassandra	cawasji	centredale	chandigarh	charman	cheltenham
cassell	cawley	cern	chandon	charmian	chen
cassidy	cawthorne	certain	chandos	charsky	cheney
cassie	caxton	ceruty	chandra	charterhouse	cheng
casson	caygill	cerv	chandrajeet	chartres	chennakesava
castalia	cayman	cervetto	chandrsekhar	chas	chepstow
castell	cazeneuve	cesar	chandrashekar	chatfield	chequers
castelnau	cb	cesarewitch	chandrashekha	chatham	cher
castile	cbc	cesario	r	chatichai	cherbourg
castleford	cbi	cessna	chandrawati	chatman	chernin
castlegate	cbs	cestrians	chanel	chatmans	chernobyl
castlereagh	cc	ceta	chang	chatrinaka	chertsey
castleton	cca	cf	changez	chatsworth	cherwell
castrey	ccc	cfa	channing	chatterjee	cheryl
castro	ccg	cfe	channon	chatterley	chesapeake
caswell	cck	cgi	chantal	chatterleys	chesarynth
catalan	ccp	cgt	chantilly	chatterton	chesford
catalonia	ccpr	chad	chaovalit	chatto	cheshire
catalonian	ccw	chadarghat	chapelle	chaturvedi	chesney
catalunya	cd	chadwick	chappelles	chatwin	chester
catania	cdc	chagall	chapeltown	chaucer	chesterfield
caterham	cdp	chagger	chaplin	chaud	chesterton
caterina	cdr	chaise	chapman	chaudron	cheung
catesby	cdu	chakradharpur	chappell	chautauqua	cheviot
catford	ce	chaleur	chapple	chavalit	chevrolet
cath	ceauscescu	chalfont	charan	chavan	cheyne

cheyt	chok	ciba	clara	clermont	cobb
chhaganlal	choksi	cicely	clare	cletter	cobbett
chi	chola	cicero	claremont	cleveland	cobham
chiam	chomsky	ciciulla	clarence	cleveleys	cobi
chiang	chong	cics	clarendon	cley	cobol
chicago	chopin	cid	clarice	cliffe	coburg
chichester	chopra	cido	claridge	clifford	coburn
chico	chorley	ciel	clarissa	clift	coca
chieh	chorlton	ciga	clark	clifton	cochran
chigwell	chou	cilla	clarke	cliftonville	cochrane
chikmagalur	choudhary	cimoli	clarks	clint	cockburn
chikmaglur	choudhury	cincinnati	clarkson	clinton	cockcroft
childebert	chow	cinderella	clarksons	clintons	cocker
childs	chowgule	cindy	clary	clio	cockermouth
chiles	chp	cinemactor	clasper	cliss	coco
chillington	chris	cingle	claud	clive	cocteau
chillingworth	chrissie	cinzia	claude	clo	coddington
chilperic	chrissy	cio	claudel	clore	cody
chiltern	christ	cipriani	claudette	clough	coe
chilterns	christa	cir	claudia	clouston	coetzee
chilton	christabel	cirencester	claudine	clovis	cohen
chiltoni	christchurch	ciskei	claudio	clowes	cohn
china	christensen	cissie	claudius	cluny	cohse
chinese	christi	cissy	clause	clutha	coimbatore
ching	christiansen	cialia	claydon	clwyd	coker
chingford	christie	citibank	clayton	clyde	cola
chinmoy	christies	citicorp	clearwater	clydebank	colaba
chinnaswamy	christina	cito	cleary	clydesdale	colah
chinnor	christine	citrine	cleator	clymer	colavito
chintakolla	christoff	citroen	cleese	cmc	colbert
chintaman	christophe	cityarts	cleeve	cmea	colchester
chinua	christopher	ciulla	clegg	cmhts	coldfield
chippenham	christy	cj	cleireacain	cmnd	colditz
chirac	chromspun	cl	clelia	cmos	coldstream
chiradeep	chrysler	clacton	clem	cnaa	cole
chiselhurst	chu	cladonjoye	clemens	cnd	colebrooke
chisholm	chubb	clair	clemente	cnn	coleby
chispas	chung	claire	clementina	cnut	coleman
chiswick	chunovic	clairmont	clements	coad	colenso
chiuchow	churchill	clairrol	clemson	coady	coleraine
chloe	churchills	clancy	cleo	coahr	coleridge
chlothar	cia	clapham	cleopatra	coatbridge	coles
cho	ciaran	clapton	clerkenwell	coates	colette

coley	compagnie	consonant	corell	cothope	cox
colgate	compaq	constance	corfu	coton	coyle
colin	compton	constantine	corinne	cotswold	cp
collett	compuadd	constantinople	corinth	cotswolds	cpez
colley	comte	constantinos	corkerton	cott	cpd
collge	comus	constanza	corky	cottam	cpl
collighan	conan	constanze	cormac	cottbus	cpp
collimore	conant	conte	cormack	cottee	cpr
collin	concacaf	contra	corman	cotten	cpre
colling	conchis	conway	corneilus	cottesmore	cprs
collinge	concorde	conwy	corneliu	cottle	cprw
collingridge	condon	cooke	cornelius	cottrell	cps
collingwood	condotta	cookson	cornell	cottrill	cpsu
collins	coney	cookstown	cornhill	cotty	cpv
collinson	cong	cooley	corning	coulby	cr
collis	congleton	coolidge	cornish	could	crabb
collo	congo	coombe	cornmill	couldn	crabbe
collor	congolese	coombes	cornwall	coulson	crabtree
colm	conigre	coombs	cornwallis	coulter	cracow
colman	coniston	cooney	cornwell	coundon	craddock
colmer	conjunction	coonoor	coronado	countian	craig
colne	conlon	coope	corrado	countin	craigavon
colombo	connachan	cooper	corrie	coupar	craigwell
colorado	connall	cooperman	corrigan	couper	craik
colquhoun	connaught	coopers	corsham	courbet	cramer
colquitt	connecticut	coote	corsica	courmont	cramlington
colrain	connell	copec	corsican	courtauld	crampton
colson	connelly	copeland	cortaulds	courtaulds	crandall
coltrane	conner	copenhagen	cortes	courtenay	cranfield
columba	conners	copernicus	cortina	courtney	crangle
columbia	connery	copley	cory	couscous	cranko
columbus	connie	coppell	corydora	cousteau	cranmer
colville	connolly	copperfield	cosby	coutts	cranston
colvin	connon	copperwheat	cose	covent	craon
colwyn	connor	cora	cosmo	coventry	crappr
comack	connors	corazon	cossiga	covington	crase
comaneci	conoco	corbett	costa	cowan	crasharts
combe	conor	corby	costain	cowes	crawford
combes	conrad	corcoran	costello	cowie	crawley
comdex	conran	cordeiro	costner	cowley	craxi
comecon	conroy	cordelia	cosworth	cowpe	cray
como	consall	corden	cote	cowper	crc
comoro	consett	cordle	cotherstone	cowpes	cre

creb	croydon	cunard	daiena	daniel	das
credal	crozier	cunliffe	daim	daniele	dascenzo
creely	crp	cunningham	daimler	danielle	dasgupta
creggan	crufts	cuomo	dain	daniels	dashwood
cremer	cruickshank	curphey	daine	dann	dataquest
cresp	crumlin	curran	daiwa	dannehower	daunbey
crespi	crumlish	currie	dakar	dannevirke	dave
cresson	crump	curtis	dakota	dannii	davenport
cresswell	crumpacker	curwen	dalai	danny	daventry
crestwood	crumwallis	curzon	dalal	danseur	davey
crete	crusoe	cusack	dalby	dante	david
crevecoeur	crutsinger	cushman	daley	danube	davide
crewe	cruz	cusick	dalgety	danvers	dauids
crichton	cruzeiros	cuthbert	dalglish	danzig	davidson
crillon	cryer	cutler	dalglish	danziger	davie
crilly	cs	cutty	dali	daphne	davies
crimea	csa	cuvier	dalit	dar	davis
crimean	csce	cuzco	dalkeith	dara	davison
cripps	csm	cw	dallas	darby	davy
crispin	cso	cy	dalles	darcy	dawes
criss	csu	cynthia	dalmatia	dardanelles	dawkins
crissy	ctc	cyprian	dalrymple	daren	dawnay
cristiani	ctl	cyrano	dalston	daridranaraya	daws
crocker	cuba	cyril	dalton	n	dawson
crofton	cuddalore	cyrus	daltrey	darius	day
crohn	cudmore	cytr	daly	darkfall	dayal
cromer	cuellar	d	dalyell	darlene	db2
crompton	culbertson	da	dalzell	darley	dbms
cromwell	culham	daak	dalziel	darlington	dce
cronin	culkin	dabeerpura	daman	darman	dcf
crore	cullam	dacourt	damascus	darnell	dcm
crores	cullen	dacre	damian	darnley	dcs
crosby	culley	dadar	damien	darrell	ddt
crosland	culloden	dadda	damocles	darren	de
crosser	culpeper	dade	damon	darrow	dea
crossick	culpepper	daedalus	dan	darshan	deadwood
crossley	culver	daf	dana	dartford	deakin
crossman	cumberland	dagenham	danapur	dartington	deana
crowe	cumbernauld	dagmar	danbury	dartmoor	deane
crowley	cumbria	dahl	danby	dartmouth	deanna
crowth	cumming	dahlburg	dandavate	darwin	dearborn
crowther	cummings	dahlink	dangerfield	darya	dearg
croxley	cummins	dai	danica	daryl	dearham

dearlove	delingpole	derrida	devore	dicky	dinkey
dearne	delius	derring	devraj	did	dinkins
dearness	delle	derry	devraux	didcot	dino
deas	delmatoff	dersingham	dewar	diderot	dinshaw
deauville	delo	derwent	dewey	didi	dion
debbie	deloitte	derwentwater	dewi	didier	dionne
debenham	deloris	deryck	dewitt	didn	dior
debenhams	delors	des	dewsbury	diefenbaker	diplock
deborah	delphi	desai	dexter	diego	dirham
debre	delray	descartes	dey	dieppe	dirk
debussy	demi	desdemona	dg	dierdriu	dirksen
decatur	demonstrative	dese	dh	dieskau	dis
decca	dempsey	desh	dha	dietrich	disa
december	dempster	deshi	dhabi	digby	disimone
decicco	denbigh	deshmukh	dhaka	diggory	disney
declan	dene	desiree	dhan	diggstown	disneyland
dee	deng	desmond	dhanbad	digha	disraeli
deeley	denham	desoto	dhanuks	dighaghat	diss
deen	denholm	despard	dharia	digvijay	dist
deenbandhu	denis	despenser	dharmapuri	dijon	ditmar
deere	denise	despite	dharna	dili	ditton
deeside	denison	dessie	dhawan	dilip	diu
definite	denning	determiner	dhempe	dillon	divac
defoe	dennis	detlef	dhindsa	dillons	divarco
defreitas	dennison	detroit	dhirendra	dilum	divonne
defries	denny	deutsch	dhoties	dilworth	diwa
degas	densmore	deutsche	dhss	dilwyn	diwakar
degnan	dentdale	deutschemark	dhyan	dimaggio	diwali
dehat	denton	deutschland	di	dimbleby	dixie
deighton	denver	dev	diana	dimitri	dixon
deirdre	denverite	devaraj	diane	dimity	dixons
dekalb	denys	devaux	dianne	dimmesdale	djakarta
dekhamhare	denzil	devenish	diaoyutai	din	djibouti
del	deodar	devereux	diaz	dinah	dk
delacroix	deoria	devi	dibben	dinar	dlp
delahanty	depardieu	devine	dick	dinard	dlr
delamere	deptford	devito	dickens	dinda	dme
delaney	dequindre	devizes	dickey	dindigul	dmitry
delaware	der	devlin	dickie	dinenage	do
delegate	derbyshire	devon	dickins	dinesh	doak
delgard	derek	devonian	dickinson	dingwall	dobbin
delhi	dermot	devonport	dickon	diniz	dobbins
delia	derosa	devonshire	dickson	dinizulu	dobbs

dobie	donne	doubleday	drosophila	duma	duque
dobson	donnell	doug	drottningholm	dumas	duran
docherty	donnelly	dougal	drs	dumbarton	durance
documenta	donny	dougan	drucker	dumfries	durand
dod	donoghue	dougie	drummond	dumfriesshire	durante
dodd	donovan	douglas	drumskin	dumont	duras
dodds	donson	douglass	drury	dunbar	urban
dodson	doody	doulton	dryden	dunblane	durcar
dodworth	doohan	dounreay	drysdale	duncan	duren
does	dooley	dover	ds	dunclon	durgan
doesn	doolittle	dovercourt	dsp	dundalk	durgans
doh	doone	dow	dss	dundas	durham
doherty	doppler	dowd	dti	dundee	durie
doi	dora	dowie	du	dundonald	durieux
doing	dorado	dowling	duane	dunedin	during
dolan	doramin	down	duart	dunfermline	durkheim
dolaucothi	doran	downes	dubai	dungannon	durkin
dollar	dorca	downey	dubcek	dungarvan	durning
doloi	dorcas	downpatrick	dublin	dunham	duroc
dolores	dorchester	dowsett	dubois	dunhill	durocher
dom	dordogne	dowson	dubovskoi	duniya	durrani
domenico	dore	doxiadis	dubrovnik	dunkel	durrant
domesday	doreen	doy	duckens	dunkeld	durwood
domingo	dorenzo	doyle	duckworth	dunkirk	dury
dominic	dorf	dp	duckworths	dunkley	dusable
dominica	dorian	dpr	dudley	dunlop	dussa
dominique	dorigo	drabble	dudulgaon	dunn	dusseldorf
don	doris	drachenfel	duesseldorf	dunna	dustin
dona	dorking	dracula	dufay	dunne	dusu
donaghy	dorling	drayton	duff	dunning	dutton
donahue	dorma	dresbach	duffield	dunnit	duval
donal	dorman	dresbachs	duffy	dunoon	duvall
donald	dorn	dresden	duggal	dunrossil	duxford
donaldson	dornan	drexel	duhagon	dunstable	dw
donato	dorniers	drexler	dui	dunstan	dwayne
doncaster	doro	dreyfus	duisburg	dunstaple	dwight
done	dorothea	dri	dukakis	dunwich	dworkin
donegal	dorothy	drinkel	dulcie	dunwoody	dwyer
donington	dorrell	drinkels	dulle	dup	dyck
donizetti	dorset	drinkhouse	dulles	dupetit	dyfed
donleavy	dorsey	drinkwater	duloc	dupont	dykstra
donn	dortmund	drogheda	dulux	duport	dylan
donna	dostoevsky	droitwich	dulwich	dupuy	dynes

dynmouth	ecj	eea	either	ellender	emilio
dysart	eckersley	eec	ekberg	ellery	emily
dyson	eckleburg	effie	ekeus	ellesmere	emirates
e	eco	efta	ekstrohm	elley	emlyn
ea	ecoli	ega	el	ellie	emma
each	ecowas	egan	elaine	ellington	emmanuel
eadie	ecr	egbert	elangovan	elliott	emmeline
eadwine	ecuador	egerszegi	elbe	elliott	emmerdale
eagleburger	ecuadorean	egerton	elbio	ellis	emmerson
ealing	ecuadorian	egf	elbridge	ellison	emmert
eames	edberg	eggar	elburn	ellsworth	emmie
eamon	edderly	eggars	elderslie	ellwood	emory
eamonn	eddie	egginton	eldon	elly	empson
eanes	eden	egham	eldorado	elmer	emrys
earle	edgar	eglevsky	eldredge	elmo	ems
earp	edgbaston	eglinton	eldridge	elmore	emyr
easington	edgebrook	egon	eleanor	elmwood	en
easingwold	edgewater	egremont	electra	eloise	ena
easson	edgeworth	egstrom	elena	elphinston	encina
eastbourne	edgware	egt	eleven	elsa	endara
easterhouse	edi	egypt	eleventh	elsie	endill
eastham	edie	egyptian	elfine	elsing	enfield
eastland	edinburgh	ehlers	elgar	elspeth	eng
eastleigh	edirne	eichman	elgin	elsworth	engel
eastman	edison	eichmann	eli	eltham	engels
easton	edith	eiffel	elias	elton	england
eastry	edmond	eiger	elie	elvis	engler
eastwick	edmonds	eight	eligio	elwell	english
eastwood	edmonton	eighteen	elijah	elwyn	englund
eaton	edmund	eighteenth	eline	ely	enid
eb	edmunds	eighth	elinor	elys	ennerdale
ebbetts	edna	eightieth	elios	elysee	ennis
eben	edouard	eighty	eliot	elzey	enniskillen
ebenezer	edrich	eileen	elisa	emanu	eno
ebert	eduard	eily	elisabeth	emanuel	enoch
ebor	eduardo	einstein	elise	embleton	enos
ebrd	edusei	eintracht	eliza	emburey	enough
ec	edvard	eire	elizabeth	emcee	enox
eccles	edvins	eireann	elizabethville	emerson	enrico
ecclesall	edward	eisa	ell	emi	enrique
ecclesfield	edwards	eiseler	ella	emil	enya
eccleston	edwin	eisenach	elland	emile	eo
ecgfrith	edwina	eisenhower	ellen	emilia	eochoaid

epa	escobar	europe	ezra	fand	faust
epervier	esher	european	f	fanny	faustus
eph	esk	eurotunnel	faber	fanon	faversham
ephesus	eskdale	eustace	fabia	fanshawe	fawcett
ephraim	esme	eustis	fabio	fantus	fawkes
ephrem	esmeralda	euston	fabiola	fao	fawley
epping	esmond	eva	fabius	faraday	fawly
eppler	esnards	evan	fabliau	farah	fay
eps	espagnol	evans	fael	farakka	faye
epsom	esquimalt	evanston	fagan	farc	fayed
epron	esrc	evegeni	fagg	farduilli	fayoud
epstein	esse	evelyn	fagin	fargo	fazio
erapalli	essen	everage	fahd	farias	fbi
erasmus	essene	everard	fahreiddin	faridabad	fc
ercp	essex	everest	fairbairn	farillon	fca
ergative	essie	everett	fairbanks	faringdon	fddi
erhard	esso	evernden	fairborther	farland	fdp
erhorn	estabrook	evert	fairbrother	farley	fe
eric	estaing	everthorpe	fairchild	farnborough	feargal
erica	este	everton	fairclough	farnham	fearon
erich	estella	every	fairfax	farnill	featherstone
erickson	estelle	everybody	fairfield	faro	feaver
ericsson	esther	everyone	fairford	faroe	february
erik	eston	everything	fairham	farouk	federico
erika	estoril	evesham	fairley	farquaad	fedorov
erin	estrich	evian	faisal	farquhar	feetham
erith	et	ewan	fakenham	farquharson	fei
erlich	etc	ewart	fakhruddin	farr	fein
ernakulam	ethel	ewen	fakrid	farraday	feinstein
ernest	etienne	ewer	falcone	farrah	feis
ernie	etna	ewing	falconer	farrar	feldman
ernst	eton	ex	faldo	farrell	feldwebel
eros	eubank	except	falk	farrer	felicia
errol	euclid	exe	falkenhayn	farringdon	felipe
ershad	eugen	exeter	falkirk	farrney	felix
erskine	eugene	exmoor	falkland	fascella	felixstowe
eruera	eugenie	exmouth	falklands	fashanu	fellini
ervin	eunice	exxon	fallon	faso	fellowes
ervine	euratom	eyadema	fallowfield	fatah	felmingham
erwin	euripides	eyfion	falmouth	fatehpur	felske
es	euro	eyre	faludi	fatima	feltham
esa	eurodisney	eyres	fam	faulkner	felton
esau	europa	eysenck	famagusta	faure	feltrim

fenari	fi	fisher	florio	forrester	frances
fendalton	fianna	fishwick	florrie	forster	francesca
fenella	fidel	fisk	flossie	forsyth	francesco
feng	fido	fiske	floy	forsythe	franchise
fenland	fiedler	fison	floyd	fortescue	franchot
fenn	fielden	fittipaldi	flr	forth	francia
fenna	fieldhouse	fitz	flynn	fortieth	francie
fennell	fienne	fitzalan	flyte	fortin	francine
fenner	fiennes	fitzgerald	fmln	fortnum	francis
fenton	fifa	fitzhugh	fn	fortran	francisco
fenwick	fifteen	fitzormonde	fo	fortuna	frank
ferdi	fifteenth	fitzpatrick	foerster	forty	franco
ferdinand	fifth	fitzroy	fogarty	fosdyke	francois
ferdinando	fiftieth	fitzwater	fogelson	fosse	franjo
fergeson	fifty	fitzwilliam	fogerty	fossway	frankel
fergie	figaro	five	fogg	fothergill	frankenstein
fergus	figone	flackwell	foggies	foucault	frankford
ferguson	fike	flaherty	foggitt	fouchard	frankfurt
fergusson	filbert	flanagan	foinavon	foudy	frankie
feringa	filey	flanders	fokine	foulger	franklin
fermanagh	filmer	flannery	folene	foulkes	franklyn
fernand	fimbra	flashman	foley	foulsham	frans
fernandes	finan	flaubert	folkestone	fountainbleau	franz
fernandez	fincara	flavia	follett	four	frapp
fernando	finchley	fleetwood	fon	fourier	fraser
fernay	findlay	fleischmann	fonda	fournier	fraserburgh
fernberger	findon	fleisher	fonta	fourteen	fratton
fernie	finlay	fleming	fontaine	fourteenth	frawley
fernside	finlayson	fletcher	fontainebleau	fourth	frayn
ferranti	finmeccanica	fleur	fontana	fowey	frazer
ferrari	finnan	fleury	fontanellato	fowler	frazier
ferraris	finnegan	flimm	fonteyn	foxton	fred
ferreira	finney	flinders	fontwell	foxx	freda
ferrers	finniston	flintshire	for	foy	freddie
ferrier	finnochio	fln	forbes	foyle	freddy
ferris	finsbury	flo	forbs	fp	frederic
ferryhill	fiona	flodden	forde	fpcr	frederica
few	firestone	florence	fordham	fpr	frederick
fewer	first	flores	forfar	fra	freedman
ffestiniog	fis	floresville	forman	frampton	freemans
fg	fischer	floreya	formby	fran	freemantle
fgd	fishbane	florian	formosa	franca	freeport
fhcima	fishbein	florida	forrest	france	freiburg

freida	furby	gallagher	garibaldi	gavin	genitive
freitas	furhmann	galliano	garin	gaviria	gennifer
french	furness	gallie	garnett	gawd	genoa
frere	fw	gallipoli	garnsey	gaya	genscher
freyer	fyfe	galloway	garrard	gayatri	gentians
frg	fylde	gallowgate	garrett	gaye	genz
frick	g	gallup	garrick	gayla	geoff
friday	ga	galpin	garry	gayle	geoffery
fridolanna	gabby	galsworthy	garson	gaynor	geoffrey
friedman	gabon	galt	garston	gaza	georg
friedrich	gabor	galton	garth	gazza	george
friedrichstadt	gabriel	galveston	gartmore	gazzer	georges
friedrichstrass	gabriela	galvone	garvell	gb	georgetown
e	gabrielle	galway	garvey	gbh	georgette
friern	gaby	gamal	garvin	gc	georgi
frind	gaddafi	gambia	gary	gcc	georgia
frith	gaekwad	gamsakhurdia	garybecker	gchq	georgiade
frito	gael	gan	garza	gdansk	georgiades
fritz	gaelic	ganahl	gascoigne	gdr	georgiana
frizington	gagarin	gandak	gascony	gdynia	georgie
frizzell	gage	gandalf	gaskell	ge	georgina
frodo	gaidar	gander	gasparini	geary	gephardt
froggatt	gail	gandhi	gass	gebrec	geraghty
froghall	gaines	gandhiji	gassendi	gec	geraghtys
from	gainesville	gandhiteit	gaston	gedanken	geraint
frome	gainsborough	ganga	gateshead	geddes	gerald
frye	gaiter	ganges	gatore	gedge	geraldine
fsa	gaitskell	ganguli	gatsby	geduld	gerard
fsln	gajdusek	gangulli	gatt	geena	gerd
fuchs	gajendra	gannon	gattling	geigy	gere
fuchsberg	gajendragadka	gaonkar	gatwick	geikie	gerhard
fuego	r	garamendi	gatz	geils	germain
fuhrmann	galahad	garbo	gauda	gelder	germaine
fuji	galapagos	garc	gauer	geldof	german
fujimori	galashiels	garcia	gauguin	geldorf	germania
fujitsu	galatians	garda	gaulle	gelson	germany
fulbright	galbraith	gardanibagh	gaumont	gemiini	geroge
fulham	galerie	gardiner	gauss	gemma	gerosa
fulke	galgenhumor	gardner	gautam	gemmell	gerrard
fullerton	galicia	garel	gautier	genet	gerrards
fulling	galilee	gareth	gav	geneva	gerry
fulton	galileo	garfield	gavaskar	genevieve	gerstenzang
funchal	gallacher	gargy	gaveston	genghis	gerstner

gert	gielgud	giovanna	glossop	goldberg	goram
gertie	giffen	giovanni	gloucester	goldenrods	goran
gertrude	gifford	gir	gloucestershir	golders	gorbachev
gerulaitis	giggs	girgory	e	goldie	gorbals
gesner	gigli	girnar	glover	golding	gorby
gestapo	gil	girobank	glyn	goldman	gorder
get	gilbert	girton	glyndebourne	golds	gordin
gets	gilbertson	gisborne	glynderbourne	goldstein	gordon
getting	gilbey	giscard	glynis	goldsworthy	gordy
getty	gilchrist	gisela	glynn	goldthorpe	gore
gettysburg	gildas	giselle	gm	goldwater	goregaon
gevurtz	giles	gittens	gmb	goldwyn	goreng
gfand	gilford	gittins	gmbh	gollancz	gorky
ggf	gilfoyle	giulio	gmc	gollapalem	gorleston
gh	gill	giuseppe	gmelin	gomes	gorman
ghali	gilles	gl	gmtv	gomez	gort
ghana	gillespie	gladiateur	goa	gomoh	gorton
ghanaian	gillette	gladstone	goan	gomorrah	gosford
ghavai	gillhaus	gladys	gobbi	gonda	gosforth
ghavri	gillian	glaise	gobi	gondal	gospic
ghaziabad	gillibrand	glamorgan	godalming	gondalalso	gosport
ghent	gillick	glanville	godber	gonz	goss
ghormley	gilligan	glasgow	goddard	gonzales	gosse
ghouse	gilling	glastonbury	godfrey	gonzalez	goswami
ghs	gillingham	glaxo	godmere	gooch	got
ghulam	gillis	glc	godolphin	goodall	gote
giacometti	gilly	glenavon	godstowe	goode	gotham
giacomo	gilman	glencoe	godwin	goodfellow	gothenburg
gian	gilmore	glenda	goeas	goodie	gottenburg
giancarlo	gilmour	glendinning	goebbel	gooding	goudhurst
giandomenico	gilroy	gleneagles	goebbels	goodis	gough
giani	gimbel	glenfield	goel	goodison	gould
gianni	gimmelman	glenmore	goering	goodman	goulding
gibb	gimmi	glenn	goethe	goodwin	gounod
gibbon	gimms	glennon	goff	goodwood	govan
gibbons	gina	glentoran	goffman	googol	gover
gibbs	ginnetti	glenys	gog	goole	goverdhan
gibraltar	ginns	glimco	gogh	goonda	govett
gibran	ginny	glison	gogol	goondas	govind
gibson	gino	gloppholm	goh	goowdas	govinda
giddens	ginsberg	gloria	goire	gopal	gow
gide	ginsburg	gloriana	goka	gopalaswami	gower
gideon	giorgio	glos	golak	gorakhpur	gowers

gowling	greece	grierson	grzesiak	gupta	habsburg
goya	greek	griff	gs	gupte	hackett
goyal	greenall	griffin	gt	gurbachan	hackney
gpa	greenaway	griffith	gu	gurd	hackstaff
gpc	greenbank	griffiths	guam	gurder	had
gr	greenberg	grigson	guangzhou	gurdwara	haddix
graaff	greene	grillet	guatemala	gurney	hades
graceland	greenham	grimaldi	guayas	gurpal	hadlee
grachev	greenhill	grimes	gubberford	gursel	hadley
gracie	greenidge	grimm	guccione	gurudwara	hadn
grady	greenland	grimma	gudonov	gus	hadrian
graefe	greenock	grimond	guerin	guscott	hafez
graeme	greenpeace	grimsby	guernica	gustav	haffkine
graf	greenville	grimwood	guernsey	gustave	haffner
grafon	greenway	grindal	guerrero	gutekunst	hagan
graham	greenwich	grindley	guestwick	gutenberg	hagen
grahame	greenwood	grinsfelder	guggenheim	guthrie	hagner
grahams	greer	grinstead	guglielmo	guyana	hague
grahamstown	greg	gris	guha	guyer	hahn
grainger	gregg	griselda	guididas	guzman	hahnemann
grainne	gregor	grobelaar	guido	gw	hai
grammar	gregorio	gromyko	guil	gwen	haifa
grammatical	gregory	groote	guildford	gwendolen	haig
grampian	gregson	gros	guillaume	gwendoline	haigam
grampians	greig	grose	guillermo	gwenlan	haigh
gramsci	grenada	grosse	guinea	gwent	haile
granada	grendon	grossman	guinness	gwr	hailes
granaski	grenfell	grosvenor	guisborough	gwyn	hailsham
granby	grenier	grosz	gujarat	gwynedd	hain
grande	grenoble	groucho	gukuldas	gwyneth	hainault
grangemouth	grenville	groupe	gulab	gwynne	haines
granger	gresham	grover	gulliver	ha	hairflair
grangetown	gresley	grovers	gummer	haaek	haislett
grantham	greta	gruagach	gundappa	haakon	haitink
granville	gretchen	grubb	gundovald	haal	hakim
grasmere	gretna	gruchy	gunn	haarhuis	haksar
grattan	greville	grumman	gunnar	haas	haksat
gravesend	grewal	grundig	gunnell	haase	hakuna
grayson	gribbon	grundy	gunny	haastrup	hal
gre	gribbons	grunte	gunson	haath	halabja
greatrex	grice	grunwald	gunter	haberma	haldane
greaves	griefswald	grunwick	gunther	habib	hales
grechko	grieg	gruoned	guntram	habibnagar	halesowen

halewood	hammett	hansard	harishchandra	haruki	havant
haley	hammond	hansen	harjinder	harvard	have
half	hammonds	hanson	harker	harvey	havel
halford	hammons	hants	harkin	harveys	havelock
halifax	hamnett	hanumakonda	harkness	harvie	havemann
halladay	hampden	hanumant	harland	harwell	haven
hallam	hampshire	hanwell	harlech	harwich	haverford
halle	hampshires	hapsburg	harlem	harwood	havers
halleck	hampson	hapsburgs	harley	haryana	havilland
halley	hampstead	haq	harling	harz	having
halliday	hampton	hara	harlingen	has	havisham
hallinan	hamptons	harald	harlow	hasan	havre
halling	han	harare	harman	hase	havvie
halliwell	hanbury	harbison	harmon	haseen	hawaii
halloween	hancock	harborough	harold	hashemi	hawaiian
hallowell	handel	harbury	harpenden	hashim	hawarden
hallward	handford	harcastle	harper	hashimi	hawes
halsall	handley	harcharan	harpercollins	hashmath	hawick
halsbury	handsworth	harcourt	harpers	haskins	hawke
halsey	hangul	hardacre	harriet	haslam	hawkes
halstead	hanington	hardaker	harriman	hasledene	hawkins
halton	hanja	hardbroom	harrington	haslemere	hawksley
halverson	hank	hardcastle	harris	hasn	hawley
hama	hankin	hardie	harrison	hassan	haworth
hamals	hanks	harding	harrity	hastings	hawthorne
hamas	hanlan	hardingham	harrod	hatcher	hayden
hambledon	hanley	hardman	harrods	hateley	haydn
hambleton	hanlon	hardwar	harrogate	hatfield	haydock
hambric	hann	hardwick	harsanyi	hathaway	haydon
hambro	hanna	hardwicke	harsnet	hatherby	hayek
hambros	hannah	harefield	hart	hatherley	hayes
hamburg	hannay	harewood	hartal	hatt	hayley
hamer	hannele	harford	hartford	hattersley	hayling
hamey	hannen	hargett	hartigan	hattie	hayman
hamid	hannibal	hargreave	hartlepool	hatton	haymarket
hamilton	hannon	hargreaves	hartley	haug	haynes
hamirbhai	hannover	hari	hartman	haugen	hayter
hamish	hanoi	haridaspur	hartmann	haughey	hayward
hamlin	hanover	harihans	hartmanns	haughton	haywards
hamlyn	hanoverian	harijan	hartnell	hauser	hayworth
hamm	hanoverians	harijans	hartselle	haute	hazelden
hammarskjo	hans	haring	hartsfield	hauxwell	hazelhurst
hammersmith	hansa	haringey	hartweger	havana	hazell

hazeman	heidrick	hendrix	herridge	hickey	hinkley
hazlemere	heil	hendry	herringswell	hickman	hinsdale
hazlitt	heilman	henfield	herrington	hickok	hinshelwood
hb	heimlich	hengesbach	herriot	hicks	hinton
hbv	heine	henke	hers	hickson	hippo
hc	heineken	henley	herschensohn	higgins	hippodrome
hcima	heinemann	hennessey	herself	higgs	hir
hcl	heinkel	hennesy	hershey	higham	hira
hco	heinrich	henning	herta	highbury	hirohito
hd	heinz	henri	herter	highfield	hiroshima
hdtv	heinze	henrietta	hertford	highgate	hirsch
he	heisenberg	henriette	hertfordshire	highgrove	hirst
headingley	heitschmidt	henrique	herts	highly	his
headington	hekmatyar	henry	herwarth	hilaire	hislop
headlam	hela	hensingham	herzogovina	hilary	hitachi
headley	helen	henson	herzog	hilbert	hitchcock
healey	helena	hepburn	heselbarth	hilda	hitchin
healy	helene	heptachlor	heseltine	hilgeman	hitler
heaney	helens	hepworth	heselton	hilgers	hiv
hearle	helfeld	hepzibah	hesketh	hillary	hl
hearn	helga	her	hesley	hillcrest	hla
hearth	heli	hera	heslop	hilley	hmi
heathcliff	helier	herbert	hess	hillier	hmso
heathcoat	heller	herbie	hesse	hillington	hmv
heathcote	hello	hercegovina	hester	hillman	hn
heathrow	hellstrom	hercule	heston	hillsboro	hnc
heaton	hellyer	hercules	hetherington	hillsborough	hnd
hebburn	helmsley	hercus	hewett	hillsdale	hoak
hebden	helmut	hereford	hewey	hillview	hoaps
heberley	helsinki	herefordshire	hewitt	hilton	hoare
hebridean	hemel	hereward	hewlett	him	hobart
hebrides	hemi	heriot	hewson	himachal	hobbes
hector	hemingway	herluin	hexham	himansu	hobbs
hedley	hemlington	herman	heyford	himmat	hobhouse
heeley	hemmings	hermann	heysham	himmler	hobson
hefner	hemphill	hermes	heywood	himself	hockaday
hegarty	hempstead	hern	hezbollah	hin	hockman
hegde	hemus	hernan	hf	hinchliffe	hockney
hegel	hencke	herne	hg	hinckley	hodder
heid	henderson	herod	hiatt	hinderclay	hoddle
heidelberg	hendon	herodotos	hibbard	hindley	hodge
heideman	hendrick	herodotus	hibbert	hindustan	hodges
heidi	hendrie	herrick	hibs	hines	hodgkin

hodgkinson	holmstrom	horseley	hoyle	humphry	hvs
hodgson	holroyd	horsham	hoyle	hun	hwim
hodkinson	holst	horsley	hoyles	hundred	hy
hodosh	holstein	horsman	hoysala	hundredth	hyam
hodson	holt	horst	hoyt	hungarian	hyannis
hoechst	holyfield	horstman	hoys	hungary	hyatt
hoekstra	holyhead	horton	hp	hungerford	hyde
hoffa	holyrood	horus	hqv	hunslet	hyderabad
hoffman	holywell	hosen	hrawi	hunterston	hydro
hoffmann	hollywood	hoskin	hrun	huntingdon	hyland
hofmann	home	hoskins	hsia	huntington	hylton
hogan	homer	hoskyns	hsien	huntingtons	hyman
hogarth	homi	hosni	htv	huntley	hyndman
hogberry	honda	hotspur	hu	huntly	hypersparc
hogg	honecker	hough	huang	huppert	hythe
hoggart	honeycutt	houghton	huascar	hur	hyundai
hoipol	honeymoon	houk	hubbard	hural	hywel
hokkaido	honeywell	hounslow	hubbell	hurd	i
hokura	hong	housman	hubble	hurley	iaaf
holabird	hongkong	houston	huber	hurray	iaea
holbein	honolulu	houtz	hubert	hurst	iago
holborn	hononegah	hovarter	hubie	hurwitz	iain
holbrook	honor	hove	huckin	hurworth	ian
holden	honoree	hovia	huddersfield	husayn	ianthe
holderness	hoogstratten	how	hudleston	huskies	iba
holdsworth	hooke	howard	hudson	husrev	ibanez
holford	hoomey	howards	huen	hussain	ibf
holgate	hooper	howarth	huey	hussainiwala	ibiza
holkham	hopkin	howden	huffman	hussein	ibn
holliday	hopkins	howdendyke	hugh	hussey	iboa
hollidaye	hopwood	howe	hughes	huston	ibrahim
hollingworth	horace	howell	hughie	hutchins	ibrox
hollins	horan	howells	hugo	hutchinson	ibsen
hollis	horatia	howes	huia	hutchison	ica
holliwell	horatio	however	hulme	hutments	icc
holloway	horbury	howey	hulstijn	hutt	ici
hollowell	hord	howie	hulton	hutton	ickes
hollywood	hornby	howle	humber	huub	icl
holm	horncastle	howlett	humberside	huw	icm
holman	horne	howsam	hume	huxley	ico
holme	horner	howson	humphrey	huy	icrc
holmes	horowitz	hoyer	humphreys	huyton	ics
holmfirth	horrocks	hoylake	humphries	hvk	id

ida	imrie	interjection	isaacson	iwc	jakarta
idaho	imro	internationale	isabel	ix	jake
idc	ims	into	isabella	ixi	jakob
idi	in	intransitive	isabelle	ixora	jakobson
idris	ina	invercargill	isaiah	izetbegovic	jaldbazi
ie	inbucon	inverness	isambard	izzedin	jaleo
iengar	ince	ioannou	isay	izzie	jalianwalabag
iepa	inder	iona	isbn	ja	jamaica
ierulli	indiana	ione	isdn	jabbar	james
ieuan	indianapolis	iowa	ise	jabir	jameson
if	indira	ip	isherwood	jacana	jamie
ifield	indo	iphig	isis	jack	jamieson
igg	indochina	ipstones	islamabad	jackie	jamila
iglehart	indrayani	ipswich	islay	jacklin	jamison
ignatius	indu	ipuky	isleworth	jackman	jammu
igor	indus	iqbal	islington	jacko	jamshedpur
ii	indutai	ira	ismail	jackson	jamshetji
iia	indycar	iran	isn	jacksons	jamuna
iii	ine	iraq	isnas	jacksonville	jana
ike	inez	iraqis	iso	jacky	janardan
il	informix	iras	isobel	jacob	janardhan
ile	ing	irb	isoud	jacobi	janata
ilea	ingard	irc	iss	jacobs	janatha
ilford	inge	ireland	issey	jacobsen	jane
ilfracombe	ingham	irena	istanbul	jacobson	janeiro
iliescu	ingleborough	irene	it	jacqueline	janes
ilkley	ingleby	irian	italian	jacques	janet
illingworth	ingleton	irigaray	italy	jacqui	janey
illinois	inglis	irina	itam	jacquie	janice
illsley	ingmar	irirangi	ite	jaffa	janie
illych	ingo	irish	itn	jaffery	janine
ilo	ingram	irkutsk	its	jag	jannie
ilona	ingres	irma	itself	jagatjit	jansen
ilsa	ingrid	ironbridge	itv	jagdish	janssen
ilyas	ini	ironside	itzhak	jageshwar	january
ilyushin	inkatha	irvin	iv	jagger	janus
imelda	inmos	irvine	ivan	jaggit	japan
imf	inna	irving	ivanhoe	jagjivan	japanese
imogen	innes	irwin	ivanisevic	jahsaxa	jaq
imperative	innsbruck	is	ives	jaideep	jaquetta
impey	intel	isa	ivoire	jaime	jardin
imran	interbank	isaac	ivor	jaipur	jardine
imre	intercity	isaacs	iwasaki	jaisimha	jared

jarman	jef	jervois	jna	jonas	juarez
jaross	jeff	jerzy	jo	jonathan	judaea
jarrett	jefferies	jespersen	joachim	jonathon	judah
jarrow	jefferson	jess	joan	jones	judas
jaruzelski	jeffery	jessamy	joanna	jong	judd
jarvis	jeffrey	jessamyn	joanne	jongh	jude
jason	jeffreys	jesse	joao	joni	judea
jasper	jeffries	jessel	joaquin	jonson	judi
jaswant	jehan	jessica	jobson	joplin	judith
jat	jehana	jessie	jocelyn	jorda	judithes
jatha	jehangir	jessop	jochen	jordan	judson
jatin	jehangirpur	jessy	jock	jordanian	judy
jatti	jehovah	jethmalani	jodami	jordanstown	juhu
java	jekub	jett	jodi	jordi	jules
javed	jekyll	jewitt	jodie	jorge	julia
javer	jem	jewkes	jody	jos	julian
javier	jemima	jewry	joe	jose	juliana
jawahar	jen	jewson	joel	josef	julie
jawaharlal	jena	jeyaretnam	joey	joseph	julien
jawans	jenkin	jez	joffre	josephine	juliet
jawarharlal	jenkins	jezrael	joffrey	josephs	juliette
jaya	jenkinson	jf	joh	josephus	julio
jayantilal	jenks	jfk	johan	josh	julius
jayaprakash	jenna	jh	johann	joshi	jullundur
jayapuram	jenner	jha	johanna	joshua	july
jayaram	jennie	jhootha	johanne	josiah	jumbe
jayawardane	jennifer	jhunjhunu	johannes	josie	junagadh
jayawardene	jennings	jiang	johannesburg	joslin	june
jayne	jenny	jill	johansson	joss	juneja
jazzbeaux	jens	jilly	john	jossy	jung
jb	jensen	jim	johnnie	jotan	junichi
jc	jeremiah	jimbo	johnny	joule	junkers
jd	jeremy	jimi	johns	jourdain	juntao
je	jeremybentha	jimmie	johnson	jovanovic	jura
jean	m	jimmy	johnsonville	jove	juras
jeanette	jerger	jin	johnston	jovi	jurges
jeanne	jericho	jinkinson	johnstone	jowar	jurnet
jeannie	jermyn	jinkwa	johnvonneuma	jowett	jussel
jed	jerome	jinnah	nn	joyce	justin
jedburgh	jerry	jinny	jon	jozsef	justine
jeddah	jersey	jitendra	jona	js	justinian
jeejeebhoy	jerusalem	jj	jonadab	jtr	juventus
jeeves	jervis	jm	jonah	juan	k2

kabaka	kan	karr	kaye	kendal	kerrville
kabir	kanara	karsan	kazin	kendall	kerry
kabul	kanawa	kartar	kc	kendra	kersey
kadamkuan	kandinsky	kas	kcl	kendrick	kershaw
kadapara	kane	kasai	kean	kendricks	kershbaum
kader	kang	kasavubu	keane	kenilworth	kerstin
kadhafi	kanin	kashi	keating	kennan	keshavan
kadur	kankakee	kashmir	keaton	kennebunkpor	kessler
kafka	kannada	kasim	keats	t	kestner
kahan	kano	kaske	kebir	kennedy	keswick
kahlil	kans	kasparov	kedzie	kennedys	ketcham
kahlo	kansas	kassem	kee	kennesaw	kettering
kahn	kant	kassim	keefe	kennet	ketti
kai	kanti	katanga	keeffe	kenneth	kettlewell
kaifu	kanu	katangan	keegan	kennington	kev
kaikohe	kapala	katangans	keele	kennsaw	kevin
kaikoura	kapil	kate	keeler	kenny	kew
kailash	kaplan	kateb	keenan	kensington	keynes
kailashpati	kapnek	kath	keeney	kent	kgb
kaimai	kappa	katharine	kegan	kentfield	kha
kaitaia	kara	katherine	kegham	kenton	khadi
kaka	karabakh	kathi	keighley	kents	khagaul
kakodkar	karachi	kathie	keir	kentucky	khalid
kalahari	karajan	kathleen	keith	kenwood	khalsa
kalam	karakoram	kathmandu	keizer	kenya	khan
kalashnikov	karel	kathryn	kel	kenyon	khandsari
kalchu	karelius	kathy	kelham	kepler	khanna
kalelkar	karen	katie	kell	ker	kharak
kallam	karena	katrina	kellard	kerala	khartoum
kalonji	karim	katy	keller	kerby	hasbulatov
kalpana	karin	katya	kellett	kercheval	khed
kalyanpur	karkal	katz	kelley	kerk	khetri
kamal	karkovice	kaufman	kellogg	kerkorian	khiang
kamaraj	karl	kaufmann	kelly	kerla	khmer
kamarck	karlstad	kaufnabb	kelsall	kermit	khmers
kamens	karnal	kaul	kelsey	kern	khojas
kaminsky	karnataka	kaunda	kelso	kernaghan	khomeini
kamlapati	karns	kavadiguda	kelvin	kernoff	khoon
kamlesh	karol	kavanagh	kemm	kerr	khouang
kamm	karori	kawasaki	kemp	kerrey	khrushchev
kampala	karpoori	kawecki	kempe	kerri	khrush
kampfner	karpov	kawhia	kempton	kerrison	khrushchev
kampuchea	karpuri	kay	ken	kerrod	khrushchov

khurana	kimberly	kirpal	knettishall	kop	krenz
khushoo	kimbolton	kirsche	knighton	kopkind	kretchmer
khushwant	kimmell	kirsten	knightsbridge	koppes	kripa
khyber	kimon	kirsty	knitmaster	kopyion	kripalani
kiah	kimpton	kisans	knokke	koraloona	krishan
kiawah	kincardine	kishan	knopper	korbut	krishna
kidd	kindersley	kishore	knorr	korda	krishnan
kidderminster	kingdon	kishwaukee	knossos	korea	krishnas
kidlington	kinghorn	kislev	knott	koresh	krishnaswamy
kidwai	kingsdale	kissak	knowles	korman	kriss
kieffer	kingsholm	kissinger	knowsley	korn	kristeva
kiefferm	kingsley	kitchencraft	knox	koro	kristiansund
kiel	kingston	kitchener	knut	kosevo	kristina
kier	kingsway	kitson	knutsford	kosovo	krobo
kieran	kingswood	kitston	knutsson	kotagiri	croger
kierkegaard	kington	kittredge	kobler	kothari	krogers
kiev	kinkel	kitzinger	koch	kothi	krol
kiff	kinnear	kivett	kochanek	kotrashetti	kronweiser
kiichi	kinney	kivu	kochaneks	kottke	krueger
kika	kinnock	kiwanis	kodaikanal	kotwali	kruger
kilavenmani	kinross	kiwomya	kodak	kou	kruize
kilbirnie	kinsey	klaus	kodur	kowalski	krunchu
kilbride	kiowa	kleber	koeries	kowloon	kruschev
kilburn	kipling	klee	koh	kozyrev	krushchev
kilda	kira	klees	kohli	kp	krushchov
kildalton	kiran	klein	kohnstamm	kpmg	ku
kildare	kirby	kleinwort	koi	kraft	kuala
kilgannon	kirgizstan	klerk	kolandaivelu	kragan	kuan
kilhour	kiri	klestil	kolb	kragen	kubala
kilian	kirk	klieman	kolchinsky	krainik	kubek
kilifi	kirkby	klift	kolhapur	krajina	kufra
kililngsworth	kirkcaldy	kline	koliwada	krakatoa	kuhn
kilimanjaro	kirkendoll	klipspringer	kolpakova	krakow	kukoc
kilkenny	kirkham	klm	kolwezi	kramer	kulandaivelu
killebrew	kirkland	kloman	kombolcha	kranish	kulbir
killeen	kirklees	klondike	kong	kraprayoon	kuldip
killingsworth	kirkpatrick	kmt	konrad	krashen	kulkarni
killion	kirkuk	knapp	konstantin	krause	kumagar
kilmarnock	kirkup	knaresborough	konstantine	kravchuk	kumar
kilpatrick	kirkwall	knauer	kooka	kray	kumasi
kilroy	kirkwood	knecht	kooning	kreike	kumbakonam
kim	kirmani	knelle	koons	kreisky	kunkel
kimberley	kirov	kneset	kooti	kreisler	kuo

kurdistan	laettner	lancashire	larne	lawson	leftfield
kuril	lafayette	lancaster	larry	lawton	lefty
kurile	lafe	lancelot	lars	layfield	leger
kuriles	lafferty	lanchester	larsen	layton	legers
kurils	lafitte	lancia	larson	lazarus	legg
kurla	lagan	lancs	las	lazenby	legge
kurmis	lagerfeld	land	lasalle	lazio	lehane
kurosawa	lago	landau	lasmo	lazlo	lehman
kursk	lagos	landbrokes	last	lazras	lehmann
kurt	laguardia	lander	laszlo	lb	lehner
kusha	laguna	landesmann	latham	lc	lehotay
kushnick	lahore	landis	lathi	lcc	lehrman
kusum	lai	landor	latika	lch	leicester
kuttner	laidlaw	landress	latimer	lcy	leicestershire
kvast	laine	landrum	latinovich	ld	leiden
kwan	laing	landry	lauda	lddc	leigh
kwazulu	lajpat	landsat	lauder	ldp	leighton
kwik	lakatos	lanfranc	lauderdale	le	leila
ky	lakeland	lang	laufer	leabrook	leinster
kydd	lakepowell	langbaurgh	laughlin	leachate	leipzig
kyi	laker	langdale	laughton	leadbetter	leitch
kyle	lakewood	langdon	laura	leafe	leith
kylie	lakhdar	lange	lauren	leah	lel
kyng	lal	langer	laurence	leakey	leland
kyoko	lala	langford	laurent	leamington	lellouche
kyoto	lalage	langham	laurie	leander	lemn
kyte	lall	langley	lausanne	leanne	len
l	lally	langsdorf	lauterberg	leapor	lena
la	lalsingh	langston	lautrec	lear	lendl
laad	lamar	langton	lautro	leary	leng
laal	lamarck	lanin	laval	leatherhead	lenin
labov	lambert	lansbury	lavaughn	leavis	leningrad
lacey	lambeth	lansdowne	laverne	leavitt	leninism
lachlan	lambie	lansing	lavers	lech	lennie
lacroix	lamborghini	lantolf	lavery	leda	lennon
lacs	lambourn	lanyon	lavin	ledford	lennox
lada	lambton	lanzarote	lavinia	lee	lenny
ladbroke	lamont	laplace	lawford	leeds	leno
ladbrokes	lampard	lapointe	lawler	leeming	lenobel
ladd	lampeter	lara	lawley	lees	lenton
ladgham	lamplugh	lardner	lawrence	leese	lenygon
ladislav	lanark	larimer	lawrenceville	lefebvre	leo
ladli	lanarkshire	larkin	lawrie	lefevre	leominster

leon	lewes	lilian	lionel	llanrumney	lomax
leonard	lewin	lilienthal	liphook	llar	lombard
leonardo	lewis	lille	lipman	ller	lombardy
leonards	lewisham	lilleshall	lipton	llewellyn	lomea
leone	lex	lilley	lirio	llewelyn	lomond
leonid	lexandro	lillian	lisa	lloyd	londesborough
leonie	lexington	lillie	lisabeth	lloyds	london
leonora	leyburn	lilliput	lisbie	lm	londonderry
leonore	leyden	lilly	lisbon	lng	longbridge
leopold	leyland	lilongwe	lisburn	loach	longchamp
leopoldville	leys	lim	lisle	lobo	longchamps
lepine	leyton	lima	liss	lochals	longfellow
lerner	leytonstone	limaye	lissa	lochan	longford
leroy	lez	limehouse	lister	lochore	longinotti
les	lf	limoge	liston	lochroe	longman
lesley	lfa	limoges	liszt	locke	longshot
leslie	lfs	limonalt	lita	lockerbie	longtime
lesotho	lh	limousin	lithuania	lockhart	longton
less	lhasa	limpar	litlewoods	lockheed	longue
lessing	li	lin	littlehampton	lockies	lonrho
lester	liam	lina	littlejohn	lockwood	lonsdale
lestor	liaquat	lincoln	littlewood	lockyer	looe
letchworth	libby	lincolnshire	littlewoods	loe	loong
letheren	liberia	lincs	liu	loeb	loosley
letitia	libor	linda	livanos	loen	lopez
letterman	lichfield	lindauer	livermore	loewe	lopham
letts	lichtenhagen	lindisfarne	liverpool	loftus	lorain
lev	lichtenstein	lindley	livesey	logan	loren
leven	liddell	lindsay	livingston	logie	lorenz
leventis	liddle	lindsey	livingstone	logue	lorenz
leverett	lieberman	lindy	liz	loha	lorenzo
leverhulme	liechtenstein	lineker	liza	lohia	loretta
levi	liefers	linfield	lizzie	loi	lori
leviable	lier	linford	lizzy	loire	lorimar
levin	liffe	lingfield	lj	lois	lorimer
levine	liffey	linighan	ljj	lok	lorlyn
levinson	ligget	linley	ljubljana	lokp	lorna
leviticus	lightfoot	linlithgow	ll	lol	lorne
levitt	liguria	linnaeus	llanberis	lola	lorraine
levittown	ligurian	linthorpe	llandrindod	loliem	lorri
lew	like	linton	llandudno	lom	lorrimer
lewcock	lil	lintott	llanelli	loman	lorrimore
lewellyn	lili	linwood	llangollen	lomas	lorton

los	lubbock	lully	lyman	macdonald	madawaska
loschky	lubor	lulu	lyme	macdonalds	maddocks
lossiemouth	lubyanka	lum	lymer	macdougall	maddox
lothar	luc	lumia	lymington	macduff	maddux
lothian	luca	luminarch	lyn	macedonia	madeira
lothians	lucan	luminoso	lynagh	macedonian	madeleine
lotta	lucas	lumley	lynda	macewan	madeline
lottie	lucca	lummus	lyndhurst	macfadyen	madge
lou	luce	lumpe	lyndon	macfarlane	madhu
louchheim	lucenzo	lumpur	lyndsey	macgregor	madhubani
loudon	luch	lumsden	lyne	macgregors	madhya
lough	lucia	lumumba	lyneham	machesney	madison
loughborough	lucian	luna	lynn	machhipur	madra
louie	luciano	lund	lynne	machiavelli	madras
louis	lucie	lundy	lynwood	machin	madrid
louisa	lucien	lunn	lyon	machynlleth	madurai
louise	lucifer	lupone	lyonnais	macina	maduro
louisiana	lucille	lurgan	lyons	macintosh	mae
louisville	lucinda	lurie	lyster	macintyre	maeve
lourdes	lucke	lusaka	lytham	mack	mafatal
louth	lucker	luscombe	lyttelton	mackay	mafatlal
louvre	lucknow	luther	lytton	mackenzie	mafatlal
lovat	lucy	luthuli	m	mackeson	maff
lovejoy	ludgate	luton	m1	mackey	mafokate
lovelace	ludlow	lutterworth	ma	mackie	mafouz
lovell	ludo	lutyens	maas	mackinnon	mafras
lovelock	ludovico	lutz	maastricht	mackridge	magan
lovet	ludwick	luxembourg	mabel	maclaren	maganlal
lowe	ludwig	luxemburg	mabs	maclean	magarrell
lowell	luebbenau	luxmoore	mac	maclennan	magdalen
lowestoft	luebke	luxor	macao	macleod	magdalene
lown	luechtefeld	luxton	macari	macmillan	magee
lowndes	luette	luz	macarthur	macmurray	magellan
lowrey	luftwaffe	luzon	macarthy	macon	magere
lowry	lugg	lvmh	macaskill	macphail	maggie
lowther	luigi	lwt	macassar	macpherson	maggiore
lr	luis	lyall	macaulay	macquillan	magherafelt
lse	luisa	lybrand	macbeth	macrae	maghreb
lsi	luise	lydia	macbride	macsharry	magiera
lta	luk	lydney	macca	macwhorter	magill
ltb	luka	lydon	maccallum	macwhyte	magilton
ltp	luke	lyell	maccarthy	madagascar	maginnis
ltte	lukic	lyle	macclesfield	madan	maginot

magna	maisie	malo	mani	marc	maris
magnus	maison	malone	manila	marcel	marisa
magnusson	mait	maloney	manilow	marcelle	marius
magritte	maithilis	malory	manipur	marcelo	marivent
maguire	maitland	malpass	manjit	march	marje
mahabalipura	maitre	maltings	manjula	marchay	marjorie
m	majhi	malton	mankowski	marcia	marjory
mahadevan	majorca	maltravers	manley	marcie	markby
mahaffey	makarova	malvern	mann	marcile	marketings
mahal	makepeace	malvolio	mannargudi	marciulionis	markham
mahals	makin	mamallapura	mannheim	marco	markhor
mahant	mal	m	manohar	marconi	markov
maharajkumar	mala	mamet	manoharan	marcos	markovic
maharanees	malamute	mammini	manolo	marcus	markus
maharani	malathion	mamur	manoranjana	mardis	marland
maharashtra	malathy	managua	mansell	mare	marlboro
mahathir	malaviya	manakau	mansfield	marehay	marlborough
mahatma	malawi	manawatu	manson	marek	marlene
mahavir	malaya	manchanda	mantegna	marella	marler
mahe	malayalam	manchester	mantell	marg	marley
mahendra	malc	manchu	manthada	margaret	marlon
mahendru	malcolm	manchuria	manton	margate	marlow
maher	maldives	mancini	manuel	marge	marlowe
mahim	maldon	mandaje	manuhiri	margery	marmaduke
mahindras	malham	mandal	manukau	margo	marmmootil
mahler	malhotra	mandalay	manvers	margolin	marne
mahmood	mali	mandale	manville	margot	maroy
mahmoud	malibu	mandapakala	many	marguerite	marquardt
mahmud	malietuina	mandela	manzanola	maria	marr
mahon	malik	mandelson	mao	mariam	marriot
mahoney	malinowski	mander	maori	marian	marriott
mai	malkovich	mandeville	maozedong	mariana	marsden
maida	mallachy	mandru	maplecrest	mariani	marseille
maidenhead	mallaig	mandy	mappin	marianne	marseilles
maidment	mallender	manescu	mapplethorpe	marie	marsicano
maidstone	mallett	manet	maradona	marietta	marske
mair	malley	manfred	marais	marilla	marston
mailer	mallia	mangal	maramarua	marilyn	marta
maine	malling	mangalore	marathe	marimuthu	martell
mainwaring	mallory	mangaroo	marathwada	marin	martha
mainz	malmesbury	manhar	marbel	marino	marti
mair	malmo	manhattan	marbella	mario	martin
mairhi	malmud	manhatton	marburg	marion	martina

martinborough	masterson	maurier	mcc	mcdowell	mckee
martindale	masterton	maurin	mccabe	mcdunn	mckendrick
martine	masud	maurine	mccafferty	mceachern	mckenna
martinelli	mata	mauritania	mccall	mcelroy	mckenzie
martinetissimo	matata	mauritus	mccallen	mcelvaney	mckeown
martinez	mather	mauro	mccallum	mcelyee	mckie
martinho	mathers	mavis	mccambridge	mcenroe	mckimmie
martinique	matheson	mawhinney	mccammon	mcevoy	mckinlay
martino	mathew	maxey	mccann	mcewan	mckinley
martins	mathews	maximilian	mccarron	mcewen	mckinnon
marton	mathilde	maximus	mccarthy	mcfall	mckinsey
marty	mathur	maxine	mccartney	mcfarland	mcknight
martyn	matilda	maxwell	mccaughan	mcfarlane	mclachlan
marudoor	matisse	may	mccauley	mcgahon	mclaren
marulanda	matisses	maya	mccaw	mcgee	mclauchlin
marv	matka	mayall	mccclair	mcgeechan	mclaughlin
marvin	matlock	maybury	mccclay	mcgehee	mclean
marwick	matsushita	mayer	mcclellan	mcgeorge	mcleish
marwood	matta	mayfair	mcclelland	mcghee	mclellan
marxism	mattes	mayfield	mccclen	mcgill	mclelland
mary	matthew	mayhew	mccclendon	mcglone	mcmore
maryland	matthews	maynard	mccloy	mcglynn	mcleod
marylebone	matthewson	mayne	mcccluskey	mcgovern	mcloughlin
maryport	matthey	maynor	mccoist	mcgowan	mclure
marys	matthias	mayo	mccconnell	mcgown	mcmahon
marzillo	mattie	mays	mccconville	mcgrath	mcmaman
masab	matties	mazda	mccormack	mcgraw	mcmamus
masai	mau	mazdoor	mccormick	mcgregor	mcmaster
masani	mauch	mzeroski	mccowan	mcgrouter	mcmillan
masciarelli	maud	mazhais	mccooy	mcguffie	mcmullen
masefield	maude	mazowiecki	mccrea	mcguinness	mcmurdo
masha	maudie	mbube	mccready	mcguire	mcnab
mashem	maudling	mca	mccreery	mcgurk	mcnair
maskell	maugham	mcadams	mcculloch	mcgwire	mcnall
masklin	maulana	mcalester	mccullough	mchale	mcnally
maslin	maundy	mcalister	mccully	mchenry	mcnamara
masnem	maung	mcallister	mcdaniel	mchugh	mcaught
massachusetts	maunganui	mcalpine	mcdermott	mcillvanney	mcaughton
massawa	maunsell	mcarthur	mcdonald	mcilroy	mcealy
massey	maura	mcauley	mcdonalds	mcinnes	mceuil
massimo	maureen	mcauliffe	mcdonnell	mcintosh	mceuil
massingham	maurer	mcavennie	mcdougal	mcintyre	mcnulty
masson	maurice	mcbride	mcdougall	mckay	mcpartlin

mcphoe	megan	menezes	metcalfe	michie	millar
mcpherson	megarry	mengistu	metropole	michigan	millard
mcquaid	meghalaya	menlo	mets	michio	millbank
mcqueen	meh	menon	mettur	mick	mille
mcrae	mehmed	menor	metz	mickey	millet
mcroberts	mehrauli	menuhin	meuse	micky	millfield
mcseveney	mehta	menzies	mewith	microsoft	millicent
mcstay	mekong	mephisto	mexico	mid	millie
mctear	mel	mera	meyer	middlebrook	milligan
mcturk	mela	mercedes	meyers	middlecoff	million
mcwhirter	melancon	mercer	meyle	middleham	millionth
mcwilliams	melanie	merchiston	meyner	middlemass	millwall
me	melba	mercia	mez	middlemiss	milly
meacher	melbourne	merck	mezza	middlemore	milne
meade	melcher	mereana	mfn	middlesborou	milner
meadowbank	melford	meredith	mgadzah	gh	milngavie
meana	melinda	merganser	mgcl	middlesbroug	milosevic
meara	melissa	meriam	mgm	h	milroy
mearns	mello	meriden	mgn	middlesex	milton
mears	mellon	merlin	mgr	middleton	milward
mecca	mellor	merlyn	mhc	middx	milwaukee
meciar	mellors	merner	mi	midgley	mimi
medau	melly	merrick	mia	midhurst	minch
medford	melrose	merrill	miami	midi	mine
medhurst	melton	merritt	miamy	midlothian	minehead
medicare	melville	merriweather	miandad	might	minerva
medici	melvilles	mers	micah	miguel	ming
medina	melvin	mersey	mich	miguelito	minh
mediterranean	melvyn	merseyside	michael	mihir	minister
medmelton	melwa	merson	michaela	mikael	miniver
medmenham	memet	merthyr	michaelis	mike	minneapolis
medoc	memons	merton	michaelmas	mikhail	minnelli
medusa	memphis	mervin	michaelparkin	milan	minnesota
medway	mena	mervyn	michaels	milanese	minnie
mee	menai	meryl	michaelson	milanesio	minogue
meehan	mendel	merymose	micel	milburn	minos
meeland	mendelssohn	mesa	micelangelo	mildred	minosso
meenambakka	menderes	mesopotamia	michele	milenoff	minshall
m	mendes	mesopotamian	micelin	miletti	minsk
meer	mendez	messer	micelis	milford	mintel
meerut	mendip	mesure	micell	miliband	minter
meg	menelek	metaesthetic	michelle	milken	minto
megadeth	menem	metairie	Michels	millan	minton

minus	moaville	mollie	montego	morgart	mounce
mir	mobil	molloy	montenegro	mori	mouncy
mira	mobuto	molly	monteverdi	moriarty	moune
miramar	mobutu	molotov	montevideo	moritz	mountbatten
miranda	moby	molton	montfort	morland	mountjoy
mirawdeli	mochan	molvar	montgomerie	morley	mountmckinle
mirgund	modal	molvedo	montgomery	morningside	y
miriam	modi	molyneux	montgomerys	mornington	mourne
miriwa	modigliani	moma	hire	morocco	moussa
miro	modofier	mombasa	monti	morpeth	mowbray
miroir	modrow	mona	montparnasse	morpurgo	mowlam
mirren	moe	monaco	montpelier	morrell	moxon
misa	moffat	monaghan	montpellier	morris	moy
mishra	moffett	monarda	montreal	morrison	moya
misra	mogadishu	mond	montrose	morrissey	moyglare
mississippi	mogg	mondale	montserrat	morse	moynihan
missoula	moggs	monday	monty	mortimer	mozart
missouri	mohallas	monet	monza	mortlake	mp
mit	mohamed	monghyr	moodie	morton	mprp
mitch	mohammed	mongolia	mooney	moscone	mrc
mitcham	mohan	monica	moorclose	moscovici	mridha
mittchell	mohican	monie	moore	moscow	mrp
mittchells	mohididi	monika	moores	moseley	mrs
mittchison	mohinder	monique	moorgate	moses	mss
mittchum	mohsina	monkhouse	moorish	moshe	mtv
mitteff	mohun	monklands	moorlake	mosher	muawad
mitford	moin	monkton	mor	mosley	mubarak
miti	moines	monmouth	morag	mossad	much
mitra	moira	monmouthshir	moran	mosse	mudd
mitsotakis	moise	e	morarji	mossley	mudgal
mitsubishi	mojor	monro	moravia	mossop	mudiger
mittal	moki	monroe	morch	mossreeba	mudigere
mitterand	molby	monrovia	morden	most	mueller
mitterrand	moldavia	mons	mordet	mothercare	mufasa
mixner	moldova	monsanto	more	motherwell	mugabe
miyake	molensingel	mont	moreau	motorola	muggeridge
miyazawa	molesworth	montagu	morecambe	motown	muir
mizell	molinari	montague	morehouse	mott	muirhead
mk	molineux	montaigne	moreland	motte	mukherjea
mkm	moliz	montaine	morenz	motunui	mukherjee
mnp	molla	montana	moreton	moudy	mukhitiar
mnr	molland	monte	morey	moulton	mukhtiar
moana	moller	montefiore	morgan	moultons	muldoon

mulholland	musgrave	nadu	naphill	navarre	nelly
mullen	musgrove	nagar	napier	navrang	nelson
muller	mushtaq	nagarwala	naples	navratilova	nenagh
mulligan	musica	nagasaki	napoleon	nawab	nene
mullin	muskegon	nagayalanka	napoleonic	nawaz	neneh
mullins	musselburgh	nagel	napper	nawroji	nenna
mulroney	mussolini	nagle	narain	naxalbari	nep
multhrop	musson	nagorcem	narasimha	nayar	neptune
mulvaney	must	nagorny	narayan	naylor	nerc
mundi	mustafa	nagy	narayanan	nazareth	nero
mundy	mustaine	nahal	narayanaswam	nbc	nesbit
munger	mustapha	nahum	i	nc	nesha
mungo	mustn	nai	narbigha	ncb	ness
muni	muthu	naickenthope	naresh	ncc	nessan
munich	muti	naidu	narita	ncm	nessie
munn	mutney	naik	narragansett	nco	nessun
munoz	mvs	naipaul	narula	ncr	nestl
munro	my	nairn	nas	nct	netaji
munster	myanma	nairne	nasa	nd	netherhall
munton	mybug	nairobi	naseby	nder	netherlands
munuswamy	mybugs	najeeb	nash	ndp	netherton
murad	myc	najibullah	nashville	neal	nethery
murchison	mycenae	najumuddin	nasik	neale	netrahin
murdo	myeloski	nakoma	nasmyth	near	nettie
murdoch	myers	nala	nasriganj	neasham	neuberger
murdock	myle	nalgo	nassau	neath	neuchatel
murf	myles	nall	nasser	nebraska	neue
murgatroyd	myra	nam	nassirya	nec	neumann
muridian	myron	namakkal	nat	ned	neusteter
muriel	myrtle	nambudiripad	nata	nederland	neusteters
murilo	myself	namesnik	natal	needham	nev
murless	n	nanaji	natalia	neelam	nevada
murli	nabil	nanak	natalie	nehf	nevil
urmansk	nabisco	nancy	natasha	nehru	neville
murphy	nacab	nand	nath	neil	nevin
murray	nachman	nandan	natham	neill	nevis
murrayfield	nadel	nanette	nathan	neilson	newall
murtach	nadell	nanking	nathaniel	neiman	newark
murtaugh	nader	nanny	nato	neither	newbold
murville	nadia	nanook	natwest	nejm	newbridge
musa	nadine	nantes	nauhn	nell	newburger
muscat	nadirpur	nantwich	nauru	nellie	newbury
muscovy	nadkarni	naomi	navajo	nellore	newby

newcastle	ni	nike	nizan	norridge	noyce
newco	niagara	nikhil	nkrumah	norris	noyes
newcombe	niall	niki	nkumbula	norristown	nozick
newcomen	niarchos	nikita	nl	norsk	np
newell	nic	nikkei	nld	northallerton	npc
newent	nicandra	nikki	nlp	northam	npfl
newfoundland	nichol	nikolai	nme	northampton	npt
newgate	nichola	nikos	nmr	northamptons	npv
newham	nicholas	nilakantan	nn	hire	nra
newhaven	nicholl	nile	no	northants	nrc
newington	nicholls	nilgiris	noah	northbrook	ns
newland	nichols	nilp	noakes	northgate	nsaid
newlands	nicholson	nilratan	nobody	northland	nsc
newley	nickau	nils	nobuyoshi	northumberland	nsf
newman	nicklaus	nilsen	noddy	d	nspcc
newmarket	nickleby	nilsson	noel	northumbria	nuadu
newnham	nicky	nimmo	nofziger	northumbrian	nucci
newport	nico	nina	nogai	northwich	nuffield
newquay	nicodemus	ninagawa	nolan	norton	nugent
newry	nicol	nine	nominative	norway	num
newsom	nicola	nineteen	nomura	norwegian	nuneaton
newsome	nicolae	nineteenth	non	norwich	nunes
newson	nicolas	ninetieth	nona	norwood	nungambakka
newsweek	nicole	ninety	none	nostradamus	m
newton	nicolette	ninian	noonan	not	nunn
newtonian	nicolo	nintendo	noone	nothing	nunthorpe
newtown	nicolson	ninth	noor	notre	nupe
newtownabbey	nicosia	nippon	noorda	notting	nuremberg
y	nicra	nirankari	nor	nottingham	nureyev
newtownards	nielsen	nirankaris	nora	nottinghamshi	nutall
newyork	nieman	nirmala	norah	re	nutting
newzealand	niemi	niro	norbert	notts	nv
next	nietzsche	nisbet	norbreck	notwithstanding	nvocc
nez	nieuil	nischwitz	nordern	g	ny
nfl	nigam	nisodemus	nordin	noun	nyberg
nfs	nigel	nissan	nordmann	nour	nyerere
nfu	niger	nissen	noreen	nova	nz
ngaio	nigeria	nithard	norell	novak	nzrfu
ngc	nigerian	nithsdale	norfolk	novell	o
ngo	nightingall	niue	noriega	november	oa
nguyen	nihore	niven	norma	novosibirsk	oakes
nh	nijinsky	nixdorf	norman	novum	oakeshott
nhs	nikau	nixon	normandy	nowak	oakland

oakley	oglala	omi	orme	otello	paan
oakwood	oglesby	omsk	ormesby	othello	pablo
oamaru	oglethorpe	on	ormond	other	pabor
oas	ognall	onassis	ormsby	others	pac
oates	oh	once	ormskirk	oti	pacey
oau	ohaewai	one	oro	otis	pachi
ob	ohakune	onie	oromo	otley	pacino
obadiah	ohio	ono	orpheus	ott	packard
obaidullah	ois	onslow	orr	ottawa	packham
oban	okhla	ontario	orrell	ottery	packington
oberwetter	okinawa	onto	orsay	otto	packwood
obispal	okla	oona	orsino	ouch	paddington
occam	oklahoma	ootacamund	orson	ouen	paddison
ocker	olaf	opcs	ortega	ought	padgett
ockleton	olazabal	opec	ortiz	oughtn	padrauna
ocs	oldfield	opelika	orton	ould	padua
oct	oldham	openshaw	orvilleredenba	oulton	pagemaker
octavia	olds	openvms	cher	oum	paget
october	oleg	ophelia	orwell	oup	pagham
oda	olerichs	opie	os	our	pahl
odd	olga	opotiki	osaka	ours	paige
odell	olinger	oppenheim	osbert	ourselves	paignton
odessa	oliveira	oppenheimer	osborn	ouse	paine
odette	oliver	oprah	osborne	out	painswick
odhar	olivet	or	oscar	over	paisa
odiham	olivetti	oran	oschersleben	overton	paise
odowd	olivia	orc	osf	ovett	paisley
oe	olivier	orcs	osiris	ovid	pak
oecd	ollie	orcutt	oskar	owain	pakistan
oed	olney	oregan	oskarmorgenst	owen	pakistani
oedipus	olsen	oregon	ern	owens	palatine
oem	olson	orelhershiser	oslo	oxbridge	palekar
of	olveston	orestes	osman	oxfam	palermo
ofahengaue	olvey	orford	ossetia	oxford	paley
off	olwyn	oriel	ossie	oxfordshire	pali
offa	olympia	orinoco	ostend	oxley	palin
offerman	olympus	orion	osvaldo	oxnard	palit
offiah	omagh	orkay	oswald	oxon	pallava
oftel	omally	orkney	oswestry	ozal	pallister
ogden	oman	orkneys	oswiu	ozawa	palma
oggsford	omar	orlando	otago	ozzie	palmberg
ogilvie	omer	orleans	otaki	p	palmer
ogilvy	omf	orly	otani	pa	palmerston

palo	parikh	patel	paxman	pender	peria
palumbo	parimutuels	paterson	paxton	pendle	pericles
pam	paris	pathet	paye	pendlebury	perkin
pamela	parishad	patiala	payne	penelope	perkins
pampa	parke	patil	payton	peng	perlman
pampore	parker	patliputra	paz	penge	perloff
panasonic	parkes	patna	pbs	penh	perloffs
panch	parkhead	patnaik	pc	penistone	permian
panchayat	parkin	patnick	pcc	penn	perot
panchayats	parkinson	paton	pcf	pennethorne	perrin
pancras	parkway	patrese	pcp	penney	perry
pande	parma	patrice	pcr	penniman	persephone
pandey	parmar	patricia	pcx	pennine	persil
pandora	parmer	patricio	pd	pennines	perth
pandurang	parminder	patrick	pdc	pennington	perthshire
pandya	parnell	patsy	pds	pennock	pertwee
panja	parr	patten	peabody	pennsylvania	perugia
pankhurst	parretti	patterson	peacheater	penrith	pesce
pankowski	parrillo	patti	peake	penrose	peshawar
panmure	parrott	pattie	pearce	penruddock	pete
panozzo	parsis	patton	pearse	penry	peter
pantas	parti	patty	pearson	pensacola	peterborough
panthis	partick	patusan	pease	penta	peterhead
paolo	particular	patwardhan	peavey	pentecost	peterlee
papandreou	partido	patwari	peckham	pentium	peters
papatoetoe	parton	pau	peddie	pentland	petersburg
papatotara	parvis	paul	pedro	pentonville	petersen
pappas	pas	paula	peebles	penzance	petersfield
papua	pasa	paulah	pegasus	pepe	peterson
papworth	pasadena	paulette	pegg	pepin	petipa
paquet	pascal	pauleys	peggy	pepsi	petit
paragraph	paschal	paulie	peking	pepys	peto
paraguay	paschall	pauline	pele	pequena	petone
parangipettai	pascoe	paulmann	pelham	per	petr
parashar	pasok	paulo	pelosi	perasso	petra
parc	passive	pauls	peltason	perce	petre
pardao	pasteur	pavarotti	peltzer	percival	petri
paree	paston	pavel	pemberton	percy	petrini
pareto	pasupati	pavey	pembroke	perdita	petrograd
parfitt	paswan	pavletich	pembrokeshire	pereira	petrouchka
parfois	pata	pavlov	penang	peres	petrovic
pargeter	patagonia	pawar	penarth	perez	pettigrew
parichy	pataudi	pawtucket	penda	pergamon	petula

petworth	phyfe	pinkerton	plo	pompidou	poseidon
peugeot	phylcia	pinkertons	plowden	ponderevo	posi
pevensey	phyllis	pinner	pltoff	pondicherry	posidonius
peyton	pia	pinochet	plummer	pons	poste
pezza	piaget	pinsk	plumpton	ponsonby	postlethwaite
pfa	piaroa	pintall	plus	pont	potsdam
pfaff	picardy	pinter	plutarch	ponte	potts
pfeiffer	picasso	pinzon	pluto	pontefract	pou
pfk	piccadilly	pioneerland	plymouth	pontiac	poulantzas
pflp	pickard	piotr	pmr	pontin	poulenc
pg	pickering	pipkin	pnc	pontius	poulter
pga	pickering	pippa	pnp	ponty	poussin
pgc	pickford	pirandello	pocock	pontypool	poutney
pge	pickfords	piraro	podmore	pontypridd	povich
ph	pickwick	pirie	poe	pooh	powell
phagwara	picton	pirmohani	poges	poole	powelliphanta
pharlap	pidgeon	pisa	pohl	pooley	powergen
phelan	piechanov	pissarro	pohly	poon	powerpc
helps	piedmont	pitlochry	poindexter	poona	powis
phena	piercy	pitman	poirot	popov	powys
phiala	piero	pitt	poisson	popper	ppd
phil	pierre	pitts	poitier	poppins	ppg
philadelphia	piersee	pittsboro	poitiers	popsi	ppl
philby	pierson	pittsburg	poitou	porfirio	ppp
philip	pieter	pittsburgh	pol	porirua	prabandhak
philippa	pietro	pius	poland	porli	prabhandak
philippe	piggot	piwkowski	polanski	porritt	prabhas
philips	piggott	pj	polaris	porsche	prabhudas
phillies	pikhoia	pkk	polisario	porsches	pradeep
phillip	pilate	pl	pollack	porta	pradesh
phillipa	pilger	pla	pollak	portadown	prado
phillips	pilkington	placido	pollard	porte	prague
phillis	pillay	plan	pollitt	porterfield	prakash
philmont	pillsbury	planck	pollock	porth	pranab
philpott	pilton	plantagenet	polly	porthmadog	prasad
phipp	pimlico	plagwyn	polswett	portia	prasanna
phnom	pimsleur	plaskett	polybius	portillo	prasanta
phoebe	pinar	plato	pomare	portland	prashanta
pholsena	pindar	platt	pomerania	portman	prasun
phouma	pinder	platts	pomiane	porto	pratap
phrase	pinero	plessey	pompano	portobello	pratapgarh
phu	pinewood	plewman	pompeii	portrush	pratapsing
phulbagan	pingel	pliny	pompey	portsmouth	pratt

prattville	priyadarshi	pujari	qatar	rab	raggett
pravda	pro	pullar	qb	rabat	raghu
prd	probyn	pullen	qbd	rabb	raghubir
prebble	procter	puller	qdm	rabelais	raghunath
precambrian	profili	pumbaa	qian	rabhu	raghuram
prechter	profumo	pumfrey	qichen	rabia	raghuramaiah
predeterminer	progressive	pumilia	qing	rabin	raglan
predicative	prokofieff	punchestown	qpr	rabindra	raheem
prefix	prokofiev	punjab	qt	rac	rahman
premadasa	pronoun	punjabi	quantock	racal	rahong
preminger	prospero	punjabis	quap	rachael	rai
premjit	prosser	punta	quattro	rachaela	raine
prentice	prost	purabiya	quax	rachel	rainer
preobrazhensk	protagoras	purcell	quayle	rachmaninov	rainey
y	protano	purdue	quebec	racine	raith
preposition	protherough	purley	queenie	rada	raitt
prepositional	proudfoot	purna	queensberry	radcliffe	raj
prescot	proust	purnell	queensferry	rader	raja
prescott	provence	purushottam	queensgate	radford	rajagopalachar
presidio	providencia	purushottamda	queensland	radhakrishnan	i
presley	prs	s	queensway	radix	rajaji
prestatyn	pru	purvis	quentin	radja	rajan
prestel	prudhoe	pushkin	quernmore	radley	rajasekhar
presthus	prue	pushpa	quex	radnor	rajashekaran
preston	prunella	pushtoonistan	quigley	radulfus	rajasthan
prestridge	prynne	putaruru	quincey	rae	rajat
prestwick	ps	putnam	quincy	raelene	raje
prestwood	psc	putney	quinim	raesz	rajendra
pretoria	psd	puttiah	quinlivan	raewyn	rajinder
pri	psion	pvc	quinn	raf	rajiv
priddy	psoe	pw	quint	rafael	rajput
priestley	pta	pwillheli	quinta	rafaelo	rajputs
prieur	pte	pye	quintal	rafel	rajya
princesse	ptolemy	pyke	quintin	rafer	rakesh
princeton	ptt	pym	quinton	rafferty	rakovsky
princetonuniv	pualani	pyongyang	quiss	rafi	raleigh
ersity	puddephat	pyreanean	quito	rafiki	ralemberg
pringle	puddingstone	pyrenees	quixote	rafiq	ralf
prinzenpalais	pudsey	pyy	quorn	rafiullah	ralph
priscilla	puerto	qaddafi	r	rafsanjani	ralston
pritam	pugh	qaddisiya	ra	rafta	ramachandra
pritchard	pugliese	qaly	ra0	ragan	ramachandran
priti	pugwash	qasim	raasay	ragd	ramadan

raman	ranulph	raymont	reese	rennie	ri
ramanathan	rao	rayner	reeve	reno	rialto
ramanathapura	raoul	raynes	reeves	renoir	riba
m	raper	raynham	reflexive	renouf	ribble
ramaswamy	raphael	raynor	reg	renshaw	ribena
ramblas	rappe	rb	regan	rensselaer	ribot
rambo	raquel	rbg	reggae	rentokil	rica
ramboat	rarotonga	rbge	reggie	renton	rican
ramchandra	rarotongan	rc	regina	renvoize	ricardo
ramesh	rasbora	rca	reginald	renwick	ricci
ramis	rashaan	rcc	regis	repton	richard
ramjichak	rashad	rcd	regnery	restigouche	richards
ramjilal	rashaud	rcn	rehnquist	retford	richardson
rammanohar	rashid	rds	reich	reuben	richardsons
ramnad	rashrapati	re	reichenberg	reuter	richelieu
ramnandan	rashtrapati	rea	reichmann	reuters	richey
ramon	raskolnikov	reagan	reichstag	rex	richie
ramos	rasmussen	reama	reid	rey	richmann
ramprakash	ratagan	rebecca	reidy	reyburn	richmond
rampton	ratcliff	redbridge	reifenrath	reyes	richmondshire
ramsay	ratcliffe	redburn	reigate	reykjavik	richpal
ramsden	rathbone	redcar	reilly	reynard	richter
ramsey	ratik	redcliffe	reily	reynolds	rick
ramsgate	ratner	redding	reims	rez	rickaby
rana	ratners	redditch	reina	reza	rickards
ranald	ratto	redenbacher	reinders	rf	rickenbacker
rance	rattray	redfern	reine	rfl	rickenbaugh
ranchi	raul	redford	reiner	rfu	rickerscote
randall	ravel	redgrave	reinhard	rhee	rickey
randerson	ravenna	redknapp	reinhold	rheim	rickie
randhawa	ravenscraig	redland	reinisch	rheims	rickman
randolph	ravi	redman	reith	rhema	rickmanswort
ranfurly	ravindra	redmond	rembrandt	rhine	h
rangaswami	ravioli	redpath	remington	rhineland	ricky
rangoon	rawat	redruth	ren	rhoda	rico
ranihan	rawlings	redskins	renault	rhode	riddell
ranji	rawls	redwood	rendell	rhodes	ridder
ranjit	raxaul	reebok	rene	rhodesia	riddick
rankin	raybestos	reece	renee	rhondda	rideout
rannoch	rayburn	reeder	renfe	rhone	ridgway
ransome	rayleigh	reedville	renfrew	rhs	ridley
rantzen	raymond	reepham	renne	rhyl	rifkind
ranulf	raymondville	rees	rennes	rhy's	riga

rigby	robbie	rodriguez	ronni	rossiter	roxy
rigg	robbins	roebuck	ronnie	rosso	roy
rigoletto	robby	roedean	ronny	rostagno	royce
riley	robemaker	roehampton	rono	rostagnos	royces
rilla	robert	roethke	ronson	rostock	royle
rimell	roberta	rogan	rooney	rostov	royston
rimini	roberto	roger	roop	rostow	roz
rimmer	roberts	rogerick	roos	rosyth	rozanov
rimswell	robertson	rogers	roosevelt	rotelli	rozelle
rincewind	robertsons	roget	roper	roth	rp
ringo	robey	roh	roquemore	rotherham	rpc
rintoul	robina	rohal	rory	rotherhithe	rpf
rio	robinson	rohan	ros	rothermere	rpr
rios	robinsonville	rohmer	rosalba	rothesay	rs
ripa	robson	rohtak	rosalind	rothmans	rsa
ripley	robyn	roirbak	rosamund	rothschild	rsc
ripon	rocard	roker	rosario	rothwell	rsfsr
ripon	rocastle	roland	rosay	rotorua	rspb
risborough	rocester	rolf	rosburg	rotowaro	rte
risc	rochdale	rolfe	roscoe	rotterdam	rtr
risley	roche	rolle	rose	rouen	ruari
rita	rochefort	rollins	roseanne	roukema	ruben
ritchie	rochelle	rollison	roseberry	round	rubens
ritschl	rochester	rollo	rosebery	rourke	rubin
ritz	rochfort	rolls	rosehill	rous	rubinfien
riva	rockall	rolnick	rosella	rousseau	rubinstein
rivera	rockefeller	roma	roseman	roussel	rubirosa
riviera	rockford	romaine	rosen	routledge	ruc
rivoli	rockhall	romano	rosenberg	roux	rudd
rix	rockies	romanov	rosendal	rover	ruddock
riyadh	rockingham	rome	rosenthal	rowallan	rudge
rizzo	rockton	romeo	rosetta	rowan	rudi
rizzuto	rockwell	romero	rosheen	rowbotham	rudnev
rm	rodd	romford	rosie	rowbottom	rudolf
rmi	roddick	rommel	roskill	rowdies	rudolph
rn	roddie	romney	roskomarkhiv	rowe	rudyard
rnli	roddy	romsey	rospa	rowland	ruffin
ro	roderick	romulus	ross	rowlands	ruffollo
roald	rodgers	ron	rossellini	rowley	ruffolo
robards	rodin	rona	rossendale	rowntree	rufus
robathan	rodney	ronald	rossetti	roxborough	rugg
robb	rodo	ronaldsay	rossi	roxburgh	ruggiero
robbe	rodrigo	ronan	rossini	roxie	

ruhr	ryley	said	salvor	sandown	sardinian
ruislip	rylie	saidabad	salzburg	sandra	sarees
ruiz	ryne	saigon	sam	sandringham	sarella
rumbelows	ryuichi	sainsbury	samachar	sandro	sargent
rumbold	ryzhkov	sainsburys	samagam	sandwell	sarin
rumford	s	saitama	samana	sandys	sark
rumsfeld	saa	sajjanlal	samant	sangatpuri	sarkar
runcie	saab	sakamoto	samantha	sanger	sarkees
runcorn	saad	sakata	samar	sangh	sarmi
rupee	saaid	sakharov	samara	sangharsh	sarnies
rupert	saanich	sakthi	samarang	sanglalore	sarojini
rushdie	saanjh	sal	samaresh	sangli	sarpanch
rushton	saatchi	salaam	samarkand	sangster	sarpanchas
ruskin	sab	salah	samarra	sanjay	sartre
russ	saba	salamanca	samastipur	sanjiva	sarum
rusa	sabah	salamandre	sambhu	sankey	sarunas
russel	sabarmati	salamuddin	samedow	sanson	sarup
russell	sabatini	saledo	samiti	sant	sas
russia	sabbarese	saleem	sammons	santa	sasha
russo	sabha	saleh	sammy	santander	sashi
russon	sabina	salem	samos	santerre	saskatchewan
rustam	sabol	salems	sampdoria	santiago	sassoon
rutgers	sabrina	salford	sampras	santis	satara
ruth	sacco	salgaonkar	sampson	santo	satchell
rutherford	sacha	salim	sampurnanand	santos	sathyamangala
ruthven	sachs	salinas	sams	santosh	m
rutland	sackville	salinger	samson	sanvordem	satiabhama
rutley	sacramento	salisbury	samsung	sanyo	satta
rutskoi	sadashivam	salle	samuel	sao	satterfield
rutter	sadat	sally	samuels	sapio	saturday
ruud	sadatullah	salman	samuelson	sapt	saturn
rw	saddam	salmond	samurai	sar	satya
rwanda	sade	salome	san	sara	satyagraha
rx	sadie	salomon	sanantonio	sarabi	satyajit
ry	sadler	saltash	sanarpatti	sarah	sau
ryan	sadnana	saltburn	sanathnagar	sarai	sauganash
rybkina	safeway	salter	sanchez	sarajevo	saul
ryde	safire	salterbeck	sandeman	sarasota	saunders
ryder	safrane	saltonstall	sanderson	sarawak	saundra
ryedale	sagittarian	saltpetre	sandford	sarbonis	saurashtra
ryker	sahni	salvador	sandhurst	sardanes	saussure
ryland	sahnoun	salvatore	sandie	sardar	sauvignon
rylands	sahrah	salvesen	sandison	sardinia	savile

saville	schlegel	scipio	sean	seligman	serps
savona	schleiermache	sco	searle	selina	serra
savoy	r	scobie	searles	selkirk	serrano
sawhney	schlesinger	scofield	seascale	sellafield	seth
sawyer	schleswig	scolatti	seato	sellars	sethi
saxe	schley	scopor	seaton	selwyn	sethna
saxony	schlichting	scorsese	seattle	sem	sethness
saxton	schmalzried	scorton	seb	sema	sethumadhava
saye	schmeichel	scot	sebastian	semele	n
sayer	schmidl	scotia	sebhat	semenov	seton
sayers	schmidt	scotland	second	seminole	seul
sayyaiddin	schmitt	scotrail	secord	semmes	seurat
sb	schnackenberg	scott	secunderabad	semple	sevakram
sbus	schneider	scottish	securicor	semtex	sevashram
scala	schoenberg	scotts	securitate	sen	seve
scalfaro	schoenman	scottsdale	sed	sena	seven
scalzi	schofield	scotty	seddon	senapati	sevenoaks
scanlan	sholes	scotvec	sedeno	sendei	seventeen
scapa	schopenhauer	scowcroft	sedgefield	sendero	seventeenth
scarabae	schramm	scrapiron	sedgwick	seneca	seventh
scarborough	schrempf	screvane	seeberg	senegal	seventieth
scargill	schroders	scripps	seebohm	senese	seventy
scaria	schrunk	scrymgeour	seekonk	seng	several
scarlett	schubert	scudamore	seelampur	senghor	severn
scarman	schultz	scudder	seerey	sengupta	severna
scarsdale	schumacher	scully	sefton	senna	seville
scarthi	schuman	scunthorpe	sega	sentence	sewa
scathach	schumann	scwarz	segal	seoul	sewell
scattareggia	schumpeter	scylla	sehra	sep	seychelles
schaap	schuster	sd	seib	september	seyid
schaefer	schuyler	sda	seidel	septimus	seymour
schaerf	schwab	sdlp	seifert	sequent	sfa
schaffer	schwartz	sdp	seine	serafin	sfezzo
scharansky	schwarz	sdr	sejm	serb	sg
schaumburg	schwarzenegg	sds	sekers	serbia	sga
schellenberg	er	se	selassie	serbs	sgml
schenk	schwarzkopf	seabrook	selborne	serc	sgr
scherer	schweitzer	seaford	selby	serebryakov	sgurr
schiftan	schwenk	seagoville	selden	serena	shaba
schiller	schweppes	seaham	seles	sergei	shabba
schillo	scicon	sealey	selfridge	sergio	shackleton
schiphol	scilly	sealink	selfridges	serova	shadayid
schlafly	scindia	seamus	selhurst	serpell	shadbolt

shadwell	shashi	sher	shortstop	siegfried	simonelli
shae	shashikala	sheraton	shostakovich	siemens	simons
shaffer	shastri	sherbet	shotton	sien	simpkins
shaftesbury	shatilov	sherborne	should	siena	simpson
shahid	shaughnessy	sherburn	shouldn	sierra	simpsons
shahn	shaun	sheridan	shoup	siet	sims
shahnay	shaw	sherif	shourie	sigismund	sinai
shai	shawcross	sheringham	shrek	sigmar	sinatra
shaikh	shawfield	sherlock	shreveport	sigmund	sinbad
shakhrai	shaws	sherman	shrewsbury	signe	since
shakir	shay	sherwood	shri	sigurd	sinclair
shalbug	shayne	sheth	shrimpton	sihanouk	sind
shall	she	shetland	shriver	sijua	sindhia
shambhu	shea	shetlanders	shropshire	sikandar	sinead
shamiana	shearman	shevardnadze	shukla	sikander	sinfield
shamir	sheba	shiel	shuler	sike	singapore
shamlou	sheehy	shiflett	shultz	silaev	singh
shan	sheeran	shih	shurg	silas	singhal
shand	sheff	shikar	shurvington	silesia	singleton
shane	sheffield	shildon	shuttleworth	sillett	sinha
shanghai	shehee	shilton	shyam	sillitoe	sinhala
shangkun	sheikha	shinn	shyama	silloth	sinise
shangri	sheila	shinwell	shyamal	silva	sinn
shankar	shek	shiona	si	silverado	sinton
shanker	shekar	shipley	siad	silverdale	siobhan
shankill	shekhar	shipman	sian	silverman	sioux
shannon	shelagh	shipton	sib	silvers	sir
shanti	shelburne	shirley	sibelius	silverstone	sircar
shantz	shelby	shiroda	sibson	silvia	sirett
shapiro	sheldon	shiromani	sichuan	silvio	sirimavo
sharad	sheldrake	shiv	siciliana	sim	sirith
sharan	sheldukher	shivaji	sickert	simba	sirius
shari	shelford	shivalakhpati	sid	sime	sisam
sharif	shelley	shivanandan	sidcup	simeon	sisk
sharjah	shelly	shoettle	siddeley	simla	sissay
sharkey	shelton	shogun	siddons	simm	sisson
sharky	sheneneh	shoji	sidi	simmel	sistik
sharma	shenzi	sholavaram	sidney	simmonds	sistine
sharon	sheo	sholom	sidon	simmons	sisulu
sharona	shep	shona	sidonius	simmons ville	sital
sharpe	shephard	shoreditch	sie	simms	siward
sharply0	sheppard	shoreham	siebern	simon	six
shartzler	shepperton	shorenstein	siegel	simone	sixsmith

sixteen	slocombe	snowdon	solway	southey	speyhawk
sixteenth	slocum	snowdonia	somboon	southfield	sphinx
sixth	sloper	snp	some	southgate	spicer
sixtieth	slorc	snyder	somebody	southland	spiegel
sixty	slough	soames	someone	southmead	spielberg
siyad	slovakia	soane	somerset	southport	spinks
sizewell	slt	soares	somerville	southville	spinoza
sizova	sm	sobe	somes	southwark	spiro
sj	smajkic	sobell	something	southwell	spitalfield
skeet	smallwood	sobhani	somme	southwold	spitalfields
skegness	smalpur	soc	sommes	southwood	spittals
skehan	smeaton	socola	sonapur	southworth	spitzbergen
skelland	smedley	socrates	sonia	souvanna	spla
skelton	smiley	soe	sonja	souza	spock
skerne	smith	soemmerda	sony	sovietunion	spode
skiddy	smithers	sofia	soo	sowerby	spooner
skinner	smithfield	softbench	sooze	soweto	spr
skinnergate	smithie	softpc	soper	spa	sprague
skipton	smithkline	sogeti	sopers	spahn	spremberg
skirton	smiths	soh	sophia	spahnie	springall
skoda	smithson	sohan	sophie	spain	springfield
skokic	smithsonian	soho	sophocles	spalding	springsteen
skol	smithy	sohrab	sophy	spaniard	sproule
skorich	smolensk	soilih	sopsaisana	spanton	sprs
skouras	smollett	sokal	sorbonne	sparc	spurdle
skvortsov	smp	sokol	soren	sparcstation	spurgeon
skydome	smuts	sol	sormander	sparkbrook	spurzheim
skye	smyth	solanio	sorrentino	sparta	sputnik
sl	smythe	solario	sotheby	spartak	spyros
slade	snaith	solaris	souder	spassoff	sqn
slagle	snapback	solarz	soule	spatz	squidgy
slater	snape	soldatov	souljah	spd	srebrenica
slaven	snark	solent	souness	spearh	srimati
slavery	snc	solihull	souphannouvo	spearman	sringeri
slavic	snead	solihull	ng	speckman	srinivas
slavonic	snelders	solimoes	souphanouvon	spedding	sro
slenczynka	snell	solingen	g	speke	srs
sligo	snellville	solly	sousa	spence	ssap
slimani	snh	solomon	souter	spencer	ssd
slimbridge	snm	solomons	southall	spencers	ssds
sloan	snmp	soloviev	southam	spennymoor	sseldorf
sloane	snodgrass	solowka	southampton	spenser	ssp
slobodan	snowden	solti	southend	speyer	ssr

ssrs	starcevic	stephens	stojko	strathspey	suffern
sssis	starkadder	stephenson	stok	stratocaster	suffix
sst	starkadders	stepney	stoke	stratten	suffolk
stableford	starkey	sterland	stokesley	stratton	sugden
stablemate	starr	sternberg	stokowski	stratus	suharto
stacey	stasi	sterne	stonebridge	strauss	suhas
stacy	staunton	steussie	stonehaven	stravinsky	suisse
staden	stavanger	steve	stonehenge	streatham	sujan
stadt	staveley	stevedore	stonehouse	streep	suk
stafford	stavrogin	steven	stoneleigh	street	sukarno
staffordshire	stavros	stevenage	stonestown	streeters	sukebind
stager	stayers	stevens	stoppard	streisand	sukey
staiger	stb	stevensen	stormont	stretford	sukhdev
staines	stc	stevenson	stornoway	stretton	sukhvinder
stainforth	stdy	stevie	storr	strickland	suki
stainton	steadman	stewart	storrie	strikeland	sulcer
staithes	stebbing	stewarts	storrington	strindberg	suleiman
stalin	stebbings	steyning	stortford	stringer	sullart
stalingr	stedeford	steytler	stott	stromness	sullivan
stalingrad	steele	stg	stoughton	stroud	sullom
stallard	steen	stibbe	stour	stu	sumatra
stallone	steeves	stich	stourbridge	stuart	sumerged
stamford	stefan	stickney	stoute	stuarts	sumitomo
stan	stefano	stilgoe	stowe	stubbs	summerchild
standish	steffi	still	stowell	studley	summerdale
stanfill	stein	stinson	stowey	stukeley	summerfield
stanford	steinbeck	stirling	stowmarket	stukely	sumner
stanhope	steinberg	stirlingshire	strabane	sturge	sunay
stanislav	steiner	stitt	strach	sturges	sunbury
stanislaw	stella	stobart	strachan	sturm	suncoast
stanistreet	stena	stockbridge	strachey	stuttgart	sundance
stanley	stendal	stockbroker	strafaci	stych	sundar
stanmore	stendhal	stockdale	straker	su	sundararajan
stannard	stengel	stocker	stram	subash	sunday
stannington	stennis	stockholm	stranahan	subordinate	sunderland
stansbery	stenson	stockley	strang	subordinating	sunil
stansfield	stepan	stockport	strangeways	subramaniam	sunley
stansted	stepanovich	stocksbridge	stranraer	subramanyam	sunman
stanton	stepen	stockton	strasbourg	subroto	sunningdale
stapleford	stephan	stockwell	strat	such	sunnyvale
stapleton	stephanie	stoddard	stratford	suchinda	sunselect
stapp	stephanotis	stoddart	strathallan	sudbury	sunsoft
stara	stephen	stodday	strathclyde	suez	

suntook	svenssons	syedna	takahiro	tappin	tay
sununu	sveri	sykes	takapuna	tapsell	tayal
supagudi	svidrigailov	syl	takeshita	tara	taylor
super	swadeshi	syllable	taki	tarakeshwari	taylors
supercalc	swaffer	syltone	takieddine	taranaki	tayside
superlative	swain	sylvania	takoradi	tarapur	tbilisi
supersparc	swale	sylvanus	tal	tarbitten	tc
suragai	swaledale	sylvester	talash	tardieu	tccb
surajit	swales	sylvia	talbot	tardis	tchaikovsky
surajmal	swamy	sylvie	taliesin	tareytown	tcr
surbiton	swanage	symington	taligent	tarikere	tdc
surere	swann	symonds	tallahassee	tariq	te
surinam	swansea	symons	tallendorf	tarquin	teague
suriname	swanson	syms	talley	tarquinian	teape
surinder	swapo	syne	tallis	tarrant	tearle
surjit	swaran	synonym	taluk	tarun	tebbit
surkov	swart	synoptics	tam	tarver	tec
surrey	swayne	syracuse	tamar	tarzan	tecs
surtees	swaziland	syria	tamara	tashie	ted
surve	swb	tabatinga	tamas	tashima	teddington
surya	swede	tabb	tamb	tashkent	teesdale
susainathan	sweden	tabitha	tambe	tasker	teesside
susan	swedish	tacitus	tambling	taskopruzade	teheran
susanna	sweeney	taczek	tammuz	tasman	tehran
susannah	sweezy	tadcaster	tammy	tasmania	tehsil
sushi	swegen	taff	tampa	tasmanian	tejeda
susie	swifts	taft	tamworth	tass	tekapo
sussex	swinbrook	taggart	tanaka	tata	telecaster
sutch	swinburne	tagore	tandon	tate	telekom
sutcliffe	swindon	taheb	tandy	tatham	telford
suter	swinstead	tahiti	tanfield	tatra	templecombe
sutherland	swinton	tahoe	tanganyika	tatum	templeman
sutton	swire	tahuri	tangerine	tatyana	templeton
suu	swiss	tai	tangier	taunton	ten
suva	switzerland	taif	tangmere	taupo	tencel
suzanne	sybase	taihape	tania	tauranga	tendulkar
suzi	sybert	tain	tanjore	taussig	tenerife
suzie	sybil	taipei	tannadice	tavistock	tenant
suzuka	sycorax	taishan	tanu	tawera	tennessee
suzuki	syd	tait	tanya	tawes	tennyson
suzy	sydenham	taj	tao	tawia	tense
sven	sydney	tajik	tapie	tawney	tenth
svend	syed	tajikistan	taplow	tawno	tepilit

teresa	the	thorne	tilak	tivoli	tonga
terence	the0	thorncroft	tilbury	tizard	tongan
teresa	thebes	thornhill	tilda	tl	tongariro
teri	theda	thornley	till	tlr	toni
termer	their	thornton	tilley	tm	tonkin
terra	theirs	thorold	tillingham	tnc	tonks
terre	thelma	thorp	tilly	tncs	tonnay
terrence	thelonius	thorpe	tim	tnf	tono
terri	them	thorpey	timaru	to	tony
terry	themistokles	thorstensen	timex	tobago	toohey
tes	themselves	those	timisoara	tobermory	toole
tesco	theo	though	timman	tobias	tootall
tess	theobald	thousand	timms	tobie	topanoora
tessa	theodor	thousandth	timmy	tobin	topliss
texas	theodora	threadneedle	timon	toby	topper
tetley	theodore	three	timoney	tocilescu	tora
tetra	there	threlfall	timor	tock	torbay
tevershall	theresa	thrice	timothy	tod	torborg
tewkesbury	therese	throgmorton	tims	todd	torino
tex	these	through	tina	todorov	torness
texaco	theseus	throughout	tindall	togo	toro
texas	thessaly	thu	tingralla	tokelau	toronto
tgwu	thetford	thucydides	tinley	tokoroa	torquay
th	they	thunderbird	tinsley	tokugawa	torrance
thacker	thiercelin	thunderbirds	tintern	tokyo	torres
thackeray	thierry	thurgood	tipperary	tola	torridon
thai	third	thurman	tipton	tolaga	torrington
thailand	thirsk	thurrock	tirana	tolby	torus
thakhek	thirteen	thursday	tiruchi	toledo	toscanini
thakre	thirteenth	thurso	tiruchirapaili	tolkien	tosh
thakur	thirtieth	thysen	tiruchirapalli	tollemache	toshiba
thakurs	thirty	thyssen	tiruchirappalli	tolonen	toshiki
thalia	this	tia	tischman	tolstoy	tossie
thame	thom	tiananmen	titania	tom	totnes
thames	thomas	tidelands	titch	tomas	tottenham
than	thompson	tidwell	titford	tombrokaw	totteridge
thana	thomson	tierney	titian	tomei	touche
thanet	thor	tiffany	titmuss	tomkins	tougas
thanjavur	thorburn	tigard	tito	tomlin	touggourt
thapa	thoreau	tigray	titov	tomlinson	toulon
that	thorfinn	tigre	tituencies	tommaso	toulouse
thatcher	thornaby	tigris	titus	tommy	toure
thatchers	thornbury	tikkun	tiverton	tonbridge	tovarisch

tovaritch	trevor	trumper	tunisian	tyne	umberto
towards	treyvone	trunchbull	tunney	tynemouth	un
townscroft	tri	truro	tunstall	tyneside	una
townsend	triassic	trythall	tuohy	tyree	under
townsends	trichieri	ts	tuppe	tyrell	undercliffe
townshend	trichy	tsarapkin	turandot	tyrol	underhill
towson	tricia	tsb	turgenev	tyron	underneath
toxteth	trickett	tschaikowsky	turgut	tyrone	underwood
toye	trier	tse	turin	tyrrell	unep
toynbee	triers	tshombe	turkey	tyson	unesco
toyota	trieste	tsitouris	turkish	u2	unfpa
tqm	trigg	tsongas	turkmenistan	uae	unhcr
tracey	trimble	tsu	turna	ubler	unicef
tracy	tring	tsunami	turnberry	ubs	uniforum
trafalgar	trinian	tualatin	turnbull	ucta	unilever
trafford	trinidad	tuan	turner	udal	uniplex
trager	triomphe	tuatapere	turpin	udas	unisys
traidcraft	tripathi	tuathal	turriff	udayar	unita
tran	tripati	tuc	turther	udc	unitedstates
tranmere	tripoli	tucano	tussard	udcs	univel
transitive	tripp	tucker	tussaud	udf	unixware
transkei	trippier	tuckett	tutankhamun	udr	unless
transvaal	tripura	tucson	tutilo	udumalpet	unlike
transylvania	trish	tudhope	tuttle	uefa	uno
traub	trisha	tudjman	tv	ugt	unsworth
travers	tristan	tue	tvei	uhles	until
travis	triste	tuesday	twain	uhry	unwin
travolta	tristram	tuffin	twelfth	ui	up
traxel	tritt	tufnell	twelve	uist	upadhyaya
treadwell	troilus	tuileries	twentieth	ujare	upakali
trekkin	trollhattan	tulane	twenty	ujjain	updike
trelawney	trollope	tuli	twice	uk	upjohn
trelawny	troon	tulle	twickenham	ukcc	upon
tremayne	tropez	tully	twiggy	ulaganathan	upton
trent	trotsky	tulmohan	two	ullapool	ur
trentham	trowbridge	tulsa	twoflower	ullman	ural
trenton	troy	tulu	twomey	ullswater	urdu
tressan	trudy	tumim	twyford	ulm	uri
tressell	trueman	tumin	tyburn	ulrich	urquhart
treuhand	truetype	tunbridge	tye	ulster	urry
trevelyan	trujillo	tung	tyerman	ulthuan	ursula
trevi	truman	tunis	tyler	ultramar	uruguay
trevino	trumbull	tunisia	tyndale	ulysses	urville

us	valmouth	vellayati	vg	vinnie	volker
usa	valois	vellely	vhs	vinny	volkes
usikov	vance	ven	vi	vinogradov	volkov
usk	vancouver	venable	via	vinson	volkskammer
usl	vander	venables	vic	vint	volkswagen
usm	vanessa	venedig	vicente	violette	volney
ussr	vang	venezia	vichy	virdon	vologsky
ustinov	vani	vengsarkar	vickers	virgil	voltaire
utah	vansittartism	venice	vickery	virginia	volvo
uthwaite	vanuatu	venkatarama	vicki	virgo	von
utley	var	ventura	vicky	virsaladze	vonneumann
utrecht	varani	venugopal	victoria	vishnuvardhan	voorhees
utrillo	varen	ver	victorian	a	vosges
uttar	vargas	vera	victorine	vishvanath	vosper
utterson	varik	verb	vidal	vishwanath	voss
uuuc	varinder	verbal	vidhan	vision	vowel
ux	variscan	verboort	vidu	viswanath	vox
uxbridge	varlaam	verde	vidyarthi	vito	vp
uzbek	varley	verdi	vieira	vitor	vr
uzbekistan	varonis	verdun	vieng	viv	vrankovic
vache	vasantdada	vere	vienna	vivaldi	vries
vaclav	vasco	verly	viennese	vivian	vroman
vadim	vasey	verma	vientiane	vivien	vue
vaidya	vass	vermeersch	viet	vivienne	vulcan
vaishnav	vasu	vermont	vieth	vivier	vulgan
vaishnavism	vatican	vern	vietnam	vizcaino	vw
vaiyapuri	vaughan	vernava	vieux	vl	vygotsky
vajapayee	vaughn	verney	vigo	vlade	vzv
vajpayee	vaux	vernon	vii	vladilen	waaf
val	vauxhall	vernor	viii	vladimir	wabash
valdez	vayalar	vero	vijay	vme	wabi
valencia	vaz	verona	vijaya	vms	wacc
valentin	vcr	veronica	vijayakumari	vo	wacklin
valentino	veasey	veronique	vijaykrishna	vocabulary	waddell
valerie	vecchi	verrone	viktor	vodafone	waddington
valery	vechey	versace	villiers	voe	wadhwan
valhalla	vedaranyam	versailles	villupuram	vogel	wadsworth
vali	vedesandur	verstandig	vilnius	voice	wagah
vallabhbai	veerappan	verulamium	vince	voiture	waggin
vallance	veerendra	verwoerd	vincent	vojvodina	wagha
vallard	vega	very	vinci	volga	wagner
valle	vegas	veuve	viner	volgograd	wagram
valmont	velasquez	vfm	vinkvoci	volk	wagstaff

waigel	walters	warwickshire	webber	wenceslas	whalley
waigwa	waltham	was	webbs	wendell	whangarei
waikato	walthamstow	wasan	weber	wendells	wharfe
wainfleet	walto	washburn	webster	wendy	wharton
wainwright	walton	washington	wedd	wensleydale	what
wairarapa	walworth	washinton	wedderburn	wentworth	whatever
waitaki	wanda	wasim	wedgwood	werder	wheatley
waitangi	wandsworth	wasn	wednesbury	were	whelan
waite	wang	wastell	wednesday	weren	wheldon
waitemata	wanganui	watanabe	weedon	werner	when
waitrose	wangemans	waterford	weekes	wes	whence
waiuku	wankhade	watergate	weetabix	wesker	whenever
wajir	wansley	waterhouse	wegener	wesley	where
wakefield	wanstead	waterloo	weggen	wessels	whereas
wakeham	wantage	waterman	weidenfeld	wessex	wherever
walbridge	wapping	waterstone	weil	westbourne	whessoe
walby	waqar	watford	weimar	westbrook	whetton
waldegrave	warangal	wath	weinberg	westbury	which
walden	warburg	watkin	weinberger	wester	while
waldheim	warburton	watkins	weiner	westfield	whillan
waldo	wardle	watkinson	weinstein	westgarth	whinfield
waldorf	wardour	watling	weinstock	westgate	whipple
waldron	wareham	watney	weiss	westinghouse	whiston
wales	wareing	watson	weissmuller	westkureuz	whitaker
walesa	warene	wattana	welbeck	westland	whitbread
walford	warhammer	waugh	welborn	westminster	whitby
walham	warhol	waukesha	welch	westmorland	whitchurch
walla	wari	wavebreaker	weldon	weston	whitechapel
wallace	waring	waveney	welensky	westover	whitecross
wallasey	waris	waverley	welford	westphalia	whitehall
waller	wark	wavertree	wella	westport	whitehaven
wallingford	warkworth	wayne	wellcome	westwood	whitehead
wallington	warley	wbc	weller	wetherall	whitehill
wallis	warne	wbo	welles	wetherby	whitehouse
wallsend	warner	we	wellesley	weu	whitelaw
wally	warnie	wea	wellingboroug	wexford	whiteley
walpole	warnock	weald	h	wexler	whiteside
walsall	warrington	wealden	wellsley	weybridge	whitey
walsh	warsaw	wearside	wellsville	weymouth	whitfield
walsham	warschauer	weatherby	welshpool	wfs	whitgiftian
walsingham	warton	weathercock	welton	wg	whither
walt	warwick	weatherford	welwyn	whadcoat	whitley
walter	warwicks	webb	wembley	whaddon	whitlock

whitman	wildridge	wilson	winthrop	wolseley	workshop
whitney	wiley	wilsons	wiremu	wolsey	wormwood
whittaker	wilf	wilstrop	wirral	wolski	worrell
whittingham	wilfred	wilton	wis	wolverhampto	worsley
whittington	wilfrid	wiltshire	wisbech	n	worthing
whitton	wilfridi	wim	wisby	wolverton	worthington
whitworth	wilhelm	wimbledon	wisconsin	won	wortley
who	wilhelmina	wimblehurst	wisden	wonderland	wotton
whoever	wilkes	wimblendon	wiseman	wong	would
whom	wilkie	wimpey	wishart	woodbine	wouldn
whose	wilkins	wimpole	wisley	woodbridge	wp
why	wilkinson	wincanton	wissahickon	woodchester	wpc
whyte	wilko	winchell	with	woodcock	wragby
wiborg	wilks	winchelsea	witham	woodford	wragbys
wichita	will	winchelsey	within	woodhall	wragge
wickens	willamette	winchester	withington	woodhead	wray
wickes	willard	winchmore	without	woodhouse	wrexham
wickham	willcox	windebrowe	witney	woodley	wright
wicklow	willem	windermere	witt	woodmansey	wrigley
widdicombe	willesden	windhagen	witteman	woodmere	wru
widdowson	willett	windows	wittgenstein	woodrow	wu
widgery	willey	windrush	witton	woodruffe	wulfhere
widmark	willi	windscale	witrock	woodside	wwf
widnes	william	windsor	wlr	woodstock	wyatt
wiegman	williamniskan	winfield	wm	woodville	wybourn
wielaard	en	winfrey	wmc	woodward	wyche
wigan	williams	wingate	wnaganui	woolf	wychwood
wigfalls	williamson	winget	woburn	woolley	wyck
wigg	willie	wingfield	wodehouse	woolton	wycliffe
wiggins	willingdon	wingrave	wofford	woolwich	wycombe
wight	willington	winifred	wogan	woolworth	wye
wigley	willis	winmor	woil	woolworths	wylie
wigmore	willoughby	winnebago	woking	woonasquatuc	wyllie
wigton	wills	winnetka	wokingham	ket	wyman
wilberforce	willsher	winnie	wolcott	woosnam	wymondham
wilbert	willy	winnipeg	wolfe	wootton	wyn
wilbur	wilma	winslow	wolfenden	worcester	wyndham
wilcock	wilmette	winsor	wolff	worcestershire	wynn
wilcox	wilmington	winstanley	wolfgang	wordperfect	wynne
wilcoxon	wilmot	winston	wolfshiem	wordstar	wynter
wilde	wilmslow	winterbottom	wolfson	wordsworth	wyoming
wildenstein	wilmsmeyer	winterslow	wollstonecraft	worid	wyong
wilding	wilshire	winterton	wollt	workington	wyre

wyresdale	xxxxi	yegor	yusuf	zeiss	zwilling
wyss	xxxxii	yehudi	yves	zeke	
wyvis	xxxxiii	yek	yvonne	zelda	
x	xxxxiv	yeltsin	zaanstroom	zemin	
xanthe	xxxxix	yemen	zacco	zen	
xavier	xxxxv	yeo	zach	zendejas	
xdesktop	xxxxvi	yeovil	zachary	zeneca	
xerox	xxxxvii	yeremi	zack	zeno	
xerxes	xxxxviii	yevgeny	zadak	zero	
xi	yacine	yicheng	zadek	zetland	
xiaoping	yadav	ying	zagreb	zeus	
xieng	yadavas	yitzhak	zahn	zhang	
xii	yadavs	ymca	zahra	zhao	
xiii	yahweh	yogendra	zail	zhelev	
ximenez	yakovlev	yogoslavia	zaire	zhivkov	
xinhua	yakunthpura	yokich	zajicek	zhukov	
xiv	yakuthpura	yokohama	zajizek	zia	
xix	yakutpura	yokomichi	zak	ziegler	
xv	yale	yolande	zamora	zig	
xvi	yalta	yom	zamzam	zimmer	
xvii	yamaha	yong	zander	zimmerman	
xviii	yamamoto	yongchaiyudh	zanne	zimmermann	
xx	yan	yorick	zanuck	zineb	
xxi	yang	york	zanucks	zinman	
xxii	yangtze	yorke	zanussi	zita	
xxiii	yanto	yorker	zanya	ziyang	
xxiv	yardley	yorkers	zanzibar	znf	
xxix	yarm	yorks	zappa	zoe	
xxv	yarmouth	yorkshire	zapt	zohra	
xxvi	yarrow	yoruba	zara	zola	
xxvii	yashpal	you	zarapkar	zollinger	
xxviii	yasmin	youngs	aslow	zorin	
xxx	yasser	younis	zazu	zoser	
xxxi	yassine	your	zborowski	zubchenko	
xxxii	yassir	yours	ze	zubero	
xxxiii	yates	yourself	zealand	zubkovskaya	
xxxiv	yavar	yourselves	zebek	zuckerman	
xxxix	yavlinsky	ypres	zechstein	zuleika	
xxxv	yaxlee	yts	zee	zulfiqar	
xxxvi	yazov	yuen	zeebrugge	zurawik	
xxxvii	ybreska	yukio	zeffirelli	zurcher	
xxxviii	yeats	yuri	zeinab	zurich	
xxxx	yegan	yussuf	zeising	zuwaya	

APPENDIX V List of all headwords in the seven textbooks

abacus	acute	agriculture	always	apartment	arrest
abandon	adapt	ahead	amaze	apex	arrive
abbreviate	adapted	aid	ambassador	apology	arrow
able	add	ail	ambition	apostrophe	art
abolish	addict	aim	ambitious	appal	artichoke
abroad	address	air	ambulance	apparent	article
abrupt	adequate	aircraft	amenity	appeal	artificial
absence	adjacent	airline	amnesty	appear	ash
absent	administer	airman	amount	appendix	ashamed
absolute	admiral	airplane	amuse	applause	ashore
absorb	admire	airport	anaemia	apple	ask
abstract	admit	aisle	anaesthetic	appliance	asleep
abuse	adopt	akin	analyse	apply	aspect
academy	adrenaline	alarm	analysis	appoint	aspirin
accent	adult	album	ancestor	appreciate	ass
accept	advance	alchemy	anchor	apprentice	assassinate
access	advantage	alcohol	ancient	approach	assault
accession	adventure	alert	anecdote	appropriate	assemble
accident	advertise	algebra	anger	appropriated	assert
acclaim	advice	alien	angle	approve	
accommodate	advise	align	angry	approximate	assess
accompany	aerobics	alike	animal	apricot	asset
accord	aerosol	alive	ankle	aquatic	assiduous
according	aerospace	alkali	annotate	arc	assign
account	affair	alley	announce	arch	assimilate
accountant	affect	alliterate	annoy	archaeology	assist
accrue	affection	allocate	annual	archaic	associate
accumulate	affinity	allow	anode	archer	assonant
accurate	affirmative	alloy	anonymous	archetype	assume
accuse	afflict	almanac	answer	archipelago	assure
ache	afford	almighty	ant	architect	asteroid
achieve	afoot	almost	antecedent	archive	asthma
acid	afraid	alone	anthem	area	astrolabe
acknowledge	afternoon	alongside	antibiotic	argon	astrology
acquaint	again	aloud	anticipate	argue	astronaut
acquire	age	alphabet	antioxidant	argumentative	astronomy
acre	agenda	already	antique	arid	astrophysics
acrobat	agent	alright	antiquity	arise	athlete
acrostic	aggresses	alter	antonym	arithmetic	atmosphere
acrylic	aggressive	alternate	anxiety	arm	atom
act	agile	alternative	anxious	army	attach
active	agony	altogether	apart	arouse	attack
actual	agree	aluminium	apartheid	arrange	attempt

attend	bag	bay	billiard	bolshy	breathe
attentive	bagpipe	beach	billionaire	bomb	breathtaking
attic	bail	bead	bin	bond	breed
attitude	bake	beak	bind	bone	breeze
attract	balance	beam	biodiversity	bonus	brew
au	bald	bean	biography	book	bribe
audience	ball	bear	biological	bookcase	brick
audiovisual	balloon	beard	biology	bookmark	bridge
audit	ballpoint	beast	biomass	bookshelf	brief
audition	baltic	beat	bird	bookshop	brigade
aunt	bamboo	beauty	birth	boost	bright
aural	ban	become	birthmark	boot	brim
author	banana	bed	biscuit	border	bring
authorise	band	bedside	bit	boring	broad
authority	bandage	bee	bite	borrow	broadcast
autobiography	bang	beef	bitter	bosom	brochure
autograph	bank	beep	bitumen	boss	bronchitis
automate	banknote	beforehand	bizarre	botany	brook
automatic	bar	beg	black	bother	broth
automobile	barbecue	begin	blacksmith	bottle	brother
autonomy	barber	behave	blame	bottom	brotherhood
autumn	bare	behaviour	blank	bough	brown
available	barge	belief	bless	bounds	browse
avenge	bark	believe	blind	bow	brush
avenue	barley	bell	blizzard	bowl	bubble
average	barn	belly	block	box	bucket
avert	baron	belong	blockhead	boy	bud
aviation	barrel	beloved	blonde	boycott	budget
avoid	barren	belt	blood	boyfriend	buffalo
await	barter	bend	blouse	brace	build
awake	basalt	benefit	blow	bracket	bulb
award	base	berry	blue	brain	bully
aware	baseball	berth	bluebell	brake	bum
awful	bashful	best	blurb	branch	buoy
axe	basin	bet	board	brand	bureau
axis	basis	betray	boast	brave	bureaucracy
baby	basket	beware	boat	brazen	burn
bachelor	bat	bible	bock	bread	bury
backache	bath	bicycle	body	break	bus
background	bathe	big	bog	breakdown	bush
bacon	battery	bike	boil	breakfast	business
bacteriology	battle	bill	bold	bream	busy
bad	battlefield	billboard	bolivia	breath	butter

butterfly	captain	causes	chat	church	clue
button	caption	caution	chauffeur	cigarette	clumsy
buy	capture	cave	chauvinist	cinema	cluster
bygone	car	caveman	cheap	circle	coach
bytes	carbohydrate	ceiling	cheat	circulate	coal
cab	carbon	celebrate	check	circumference	coast
cabbage	card	celebrity	checklist	circumstance	coat
cable	cardboard	celestial	cheek	circus	coca
cache	cardinal	cell	cheekbone	cite	cock
cafe	cardiology	cellar	cheer	citizen	cocktail
cafeteria	cardiovascular	cement	cheese	city	cocoa
cage	care	cemetery	cheetah	civil	code
cake	career	census	chef	civilian	coffee
calcium	caretaker	cent	chemical	civilise	coherence
calculate	cargo	centigrade	chemist	claim	coherent
calculator	caricature	centimetre	chemistry	clap	coil
calendar	carpet	centre	cheque	clarify	coin
calf	carrot	century	cherry	class	coincide
caliph	carry	cereal	chess	classic	cola
call	cart	ceremony	chest	classify	cold
calm	carton	certificate	chew	classmate	colic
calorie	cartoon	cfc	chicken	clean	colitis
camel	carve	chain	chickpea	clear	collaborate
camera	case	chair	chief	clergy	collapse
camp	cash	chairman	child	clerk	collate
campaign	casserole	chairperson	chilli	clever	colleague
campus	cassette	chalk	chimney	click	collect
canal	cast	challenge	chip	client	college
cancel	caster	chamber	chivvy	climate	collide
cancer	castle	champion	chloride	climb	collie
candidate	casualty	chance	chlorophyll	cling	collier
candy	cat	chandler	chocolate	clinic	collision
cane	catalogue	change	choice	clip	collocate
canny	catastrophe	channel	choke	clock	colloquial
canoe	catch	chapter	cholera	clog	colonel
canon	categorical	character	cholesterol	clone	colony
canteen	category	charge	choose	close	colossal
cap	caterpillar	charitable	chop	cloth	colossus
capable	cathedral	charity	chore	clothe	colour
capacity	cathode	charm	christmas	cloud	column
capita	cattle	chart	chronic	clove	columnist
capital	cauliflower	charter	chronology	clown	combat
capitalise	cause	chase	chuck	club	combine

come	compress	consequent	convert	country	crowd
comeback	comprise	conservative	convey	countryman	crown
comedian	compromise	conserve	convict	countryside	cruel
comedy	compulsory	consider	convince	couple	crusade
comet	compute	considerable	cook	courage	cry
comfort	con	considerate	cookie	courgette	crystal
comic	concentrate	consign	cool	course	cube
command	concern	consist	cooperate	court	cubism
commend	concert	consistent	coordinate	cousin	cucumber
comment	concerto	console	cope	cover	cue
commerce	concession	consolidate	copper	cow	cultivate
commission	concise	conspire	coppice	coward	culture
commit	conclude	constant	coptic	cowboy	cup
committee	concrete	constitute	copy	cradle	cupboard
common	condemn	constitution	copybook	craft	cure
commonwealth	condense	constrain	copyright	cram	curie
communicate	condition	construct	coral	crash	curious
communism	condole	consul	cordial	crater	curl
communist	conduct	consult	core	crawl	currency
community	conduit	consume	coriander	crazy	current
compact	cone	contact	cork	cream	curricular
companion	coney	contain	corn	create	curriculum
company	confer	contaminant	corner	creature	curry
compare	confidence	contaminate	cornflake	credible	curse
compass	confidential	content	corporate	credit	cursor
compassion	confine	contest	correct	crescent	curtain
compel	confirm	context	correspond	crew	curve
compensate	conflict	continent	correspondent	cricket	custom
compete	confront	contingency	corrugate	crime	customer
competition	confuse	contingent	corrupt	criminal	cut
compile	congenial	continue	cortisol	cripple	cycle
complain	congested	contract	cosmetic	crisis	cylindrical
complaints	congratulate	contraction	cost	critic	dad
complete	congress	contradict	costume	criticise	daffodil
complex	connect	contrary	cot	criticism	dairy
complexion	conquer	contrast	cottage	crocodile	daisy
compliment	conquest	contribute	cotton	croissant	dam
comply	conquistador	control	cough	crop	damage
compose	conscientious	controversy	council	cross	dance
compound	conscious	convene	counsel	crossroads	danger
comprehend	consensus	convenience	count	crossword	dare
comprehensiv	consent	convention	counter	crouch	dark
e	consequence	converse	counterfeit	crow	darling

dart	defy	destroy	diphthong	distinguish	drake
data	degrade	detail	diploma	distract	drama
date	degree	detect	direct	distribute	drastic
daughter	dei	detergent	direction	district	draughtsman
dawn	deity	deteriorate	directory	disturb	draw
day	delay	determine	dirty	dive	drawback
dead	delicious	detest	disabled	diverse	dread
deaf	delight	devastate	disadvantage	diversion	dreadful
deal	deliver	develop	disappear	divide	dream
dean	delta	device	disappoint	dividend	dreary
dear	demand	devise	disapprove	divine	dress
death	democracy	devote	disaster	divorce	drift
debate	demoiselle	devour	disc	dock	drink
debris	demonstrate	diabetes	discard	doctor	drive
debt	denmark	diagram	discipline	doctrine	droll
decade	dense	dial	disco	document	drop
decay	dent	dialect	discord	documentary	drought
deceased	dentist	dialogue	discount	documents	drown
deceive	deny	diameter	discourage	doe	drug
decent	depart	diamond	discourse	dog	drum
deception	department	diarrhea	discover	doll	dry
decide	departure	diary	discriminate	dolphin	duct
decipher	depend	dictate	discuss	domain	due
decision	deplete	diction	disease	domestic	duke
decisive	deposit	dictionary	disguise	domino	dump
declare	depress	die	disgust	donate	dun
decline	derive	diet	dish	donkey	dune
decrease	descend	differ	dishwasher	door	duplicate
dedicate	describe	difference	disinfect	doorway	dusk
deduce	desert	difficult	disk	dormitory	dust
deduct	desertification	dig	disobedient	dose	dustbin
deed	deserve	digest	disorder	dot	dutch
deep	design	digit	disperse	double	duty
defeat	designate	dignity	displace	doubt	dwarf
defect	desire	digraph	display	dough	dwell
defence	desk	dilemma	dispose	doughnut	dye
defend	desolate	dim	disprove	dour	dynamo
defer	despair	dimension	dispute	download	dynasty
deficiency	despatch	diminish	dissatisfy	dozen	eager
define	despise	dine	dissension	draft	eagle
deforest	dessert	dinner	dissipate	drag	ear
defraud	destination	dinosaur	distance	dragon	earl
defrost	destiny	dioxide	distinct	drain	early

earn	elevate	enormous	etiquette	expense	fairy
earring	elicit	enquire	eureka	experience	fairytale
earth	eliminate	enrich	evaluate	experiment	faith
earthenware	elope	ensure	evaporate	expert	fake
earthquake	eloquent	enter	evasion	expire	fall
ease	else	entertain	even	explain	fallout
east	email	enthusiastic	evening	explicit	false
easter	emancipate	entire	event	explode	fame
easy	embarrass	entitle	eventual	exploit	familiar
eat	embassy	environment	ever	explore	families
eave	ember	envy	evermore	explosion	family
ebb	embezzle	epidemic	evidence	export	famine
ebony	embrace	epileptic	evident	expose	fan
echo	emerge	episode	evil	expository	fancy
eclipse	emergency	epoch	evocative	express	fang
ecological	emigrate	equal	evoke	extend	fantastic
ecology	eminent	equation	evolve	exterior	fantasy
economy	emir	equator	exact	exterminate	far
ecosystem	emit	equip	exaggerate	extinct	farce
edge	emotion	equivalent	examine	extra	farm
edit	emotive	eradicate	example	extract	fascinate
educate	emperor	erase	exceed	extraordinary	fashion
efface	emphasis	ere	excellent	extraterrestrial	fast
effect	emphatic	erect	excess	extravagant	fasten
efficient	empire	erode	exchange	extreme	fat
effort	employ	erosion	excite	extrovert	fatal
eg	empty	err	exclaim	eye	father
egg	enable	error	excursion	eyebrow	fatigue
elaborate	enclose	erupt	excuse	eyesight	fault
elbow	encourage	escalator	execute	eyewitness	favour
elder	encyclopedia	escape	executive	fa	fax
elect	end	escort	exempt	fable	fear
electric	endorphin	eskimo	exercise	face	feast
electromagneti	endorse	especial	exert	facility	feather
c	endow	essay	exhibit	facsimile	feature
electron	enemy	essential	exhort	fact	federal
electronic	energy	establish	exist	factor	federate
elegant	engage	estate	exit	factory	fee
elegy	engine	esteem	exorcise	faculty	feed
element	engrave	estimate	exotic	fahrenheit	feedback
elemental	enjoy	eternal	expand	fail	feel
elementary	enlist	ethical	expect	faint	felled
elephant	enliven	ethnic	expedition	fair	fellow

female	flake	force	friend	gaze	go
feminine	flame	fore	fright	gazelle	goal
fence	flap	forebear	fro	gcse	goat
fend	flare	forecast	frog	gdp	god
ferocious	flash	forehead	front	gee	gold
fertile	flashy	foreign	frost	gene	goo
festival	flat	foreman	frown	general	good
fetch	flattering	foresee	fruit	generate	goodbye
fete	flavour	forest	fruitful	generator	goods
feudal	fleet	foreword	frustrate	generous	goose
fever	flesh	forge	fry	genetic	gourmet
fibre	flight	forget	fuel	genius	govern
fiction	fling	forgive	fulfil	genre	governance
fie	float	fork	full	gentle	gown
field	flock	form	fume	gentleman	grace
fierce	flood	formal	fun	genuine	grade
fight	floor	former	function	geography	gradual
figurative	flop	formula	fund	geometry	graduate
figure	floppy	fort	fundamental	geothermal	grain
file	flour	forthcoming	funeral	gerund	gram
files	flourish	fortify	funk	gesture	grand
fill	flow	fortnight	fur	ghost	granny
fillet	flower	fortune	furnish	ghoul	grant
film	flu	forum	furniture	gi	grape
filter	fluctuate	forward	further	giant	grapefruit
fin	fluent	fossil	furthermore	gift	graph
final	fluid	foul	fuse	ginger	graphic
finance	flush	found	future	giraffe	grasp
find	flute	foundation	gadget	girl	grass
fine	fly	founded	gain	girlfriend	grate
finger	fob	fox	galaxy	gist	grateful
finish	focus	fracture	gallery	give	gratitude
finished	foe	fragment	gallon	glacier	grave
fir	fog	frame	game	glad	graveyard
fire	folder	framework	gang	glass	gravity
firefight	follow	fraud	gap	gleam	graze
fireside	fond	free	garage	glimpse	great
firewood	font	freeze	garbage	gloat	greed
firm	food	freq	garden	globe	green
fish	foodstuff	frequent	garlic	glove	greengrocer
fit	fool	fresco	gas	glue	greenhouse
fix	foot	fresh	gate	glutton	greet
flag	forbid	freshman	gather	glycerin	grey

greyhound	hamburger	heat	hometown	humanitarian	impact
grid	hamlet	heaven	homeowner	humid	impatience
grief	hammer	heavy	homework	humour	impel
grill	hamster	heck	homonym	hunger	implement
grim	hand	height	homophone	hunt	implicate
grind	handbag	helicopter	honest	hunter	implore
grip	handicap	hell	honey	hurricane	imply
grit	handicraft	helmet	honour	hurry	impolite
grocer	handkerchief	help	hood	hurt	import
grocery	handle	hen	hook	husband	important
gross	handlebar	hence	hooray	hush	impose
grotesque	handsome	herb	hop	hut	impress
ground	handy	here	hope	hydra	improve
group	hang	hereby	horizon	hydroelectric	inch
grove	haphazard	heritage	horizontal	hygienic	incidence
grow	happen	hermetic	hormone	hymn	incident
grumpy	happy	hero	horn	hyphen	incinerate
guarantee	harbor	heroine	horoscope	hypothesis	include
guard	hard	hesitate	horrible	i	income
guess	hardly	hew	horror	ice	incorporate
guest	hardy	hiccup	horse	icing	increase
guide	hare	hide	horseback	icon	incredible
guidebook	harm	hideous	horseman	idea	indecision
guideline	harmony	hieroglyphic	hospitable	ideal	indeed
guilty	harry	high	hospital	identify	independent
guitar	harsh	highlight	hospitality	ideology	index
gulf	harvest	hike	host	idiom	india
gum	hat	hilarious	hostel	idle	indicate
gun	hate	hill	hostess	igloo	indifferent
gunpowder	hazard	himalaya	hostile	ignorant	indigenous
guy	head	hind	hot	ill	indigestion
gymnasium	headache	hindu	hotel	illusion	individual
gymnastics	headline	hint	hour	illustrate	indonesia
habit	headmaster	hip	house	image	industry
habitat	headquarters	history	household	imagine	infant
hack	headwords	hit	housewife	imam	infect
hair	heal	hitch	hover	imitate	infer
haircut	health	hitchhike	howl	immediate	inferior
hairstylist	heap	hobby	hug	immemorial	infirmary
hairdresser	hear	hold	huge	immerse	influence
hairdryer	heart	hole	hullo	immigrate	inform
hall	heartbreak	holiday	hum	immoral	informal
hallow	heartily	hollow	human	immune	infrastructure
ham					

infringe	interject	jack	ketchup	lady	lean
infuse	interlocutor	jackal	kettle	lag	leap
ingratitude	interminable	jacket	key	lagoon	learn
ingredient	intermittent	jail	keyboard	lake	lease
inhabitant	internal	jam	kick	lakeside	leather
inherit	international	jangle	kickback	lam	leave
initial	interpret	jar	kid	lamb	lecture
initiate	interrogate	jasmine	kidding	lame	leech
injure	interrupt	jazz	kidney	lament	leek
ink	intervene	jealous	kill	lamp	left
inn	interview	jeans	kilo	land	leftover
innovate	intimate	jeer	kilogram	landlady	leg
input	intone	jelly	kilometre	landmine	legal
inquire	introduce	jeopardy	kilt	landscape	legend
inquisition	introvert	jewel	kimono	lane	lei
inscription	inuit	job	kin	language	leisure
insect	inundate	jog	kind	lap	lemon
insecticide	invade	join	kindergarten	lapel	lemonade
insert	invasion	joke	king	large	lend
inside	invent	jot	kingship	late	length
insist	invest	journal	kiss	latin	lenient
insomnia	investigate	journalism	kit	latter	lens
inspect	invite	journey	kitchen	laud	leopard
inspire	invoice	joy	kiwi	laugh	lesson
instance	involve	judge	knave	laughter	let
instead	ion	judo	knee	launch	letter
institute	iq	jug	kneel	laundry	lettuce
instruct	iran	juice	knickers	laundry	level
instrument	irk	jumble	knife	laureate	lexical
insulin	iron	jump	knight	laurel	liable
insult	irony	jumper	knit	lava	liberal
insure	irrigate	jungle	knob	lavatory	liberty
integrate	irritable	junior	knock	law	libran
integrity	irritate	junk	knot	lay	library
intellectual	islam	just	know	layer	licence
intelligence	island	justice	knuckle	layout	lid
intend	isle	justify	label	lazy	lie
intense	isolate	kabuki	laboratory	le	lieutenant
intent	issue	kangaroo	labour	lea	life
interact	italic	karaoke	lack	lead	lifeguard
interest	itch	keen	lad	leaf	lifestyle
interfere	item	keep	ladder	leaflet	lift
interior	ivory	kerosene	ladle	league	light

lightning	log	magnitude	mascara	mental	minor
likely	logic	maid	mask	mention	mint
liken	logo	mail	mass	mentor	minute
likeness	lone	main	mast	menu	mirror
lilt	long	maintain	master	merchandise	mis
lily	longing	major	mat	merchant	miscellany
lime	look	make	match	mercury	misery
limerick	loon	malaria	matchstick	mere	miss
limit	loose	male	mate	merge	mission
line	lord	mall	material	mesopotamian	mistake
linen	lorry	malnutrition	maternity	message	mister
lines	lose	malpractice	mathematics	metal	mistreat
linger	loss	malt	matter	metaphor	mix
link	lot	man	mature	mete	mobile
lion	loud	manage	mausoleum	meteor	mock
lip	lounge	manifesto	maximum	method	model
lipid	lout	manioc	maybe	metre	modern
lipstick	love	manipulate	mayonnaise	metro	modern
liquid	low	manner	mayor	metropolis	modest
lire	lowland	mantis	mb	mew	modify
list	loyal	manual	md	microcompute	moment
listen	lubricant	manufacture	meal	r	monarch
lists	luck	map	mealttime	mid	monetary
literacy	luggage	maple	mean	middle	money
literary	lull	mar	meaning	mien	monitor
literature	luminous	marathon	measure	mighty	monkey
litre	lump	marble	meat	migrate	mono
litter	lunar	mare	mechanic	mild	monopoly
little	lunch	margarine	media	mile	monotony
live	lung	marge	medical	militant	monoxide
livelihood	lust	margin	medication	military	monster
liver	lyric	marina	medicine	militate	month
load	machine	marine	meditate	milk	monument
loaf	mad	maritime	medium	mill	mood
loan	madam	mark	meet	mime	moon
lobby	madame	market	melancholy	mince	moor
local	madeleine	marmalade	melon	mind	moral
locality	magazine	marry	melt	miner	moreover
locate	magic	marsh	member	mineral	morning
loch	magma	martian	memo	mini	mortal
lock	magnesium	martyr	memorable	minimum	moslem
locker	magnet	marvel	memory	minister	mosque
lodge	magnify	marx	mend	ministry	mosquito

moth	narrate	night	objective	option	ozone
mother	narrow	nitrogen	oblige	oral	pace
motion	nation	nobel	observe	orange	pacific
motive	native	nod	obtain	orbit	pacifist
motor	nature	node	obtuse	orchard	pack
motorcycle	nausea	nodes	obvious	order	paella
motorway	nautical	noise	occasion	ordinal	page
mount	naval	nomad	occupation	ordinary	pail
mountain	nave	nominal	occupy	organ	pain
mouse	navigable	non	occur	organic	paint
moussaka	navigate	noon	ocean	organism	pair
moustache	navy	norm	odd	organize	paisley
mouth	neat	normal	ode	orient	pal
move	necessary	north	offence	origin	palace
movie	necessity	nose	offer	otherwise	pale
mow	neck	nostalgia	office	ounce	palm
mph	nee	note	official	outcome	palsy
muck	need	notebook	often	outgoing	pan
mud	needlework	notice	ogre	outing	pancake
mule	needy	notion	oil	outline	panda
multinational	negative	nougat	ointment	outstanding	panel
multiple	neglect	novel	okay	oval	panic
multiply	negro	now	old	oven	panther
murder	neighbour	nuclear	olive	overall	pants
muscle	nepal	nudge	olympic	overalls	paper
muscular	nephew	number	omelette	overdose	par
museum	nepotism	numeral	omit	overflow	parabola
mushroom	nerve	numerous	onion	overhead	parachute
music	nervous	nurse	online	overland	parade
muslim	nest	nursery	only	overnight	paraffin
mute	net	nut	onwards	overseas	parallel
mutton	network	nutmeg	open	oversee	paralyse
mutual	neutral	nutrient	opera	overtake	paramedic
mynah	never	nutrition	operate	overtime	paraphrase
mystery	new	nutritious	opinion	overturn	parasite
myth	news	nylon	opportunity	overweight	parasympathet ic
mythic	newspaper	oak	oppose	owe	
nadir	newsstand	oasis	opposite	owl	pardon
nail	newton	oast	oppress	own	parent
naked	nibble	obese	opt	ox	park
name	nice	obey	optic	oxide	parlour
napkin	nickname	obituary	optimism	oxygen	parrot
nappy	niece	object	optimist	oyster	part

partiality	penalty	phlegmatic	plant	portrait	preserve
participate	pencil	phonetic	plantation	portray	president
participle	penguin	phosphorus	plastic	portugal	press
particle	penicillin	photograph	plate	pose	pressure
partner	peninsula	photosynthesis	plateau	position	prestige
party	penny	photovoltaic	play	positive	presume
pass	pension	physical	playwright	possess	pretend
passage	people	physician	please	possible	pretty
passenger	pepper	physics	pleasure	post	prevent
passion	percent	piano	pledge	postcard	preview
passport	perfect	piccolo	plenty	postman	previous
password	perforce	pick	plot	pot	price
past	perform	pickle	plug	potato	prick
pasta	perfume	picnic	plum	potential	pride
paste	perhaps	picture	plumb	pothole	priest
pasteurise	peril	picturesque	plural	pottery	primary
pastime	period	pie	pneumonia	poultry	prime
pastry	permanent	piece	poach	pound	prince
pasture	permit	pile	pocket	pour	princess
pat	perpendicular	pill	poem	poverty	principle
patent	persia	pilot	poet	powder	print
path	person	pin	point	power	priority
patient	persona	pincer	poison	pox	prison
patriot	personalise	pinch	polar	practical	private
patriotic	personality	pink	pole	practise	privilege
patron	personnel	pioneer	police	prairie	prize
pattern	perspire	pious	policy	pray	probable
pause	persuade	pipe	polio	pre	probity
pay	pertinent	piracy	polish	preach	problem
pea	peru	pirate	polite	preamble	proceed
peace	pessimist	pisces	politic	precaution	process
peacekeeping	pest	piston	pollute	precede	processing
peach	pesticide	pit	polo	precious	produce
peacock	pet	pitch	polyester	precise	product
peak	petition	pity	pond	predict	profession
peanut	petrol	pizza	pool	preface	professional
pear	petroleum	place	poor	prefer	professor
pebble	pharaoh	plague	pop	prehistoric	profile
peculiar	pharmaceutica	plaid	popular	prejudice	profit
pedestal	l	plain	population	premium	programme
pedestrian	phenomenon	plan	pork	prepare	progress
peek	philippine	plane	port	prescribe	prohibit
pen	philosophy	planet	portfolio	present	project

promise	punctual	radioactive	reception	rehearse	resemble
promote	punctuate	radium	recess	reign	reserve
prompt	punish	rage	recipe	rein	reshuffle
prone	punk	ragged	recipient	reinforce	reside
pronounce	pupil	raid	recite	reject	residue
pronunciation	purchase	rail	reckon	relation	resign
proof	pure	rain	recluse	relative	resist
proper	puree	raincoat	recognize	relativity	resolution
property	purple	rainfall	recommend	relax	resolve
prophet	purpose	rainforest	record	release	resort
proportion	purse	rainstorm	recount	relevant	resource
propose	pursue	raise	recover	relief	respect
prose	push	rally	recreation	relieve	respective
prospect	put	ram	recruit	religion	respiration
prospectus	puzzle	range	rectangle	reload	respond
prosper	pyjamas	ranges	recuperate	rely	response
prostate	pyramid	rank	recur	remain	responsible
protect	qualify	rap	recycle	remark	rest
protein	quality	rapid	red	remedy	restaurant
protest	quantity	rare	redress	remember	restore
proton	quarrel	raspberry	reduce	remind	restrict
protractor	quarter	rat	reed	reminisce	result
proud	quartet	rate	refer	remote	resume
prove	quartz	rather	referee	remove	retain
proverb	queen	ratify	referendum	rent	retire
provide	question	ravage	refine	repair	return
provident	questionnaire	rave	reflect	reparation	reunion
provoke	queue	raw	reflex	repay	reveal
psychiatry	quick	ray	refrain	repeat	revenge
psychology	quid	raze	refresh	repel	revenue
psychotherapy	quiet	razor	refrigerate	repertory	reverberate
pub	quinine	reach	refuge	replace	revere
public	quit	react	refund	reply	reverse
publication	quite	read	refuse	report	review
publicize	quiz	ready	refute	represent	revise
publish	quote	real	regain	reproach	revive
puce	rabbit	realise	regard	reproduce	revolt
pudding	rabies	really	regime	republic	revolution
pull	race	reason	region	reputation	revolve
pun	racket	recapitulate	register	request	reward
punch	radiate	receipt	regret	require	rheumatism
punchline	radical	receive	regular	rescue	rhinoceros
punchy	radio	recent	regulate	research	rhyme

rhythm	route	sardine	scrimp	sequel	shine
rice	routine	satchel	script	sequence	ship
rich	row	sate	scrutiny	serene	shirt
rid	rowan	satellite	sculpt	series	shock
riddle	royal	satisfy	scurry	serious	shoe
ride	rub	sauce	sea	serpent	shoot
rife	rubber	saucer	seafarer	serve	shop
rig	rubbish	sauna	seafood	service	shore
right	rubric	sausage	seal	session	short
rile	rude	savage	search	set	shortlist
rill	rue	save	seashore	setter	shorts
rim	rugby	savour	season	settings	shoulder
rime	ruin	say	seat	settle	shout
ring	rule	scale	seawater	severe	show
rinse	rum	scan	seconds	sew	shower
ripe	run	scandal	secret	sewage	shows
rise	rural	scar	secretariat	sex	showy
risk	rush	scarce	secretary	sf	shrimp
rite	sack	scare	section	shabby	shrug
rival	sacrifice	scatter	sector	shade	shut
river	sad	scenario	secure	shadow	shy
road	saddle	scene	see	shaft	sibling
roast	safe	schedule	seek	shake	sic
rob	saffron	scheme	seem	shakespeare	sick
robe	sail	scholar	segment	shame	side
robin	saint	scholarship	seize	shampoo	sieve
robot	sake	school	seizure	shape	sift
rock	salad	schoolgirl	seldom	share	sight
rod	salary	science	select	shareholder	sightsee
roger	sale	scissors	self	shark	sign
roil	salesman	scope	sell	sharp	signal
role	salmon	score	semi	shatter	signature
roll	salt	scoreboard	senate	shave	significant
roof	salute	scout	senator	shear	silence
room	same	scrabble	send	sheep	silicon
roost	sample	scramble	senior	sheer	silk
root	sanction	scrap	sensation	sheet	silver
rope	sand	scrapbook	sense	shelf	similar
roses	sandal	scratch	sensible	shell	simile
rotate	sandwich	scream	sensitive	shelter	simple
rotor	sane	screen	sentencing	shepherd	simulate
rough	sans	screenshot	sentiment	shift	simultaneous
rouse	sapphire	scribble	separate	shilling	sin

sincere	small	somewhat	spice	static	strain
sing	smallpox	son	spill	station	strait
single	smart	sonata	spirit	statistic	strange
singular	smell	song	spite	statue	strap
sinister	smile	soon	splash	status	strategy
sink	smog	sore	splendid	stay	strawberry
sion	smoke	sorrow	split	stead	stream
sister	smooth	sorry	spoil	steady	street
sit	smuggle	sort	sponsor	steak	strength
site	snack	sos	spoon	steal	stress
situate	snail	sot	sport	steam	stretch
size	snake	soul	sportsman	steamship	stretcher
skate	snatch	sound	sportswear	steel	strict
skeleton	sneak	soup	spot	steep	strike
sketchbook	sneer	source	spout	stein	string
ski	snippet	south	sprain	stem	stringent
skill	snivel	southeast	spray	step	strip
skim	snow	southwest	spread	stepdaughter	strong
skin	snowflake	souvenir	spring	stepmother	structure
skinny	snowman	sow	sprinkle	sterile	struggle
skip	snowstorm	space	spy	stethoscope	stubborn
skirt	so	spaceship	square	stew	stud
sky	soak	spaghetti	stables	steward	student
slab	soap	spam	stadium	stick	studio
slang	soccer	span	staff	stiff	study
slate	social	spank	stage	stimulate	stuff
slave	socialise	spanner	stagnate	sting	stumble
sleaze	society	spare	stairs	stir	stupid
sleep	sociology	spark	stakeholder	stock	style
sleeve	sock	speak	stall	stocking	subdivide
slender	soda	spear	stamp	stoke	subdue
slice	soft	special	stand	stomach	subject
slide	software	species	standard	stomp	subjective
slight	soil	specific	standby	stone	submarine
slim	solar	specify	stanza	stool	submerge
slime	soldier	spectator	staple	stop	submit
slip	solemn	spectroscope	star	store	subsequent
slippery	solid	speculate	stare	stork	subsidy
slog	solidarity	speech	start	storm	substance
slogan	solitude	speed	startle	stormy	subtract
slope	solution	spell	starve	story	suburb
slow	solve	spend	state	storyline	subway
sly	someday	sphere	states	straight	succeed

suck	surgery	tactic	televise	thirst	tolerant
sudden	surname	tag	television	thorough	tolerate
sue	surplus	tail	tell	thou	toll
suffer	surprise	take	teller	thread	tomato
sufficient	surround	taken	temperament	threat	tomb
suffocate	survey	tale	temperate	thrill	tomorrow
sugar	survive	talent	temperature	thrive	ton
suggest	susceptible	talk	temple	throat	tone
suit	suspect	tall	temporary	throb	tongue
suitcase	suspend	tan	tenant	throw	tonight
suite	suspense	tangible	tend	thumb	too
sulphur	suspicious	tank	tender	thunder	tool
sulphuric	sustain	tap	tennis	thus	tooth
sum	swagger	tape	tent	thwart	toothache
summary	swallow	tare	term	tick	top
summer	swamp	target	terminate	ticket	topic
sun	swap	tarn	terrace	tide	topsoil
sunbathe	sweater	tart	terrible	tidy	tor
sunburn	sweatshirt	task	terrific	tie	torch
sunglasses	sweatshop	taste	territory	tiger	torrid
sunlit	sweep	tatty	terror	tight	torture
sunset	sweet	taurean	tertiary	tile	toss
super	swerve	tax	test	timber	total
superb	swift	taxi	text	time	totalitarian
superfluous	swim	tea	textile	timetable	touch
superior	swirl	teach	thank	tin	tough
superman	switch	team	thanksgiving	ting	tour
supermarket	sword	teamwork	theatre	tint	toward
superstition	symbol	teapot	theft	tiny	tower
supervise	sympathise	tear	theme	tip	town
supper	sympathy	tease	then	tire	toxic
supplement	symphony	teaspoon	theorem	tiresome	toxin
supply	synthesis	technical	theory	tissue	toy
support	synthetic	technique	therapy	tit	trace
suppose	system	techno	therefore	title	track
supposition	t	technology	thermometer	toast	trade
suppress	tab	tedious	thermos	tobacco	tradition
supreme	table	teenage	thesis	today	traffic
sure	tablespoon	telecommunic	thick	toe	tragedy
surf	tablet	ation	thief	toga	tragic
surface	tabloid	telegraph	thin	together	trail
surge	taco	telephone	thing	toiletory	train
surgeon	tact	telescope	think	tokens	trait

trample	tsar	upbringing	vegetate	volunteer	wed
trans	tub	update	vehicle	vomit	week
transcribe	tube	upgrade	vein	vote	weekday
transcript	tuberculosis	upkeep	vend	vow	weep
transfer	tulip	upload	vent	voyage	weigh
transform	tumble	upper	ventilate	wafer	weight
transit	tuna	upright	venus	wage	welcome
transition	tune	uproot	verify	wagon	welfare
translate	tunnel	upset	verse	wait	well
transmit	tunny	upwards	version	wake	west
transparency	turban	uranium	vertical	walk	western
transparent	turn	urban	vest	walkman	wet
transplant	turnip	urn	vet	wall	whale
transport	turnover	use	viable	wallaby	wheat
transpose	tusk	used	victim	wallet	wheel
trap	twin	usual	victory	wane	wheelchair
trash	twist	utensil	video	want	whether
travel	txt	utility	view	war	whirl
treasure	type	uv	viewpoint	ward	whirr
treat	types	vacancy	vigour	warder	whisker
treatise	typewrite	vacant	viking	warehouse	whistle
treaty	typical	vacation	villa	warfare	white
tree	ufo	vaccinate	village	warlike	whole
tremble	ugly	vaccine	villain	warm	wholemeal
tremendous	ultra	vacuum	vinegar	warn	wicked
trend	umbrella	vain	violate	warp	wide
triangle	uncle	valentine	violent	warrior	widespread
tribe	undergo	valley	violet	wash	widow
tribune	underground	value	violin	washboard	wife
trick	underlie	van	virtual	waste	wig
trigger	underline	vandal	virtue	watch	wild
trip	undermine	vanguard	visage	water	win
triumph	understand	vanilla	visit	waterproof	wind
troop	unicorn	vanish	vita	wave	window
tropics	uniform	variety	vital	wavy	wine
trouble	unify	various	vitamin	way	wing
trousers	union	vary	vocation	weak	wink
trowel	unique	vase	void	wealth	winter
truck	unit	vast	volcano	weapon	wipe
true	unite	vd	volley	wear	wire
trumpet	universe	veal	volleyball	weary	wise
trust	university	vegetable	volume	weather	wish
try	unwind	vegetarian	voluntary	web	wishful

wit	yoghurt
withstand	yogurt
witness	yolk
wolf	yore
woman	young
wonder	youth
wonderful	zebra
wood	zenith
woodwork	zip
wool	zone
word	zoo
words	zulu
work	
workforce	
workload	
workshop	
world	
worldwide	
worm	
worry	
worse	
worship	
worth	
wound	
wreck	
wrestle	
wrinkle	
write	
written	
wrong	
wrongdoing	
xerox	
xylophone	
yacht	
yak	
yard	
year	
yeast	
yellow	
yes	
yesterday	
yet	
yield	

APPENDIX VI

List of words from the First -1000 word list of most frequent words occurring all Algerian textbooks (850 words)

able	average	captain	contain	determine	enemy
accept	bad	car	content	develop	enjoy
accord	ball	care	continue	die	enter
according	bank	carry	control	differ	equal
account	bar	case	corn	difference	escape
accountant	base	castle	cost	difficult	even
act	basis	catch	cotton	direct	evening
active	battle	cause	council	direction	event
actual	bear	causes	count	disadvantage	ever
add	beauty	centre	country	disappear	example
address	become	chance	course	discover	exchange
admit	bed	change	court	distance	exercise
adopt	begin	character	cover	distinguish	exist
advance	belief	charge	cross	district	expect
advantage	believe	chief	crowd	divide	expense
adventure	belong	child	crown	doctor	experience
affair	best	choice	cry	dog	experiment
again	big	choose	current	door	explain
age	bill	church	cut	doubt	express
agent	bird	circle	dad	draw	extend
agree	black	citizen	danger	dream	eye
air	blood	city	dark	dress	face
allow	blow	claim	date	drink	fact
almost	blue	class	daughter	drive	factory
alone	board	classify	day	drop	fail
already	boat	clear	dead	dry	fair
always	body	close	deal	due	faith
amount	book	cloud	dear	duty	fall
ancient	box	coal	death	ear	familiar
animal	boy	coast	decide	early	families
answer	branch	coin	decision	earth	family
appear	bread	cold	decisive	earthquake	far
apply	break	college	declare	east	farm
appoint	bridge	colony	deep	easy	fast
arise	bright	colour	defeat	eat	father
arm	bring	come	degree	effect	favour
army	broad	command	demand	efficient	fear
arrive	brother	committee	department	effort	feel
art	build	common	depend	egg	fellow
article	burn	company	describe	elect	field
ask	business	complete	desert	else	fight
associate	buy	concern	desire	empire	figure
attack	call	condition	destroy	employ	fill
attempt	capital	consider	detail	end	find

fine	go	industry	lip	miss	open
finish	god	influence	listen	mister	operate
finished	gold	instead	literary	modern	opinion
fire	good	interest	literature	moment	opportunity
fish	goodbye	introduce	little	money	order
fit	great	iron	live	month	ordinary
fix	green	join	local	moon	organize
flight	ground	joy	long	moral	otherwise
floor	group	judge	look	moreover	outstanding
flow	grow	just	lord	morning	owe
flower	hand	justice	lose	mother	own
fly	hang	keep	loss	motor	page
follow	happen	kill	love	mountain	paint
food	happy	kind	low	mouth	paper
force	hard	king	machine	move	part
foreign	hardly	know	main	music	particle
forest	head	lack	make	name	party
forget	hear	lady	man	nation	pass
form	heart	lake	manner	native	past
former	heat	land	manufacture	nature	pay
fortune	heaven	language	mark	necessary	peace
found	heavy	large	market	necessity	people
free	help	late	marry	need	perhaps
fresh	here	latter	mass	neighbour	permit
friend	high	laugh	master	never	person
front	hill	laughter	material	new	picture
full	history	law	matter	news	piece
furnish	hold	lay	maybe	newspaper	place
furniture	honour	lead	mean	night	plain
further	hope	learn	meaning	north	plan
future	horse	leave	measure	note	plant
gain	hot	left	meet	notice	play
game	hour	length	member	now	please
garden	house	let	memory	number	pleasure
gas	human	letter	mention	numerous	point
gate	husband	level	mere	object	poor
gather	i	library	metal	observe	popular
general	idea	lie	mid	occasion	population
gentle	ill	life	middle	offer	position
gentleman	important	lift	mile	office	possess
gift	inch	light	milk	official	possible
girl	include	likely	mind	often	post
give	increase	limit	miner	oil	pound
glad	indeed	line	minister	old	poverty
glass	independent	lines	minute	only	power

prepare	receive	scarce	situate	stone	throw
present	recent	scene	size	stop	thus
president	recognize	school	sky	store	time
press	record	science	sleep	story	today
pressure	red	sea	small	strange	together
pretty	reduce	season	smile	stream	ton
prevent	refuse	seat	snow	street	too
price	regard	secret	so	strength	top
private	relation	secretary	social	strike	total
problem	relative	see	society	strong	touch
produce	religion	seem	soft	struggle	toward
product	remain	sell	soldier	student	town
profit	remark	send	son	study	trade
progress	remember	sense	song	subject	train
promise	reply	sensible	soon	substance	travel
proof	report	sensitive	sort	succeed	tree
proper	represent	separate	soul	suffer	trouble
property	republic	serious	sound	suggest	true
propose	reserve	serve	south	summer	trust
protect	respect	service	southeast	sun	try
prove	respective	set	southwest	supply	turn
provide	rest	settle	space	support	type
public	result	shadow	speak	suppose	types
pull	return	shake	special	sure	understand
purpose	rich	shape	speech	surface	union
put	ride	share	speed	surprise	unit
quality	right	shine	spend	surround	unite
quantity	ring	ship	spirit	sweet	university
quarter	rise	shoot	spite	sword	use
queen	river	shore	spot	system	used
question	road	short	spread	table	usual
quite	rock	shoulder	spring	take	valley
race	roll	show	square	taken	value
raise	room	shows	stage	talk	variety
rank	rough	side	stand	tax	various
rate	royal	sight	standard	teach	victory
rather	rule	sign	star	tear	view
reach	run	signature	start	tell	village
read	safe	silence	state	temple	virtue
ready	sail	silver	states	term	visit
real	sale	simple	station	test	vote
realise	salt	sing	stay	then	wage
really	same	single	steel	therefore	wait
reason	save	sister	step	thing	walk
receipt	say	sit	stock	think	wall

want
war
watch
water
wave
way
wealth
wear
week
welcome
well
west
western
whether
white
whole
wide
wife
wild
win
wind
window
winter
wise
wish
woman
wonder
wonderful
wood
word
words
work
world
worth
wound
write
written
wrong
year
yes
yesterday
yet
yield
young
youth

APPENDIX VII

List of words from Second -1000 word list of most frequent words occurring all Algerian textbooks

abroad	asleep	blind	cave	connect	deception
absence	attend	block	cent	conquer	decrease
absent	attract	boast	centimetre	conquest	deed
absolute	audience	boil	century	conscious	defence
accident	aunt	bold	ceremony	convenience	defend
accuse	autumn	bone	chain	cook	delay
ache	avenue	border	chair	cool	delight
admire	avoid	borrow	chairman	copper	deliver
advertise	awake	bottle	chalk	copy	descend
advice	axe	bottom	charm	cork	deserve
advise	baby	bounds	cheap	corner	desk
afford	bag	bow	cheat	correct	despair
afraid	bake	bowl	check	cottage	diamond
afternoon	balance	brain	cheer	cough	dictionary
agriculture	band	brave	cheese	courage	dig
ahead	barber	breakfast	cheque	cousin	dine
aim	bare	breath	chest	cow	dinner
airplane	barrel	breathe	chicken	coward	dirty
alike	basin	bribe	chimney	crash	disappoint
alive	basket	brick	christmas	cream	disapprove
aloud	bath	broadcast	civilise	creature	discipline
altogether	bathe	brown	clean	crime	discuss
ambition	bay	brush	clerk	criminal	disease
ambitious	beak	bucket	clever	critic	disgust
amuse	beam	bury	climb	crop	dish
anger	bean	bus	clock	cruel	dissatisfy
angle	beard	bush	cloth	cultivate	disturb
angry	beast	busy	club	cup	dive
annoy	beat	butter	coat	cupboard	donkey
anxiety	beg	button	coffee	cure	dot
anxious	behave	cage	collect	curious	double
apart	behaviour	cake	combine	curl	dozen
apology	bell	calculate	comfort	curse	drag
applause	belt	calculator	commerce	curtain	drown
apple	bend	calm	companion	curve	drum
approve	berry	camera	compare	custom	dust
arch	bicycle	camp	compete	customer	eager
argue	bind	canal	competition	damage	earn
arrange	birth	cap	complain	dance	ease
arrest	bit	card	complaints	dare	edge
arrow	bite	cart	compose	deaf	educate
artificial	bitter	cat	confidence	debt	elder
ash	blame	cattle	confuse	decay	electric
ashamed	bless	caution	congratulate	deceive	elephant

empty	fool	hate	journey	meal	pain
enclose	foot	heal	juice	meat	pair
encourage	forbid	health	jump	mechanic	pale
engine	forgive	heap	key	medicine	pan
enquire	fork	hesitate	kick	melt	pardon
entertain	formal	hide	kilogram	mend	parent
entire	forward	hit	kilometre	merchant	park
envy	frame	hole	kiss	message	passage
especial	freeze	holiday	kitchen	metre	passenger
essential	frequent	hollow	knee	mild	paste
evil	fright	honest	kneel	mill	path
exact	fruit	hook	knife	mineral	patient
examine	fry	horizon	knock	mistake	patriotic
excellent	fun	hospital	knot	mix	pattern
excess	funeral	host	ladder	model	pause
excite	fur	hotel	lamp	modest	peculiar
excuse	gallon	hullo	lazy	monkey	pen
explode	gap	hunger	leaf	motion	pencil
explore	garage	hunt	lean	mouse	penny
explosion	generous	hurry	leather	mud	perfect
extra	goat	hurt	leg	multiple	perform
extraordinary	govern	hut	lend	multiply	permanent
extreme	grace	ice	lesson	murder	persuade
faint	gradual	ideal	liberty	mystery	pet
false	grain	idle	lid	nail	photograph
fan	gram	imagine	liquid	narrow	pick
fancy	grand	imitate	list	neat	pile
fashion	grass	immediate	lists	neck	pin
fasten	grateful	improve	litre	neglect	pinch
fat	grave	inform	load	nephew	pink
fault	greed	informal	loaf	nest	pipe
feast	greet	ink	loan	net	pity
feather	grey	inn	lock	nice	plane
female	grind	inquire	log	niece	plate
fence	guard	insect	lone	noise	plenty
fever	guess	inside	loose	noon	plural
fierce	guest	instrument	lot	nose	pocket
film	guide	insult	loud	nurse	poem
finger	guilty	insure	loyal	nut	poet
firm	gun	intend	luck	obey	poison
flag	habit	intent	lump	ocean	police
flame	hair	interfere	lunch	omit	polish
flash	hall	international	lung	onwards	polite
flat	hammer	interrupt	mad	oppose	pool
flavour	handkerchief	invent	mail	opposite	pot
flesh	handle	invite	male	orange	pour
float	harbor	island	manage	organ	powder
flood	harm	jealous	map	origin	practical
flour	harvest	jewel	mat	outline	practise
fond	hat	joke	match	pack	pray

preach	repeat	severe	spill	tender	uncle
precious	replace	sew	splendid	tent	universe
prefer	reproduce	shade	split	terrible	upper
prejudice	reputation	shame	spoil	thank	upright
preserve	request	sharp	spoon	theatre	upset
pretend	rescue	shave	sport	thick	upwards
pride	resign	sheep	staff	thief	vain
priest	resist	sheet	stairs	thin	verse
print	responsible	shelf	stamp	thirst	violent
prison	restaurantire	shell	steady	thorough	voyage
prize	revenge	shelter	steal	thread	wake
probable	review	shilling	steam	threat	warm
profession	reward	shirt	steep	throat	warn
programme	rice	shock	stem	thumb	wash
prompt	rid	shoe	stick	thunder	waste
pronounce	ripe	shop	stiff	ticket	weak
proud	risk	shout	sting	tide	weapon
punctual	rival	shower	stir	tidy	weather
punish	roast	shut	stocking	tie	weigh
pupil	rob	sick	stomach	tight	weight
pure	rod	signal	storm	tin	wet
purple	roof	silk	straight	tip	wheat
push	root	sincere	strap	tire	wheel
puzzle	rope	sink	stretch	title	whistle
qualify	row	skill	strict	tobacco	wicked
quarrel	rub	skin	string	toe	widow
quick	rubber	skirt	strip	tomorrow	wine
quiet	rubbish	slave	stuff	tongue	wing
rabbit	rude	slide	stupid	tonight	wipe
radio	ruin	slight	suck	tool	wire
rail	rush	slip	sudden	tooth	witness
rain	sacrifice	slope	sugar	tough	wool
rapid	sad	slow	suit	tour	worm
rare	saddle	smell	supper	tower	worry
rat	sake	smoke	suspect	toy	worse
raw	salary	smooth	suspicious	track	worship
ray	sample	snake	swallow	translate	wreck
razor	sand	soap	sweep	trap	yard
recommend	satisfy	sock	swim	treasure	yellow
refer	sauce	soil	sympathy	treat	
reflect	saucer	solemn	tail	tremble	
refresh	scale	solid	tall	tribe	
regret	scatter	solution	tap	trick	
regular	scissors	solve	taste	trip	
relief	scratch	sore	taxi	tube	
relieve	screen	sorry	tea	tune	
remedy	search	soup	telegraph	twist	
remind	seize	sow	telephone	typical	
rent	seldom	spare	temperature	ugly	
repair	self	spell	tend	umbrella	

APPENDIX VIII

Words from the Academic Word List occurring in the

abandon	assume	compile	convert	domestic
abstract	assure	complex	convince	draft
academy	attach	compound	cooperate	drama
access	attitude	comprehen	coordinate	economy
accommod	author	sive	core	edit
ate	authority	comprise	corporate	element
accompany	automate	compute	correspond	eliminate
accumulate	automatic	concentrate	couple	emerge
accurate	available	conclude	create	emphasis
achieve	aware	conduct	credit	emphatic
acknowled	benefit	confer	culture	enable
ge	bond	confine	currency	energy
acquire	brief	confirm	cycle	enormous
adapt	capable	conflict	data	ensure
adapted	capacity	consensus	debate	environmen
adequate	category	consent	decade	t
adjacent	challenge	consequenc	decline	equation
adult	channel	e	deduce	equip
affect	chapter	consequent	define	equivalent
aid	chart	considerabl	demonstrat	erode
allocate	chemical	e	e	erosion
alter	circumstan	consist	deny	error
alternate	ce	consistent	depress	establish
alternative	cite	constant	derive	estate
analyse	civil	constitute	design	estimate
analysis	clarify	constitution	detect	ethical
annual	classic	constrain	device	ethnic
anticipate	code	construct	devote	evaluate
apparent	coherence	consult	dimension	eventual
appendix	coherent	consume	diminish	evidence
appreciate	coincide	contact	discriminat	evident
approach	collapse	context	e	evolve
appropriate	colleague	contract	displace	exceed
approximat	comment	contradict	display	exhibit
e	commissio	contrary	dispose	expand
area	n	contrast	distinct	expert
aspect	commit	contribute	distribute	explicit
assemble	communica	controversy	diverse	exploit
assess	te	convene	document	export
assign	community	convention	documents	expose
assist	compensate	converse	domain	extract

facility	incidence	licence	overseas	react
factor	incident	link	panel	recover
feature	income	locate	parallel	refine
federal	incorporate	logic	participate	regime
fee	index	maintain	partner	region
file	indicate	major	percent	register
files	individual	manipulate	period	regulate
final	infer	manual	phenomeno n	reinforce
finance	infrastructu re	margin	philosophy	reject
fluctuate	initial	mature	physical	relax
focus	initiate	maximum	policy	release
formula	injure	media	pose	relevant
forthcomin g	innovate	medical	positive	rely
foundation	input	medium	potential	remove
founded	insert	mental	precede	require
framework	inspect	method	precise	research
function	instance	migrate	predict	reside
fund	instance	military	presume	resolution
fundamenta l	institute	minimum	previous	resolve
furthermore	instruct	ministry	primary	resource
generate	integrate	minor	prime	respond
globe	integrity	modify	principle	response
goal	intelligence	monitor	priority	restore
grade	intense	motive	proceed	restrict
grant	interact	mutual	process	retain
guarantee	internal	negative	processing	reveal
guideline	interpret	network	professiona l	revenue
hence	intervene	neutral	prohibit	reverse
highlight	invest	norm	project	revise
hypothesis	investigate	normal	promote	revolution
identify	involve	notion	proportion	role
ideology	isolate	nuclear	prospect	route
ignorant	issue	objective	psychology	scenario
illustrate	item	obtain	publication	schedule
image	job	obvious	publish	scheme
immigrate	journal	occupation	purchase	scope
impact	justify	occupy	pursue	section
implement	label	occur	quote	sector
implicate	labour	odd	radical	secure
imply	layer	option	range	seek
impose	lecture	orient	ranges	select
	legal	outcome		sequence
	liberal	overall		series

sex	stress	symbol	topic	unify
shift	structure	tape	trace	unique
significant	style	target	tradition	utility
similar	submit	task	transfer	vary
simulate	subsequent	team	transform	vehicle
site	subsidy	technical	transit	version
somewhat	sufficient	technique	transition	violate
source	sum	technology	transmit	virtual
specific	summary	temporary	transport	volume
specify	supplement	terminate	trend	voluntary
sphere	survey	text	trigger	volunteer
statistic	survive	theme	undergo	welfare
status	suspend	theory	underlie	widespread
strategy	sustain	thesis	uniform	

ملخص

تعتبر الكتب المدرسية لتدريس الانجليزية في الجزائر المصدر الأهم إن لم تكن الوحيد لتعلم الألفاظ. لذا تناول هذا البحث دراسة هذه الكتب والتي عددها سبعة من ناحية التغطية المعجمية والمقروئية. فصمم هذا البحث لتشكيل (أ) مجموعة معجمية، (ب) مقارنة الكتب من ناحية التغطية المعجمية و (ج) تحديد المقروئية. كان الهدف الأساسي والرئيس هو ان كان مستوى مقروئية الكتب يتماشى و التغطية المعجمية أي تحديد إن كان في المستوى القرائي المستقل، أم المستوى القرائي التعليمي ن أم المستوى القرائي الإحباطي. اضافة الى ذلك تم مقارنة قائمة كل الألفاظ الواردة في الكتب السبع لقوائم الألفاظ General Service List و Academic Word List لتقييم ما إذا كانت الكتب المدرسية قادرة على توفير مفردات كافية ومفيدة ومناسبة. وكان آخر هدف هو تزويد مدرسي اللغة الإنجليزية بوسيلة (البرمجيات) لمقارنة المستويات حتى يتسنى تحديد المفردات المطلوبة لتكميلة للوصول الى نسبة 95 % من الفهم. تتميز المنهجية المعتمدة بتعدد الوسائل (برامج الكمبيوتر) حيث تم استخراج قوائم كل مفردات الكتب المدروسة. وقد أظهرت النتائج أن مستوى المقروئية والتغطية المعجمية في جميع الكتب منخفض. كما تبين من التغطية المعجمية أن فرصة تعلم مفردات كافية ومفيدة ومناسبة غير متوفرة للطلبة. وخلصت الدراسة بوصيات بشأن كيفية معالجة هذه المشكلة.

Résumé

Mots-clé: lisibilité - couverture lexicale - compréhension de la lecture - manuels d'anglais langue étrangère. En Algérie, les manuels d'anglais langue étrangère sont la principale, sinon la seule source d'entrée lexicale pour les élèves. Cette recherche a pour but d'examiner les sept manuels en usage dans ces écoles afin d'en déterminer la couverture lexicale et la lisibilité. Le projet de recherche a été conçu pour i) compiler un corpus des manuel, ii) comparer leur couverture lexicale des manuels scolaires, et iii) d'évaluer leur lisibilité. La principale préoccupation était de déterminer si le niveau de lisibilité des manuels correspondait ou non à la plage de la couverture lexicale de l'étudiant.

Par ailleurs, les listes des items lexicaux ont été comparées à la *General Service List* et à la *Academic Word List* pour évaluer si les manuels scolaires sont à même de fournir un vocabulaire suffisant, utile et approprié. Un autre objectif était de fournir aux enseignants d'anglais et aux pédagogues en général un moyen (logiciel) pour comparer les niveaux de vocabulaire des textes de lecture destinés aux apprenants algériens de l'anglais afin d'en déterminer les niveaux de lisibilité et de vocabulaire, et quel vocabulaire supplémentaire était requis pour ces apprenants atteignent le taux de 95% de compréhension.

La méthodologie adoptée pour explorer la couverture lexicale a été caractérisée par une approche multi-instrument impliquant des programmes informatique. Les ensembles des manuels scolaires ont été traités pour générer des listes de mots qui y figurent.

Les résultats ont montré que tous les manuels en usage ont aussi bien un faible niveau de couverture lexicale et de lisibilité, les plaçant au niveau de frustration. Hormis les trois premiers manuels, il existe une divergence totale en termes de couverture lexicale entre les quatre autres manuels vu que le taux de vocabulaire commun à travers les manuels scolaires est très faible.

La comparaison de la couverture lexicale des sept manuels aux listes de vocabulaire ont révélé que les élèves algériens ne sont pas exposés a un vocabulaire suffisant, utile et approprié, étant donne qu'ils sont exposés à une faible proportion de mots de haute fréquence. L'étude se termine avec des implications et des recommandations quant à la façon de remédier au problème.