

L2 Competence of Young Language Learners in
Science and Arts CLIL and EFL Instruction
Contexts.
A Longitudinal Study.

Anna Vallbona González

DOCTORAL DISSERTATION

SUPERVISED BY: Dr Elisabet Pladevall Ballester

Doctorat en Filologia Anglesa

Departament de Filologia Anglesa i de Germanística

Facultat de Filosofia i Lletres

Universitat Autònoma de Barcelona

To Conor, Laura and Emma

ACKNOWLEDGEMENTS

First and foremost, I would like to thank my research supervisor, Dr Elisabet Pladevall, for her constant support throughout my PhD. She has been an exceptional tutor from beginning to end. Her dedication, patience and time, often out of hours, her critical and thoughtful comments on my drafts as well as her support throughout the whole process, have been very much appreciated. She has been a warmly encouraging force, pushing me to complete this project.

I also owe my deepest gratitude to the late Mia Victori, who gave me invaluable assistance and advice at the beginning of this project.

I also would like to show my appreciation for the University in Vic in allowing me a period of paid leave in order to finish my PhD. A special mention has to go to the head of my department, Dr Maica Bernal for the extra time that she made available by being understandable with my timetable.

A huge thanks also to Oliver Valero, in the Servei d'Estadística in the UAB, for his invaluable help in creating the statistics model that I was able to use to analyse my data and his brilliant educational instructions as to the interpretation of the statistical results.

It goes without saying that this project would not have been possible without the cooperation and support of a number of other people, to whom I would like to express my everlasting gratitude. First of all, I am obliged to the members of the CLIL-SLA Project: Elisabet Pladevall, Teresa Navés, Alexandra Vraciu, Anna Bret and Marc Miret. I would also like to thank Mary Louise Walsh, Adriana Rodriguez and Berta Torras for their help at data collection times and for helping me to type the children's hand-written compositions.

I also feel extremely grateful to the schools and the teachers that participated in my project: *Escola Pare Coll* in Vic, *Escola Mare de Déu del Carme* in Cardona and *Escola*

Mare de Déu del Roser in Guissona. A big thanks too to the teachers Mariona Fletas, Mercè Morales and Rosa Maria Morros for their patience, cooperation and helpful arrangements made at data collection times and especially for allowing me to do class observations during their CLIL sessions. A project like this would not have been possible without the cooperation of the informants. I owe a great deal to the pupils in the schools for taking part in the tests and answering my questionnaires, as well as for sharing their thoughts and their lives with me in their compositions.

I am greatly indebted to Carme Florit from Escola Santa Anna in Mataró, and Anna Bret from Escola Paidos, in Sant Fruitós de Bages, for allowing me to pilot the tests before the beginning of my project.

A special mention must also go to my colleagues Sarah Khan, Àngels Pinyana, Núria Medina, Àngel Raluy, Laura Vilardell and Toni Portell for their encouraging comments and their support at all times.

Finally, I would like to express my sincere gratitude to my family and friends, who supported me throughout my study, bore with me throughout my emotional ups and downs, and pushed me to the final completion of this research project. A big thanks to Conor, Laura and Emma for their comprehension and patience, especially during the final, strenuous stages of this project.

TABLE OF CONTENTS

ACKNOWLEDGEMENTS	vii
TABLE OF CONTENTS	xi
LIST OF TABLES	xix
CHAPTER 1 Introduction and Research Questions	1
1.1 Content and Language Integrated Learning (CLIL)	1
1.2 Justification for the Study	2
1.3 Research questions	6
1.4 Organisation of the Thesis	11
CHAPTER 2 Content and Language Integrated Learning (CLIL)	15
2.1 CLIL as a European Construct	15
2.2 The Language Situation in Catalonia. The Emergence of CLIL	23
2.3 Successful CLIL Programmes. Characteristics and Outcomes	27
2.3.1 Insights into CLIL success factors	27
2.3.2 CLIL Research in Europe	38
2.3.3 CLIL in Spain. An overview of research	49
2.3.4 CLIL Research in Catalonia	56
2.4 The Implementation of CLIL in Primary Education in Catalonia	59
CHAPTER 3 Young Language Learners	63
3.1 Young Learners. Definition and Characteristics	64
3.2 The Child as a Language Learner. Influential Factors	68
3.2.1 Age and the Critical Period Hypothesis	69
3.2.2 Motivation and Aptitude in Young Language Learners	72
3.2.3 The learning context	76
3.3 Review of some Psycholinguistic Principles in SLA	78
3.3.1 Exposure to Input. The Input Hypothesis	79
3.3.2 The processing of Meaning	80

3.3.3 The processing of Form. The Noticing Hypothesis	82
3.3.4 Language Production. The Output Hypothesis.....	85
3.4 Learning Foreign Languages at an Early Age.....	87
3.4.1 Early Foreign Language Learning Programmes.....	88
3.4.2 The Teacher Factor in Early Language Learning.....	91
3.4.3 Learning a Foreign Language in the School Context	97
3.4.4 Learning Languages through Content. Content-based Models	101
CHAPTER 4 Method.....	109
4.1 Design of the study.....	109
4.2 School Contexts and Participants.....	109
4.2.1 School A.....	111
4.2.1.1 Control Group (School A)	111
4.2.1.2 CLIL group (School A)	113
4.2.2 School B.....	114
4.2.2.1. Control group (School B)	114
4.2.2.2 CLIL group (School B)	115
4.2.3. School C.....	116
4.2.3.1 Control group (School C)	116
4.2.3.2 CLIL group (School C)	117
4.2.4 Summary of Participants	118
4.2.5 CLIL Teachers.....	120
4.3. Data Collection times and Instruments.....	124
4.3.1 Data Collection Times	124
4.3.2 Data Collection Instruments and procedure	131
4.3.2.1 Language Proficiency Tests.....	131
4.3.2.1.1 Listening tests	134
4.3.2.1.2 Reading tests	135
4.3.2.1.3 Written composition.....	136

4.3.2.2 Students' Background Questionnaire.....	138
4.3.2.3 CLIL Teachers' Background Questionnaire.....	138
4.3.2.4 CLIL Teachers' Opinion Questionnaires.....	139
4.3.2.5 Class Observations.....	140
4.3.3 Writing Measures.....	142
4.4 Data analysis and statistical procedures.....	149
CHAPTER 5 Results.....	153
5.1 CLIL Science results.....	154
5.1.1 Listening test results.....	155
5.1.2 Reading test results.....	160
5.1.2.1 Reading question 1.....	164
5.1.2.2 Reading question 2.....	165
5.1.2.3 Reading question 3.....	167
5.1.2.4 Reading question 4 (Comprehension).....	169
5.1.2.5 Reading question 4 (Accuracy).....	170
5.1.3 Summary of Science Listening and Reading Results.....	172
5.1.3.1 Summary of Listening and Reading Achievement Results.....	172
5.1.3.2 Summary of Listening and Reading Improvement Results.....	174
5.1.4 Writing.....	176
5.1.4.1 Fluency.....	177
5.1.4.1.1 Total Number of Words.....	177
5.1.4.1.2 Total Number of Words in English.....	179
5.1.4.1.3 Ratio: Total Number of Words in English/Total Number of Words.....	182
5.1.4.1.4 Total Number of Units.....	184
5.1.4.2. Accuracy.....	188
5.1.4.2.1 Total Number of Error Free Units.....	189
5.1.4.2.2 Ratio: Total Number of Error Free Units/Total Number of Units.....	192
5.1.4.3 Complexity.....	196

5.1.4.3.1 Lexical Complexity	196
5.1.4.3.1.1 Total Number of Lexical Verbs.....	197
5.1.4.3.1.2 Ratio: Total Number of Lexical Verbs/ Total Number of words in English	200
5.1.4.3.1.3 Total Number of Adjectives	202
5.1.4.3.1.4 Ratio: Total Number of Adjectives/Total Number of Words in English	206
5.1.4.3.2 Syntactic Complexity	208
5.1.4.3.2.1 Instances of Coordinated Units.....	208
5.1.4.3.2.2 Instances of Subordinate Units	211
5.1.5 Summary of CLIL Science Writing Results.....	215
5.1.5.1 Summary of Writing Achievement Results	215
5.1.5.2 Summary of Writing Improvement Results	219
5.2 CLIL Arts & Crafts results	222
5.2.1 Listening test results	223
5.2.2 Reading test results	227
5.2.2.1 Reading question 1	231
5.2.2.2 Reading question 2.....	233
5.2.2.3 Reading question 3	235
5.2.2.4 Reading question 4 (Comprehension)	236
5.2.2.5 Reading question 4 (Accuracy).....	238
5.2.3 Summary of Arts & Crafts Listening and Reading Results.....	240
5.2.3.1 Summary of Listening and Reading Achievement Results	240
5.2.3.2 Summary of Listening and Reading Improvement Results	241
5.2.4 Writing	243
5.2.4.1 Fluency	243
5.2.4.1.1 Total Number of Words.....	244
5.2.4.1.2 Total Number of Words in English	246
5.2.4.1.3 Ratio: Total Number of Words in English/Total Number of Words	250

5.2.4.1.4 Total Number of Units	252
5.2.4.2. Accuracy	255
5.2.4.2.1 Total Number of Error Free Units	256
5.2.4.2.2 Ratio: Total Number of Error Free Units/Total Number of Units	259
5.2.4.3 Complexity	261
5.2.4.3.1 Lexical Complexity	261
5.2.4.3.1.1 Total Number of Lexical Verbs	261
5.2.4.3.1.2 Ratio: Total Number of Lexical Verbs/Total Number of Words in English	265
5.2.4.3.1.3 Total Number of Adjectives	266
5.2.4.3.1.4 Ratio: Total Number of Adjectives/Total Number of Words in English	270
5.2.4.3.2. Syntactic Complexity	271
5.2.4.3.2.1 Instances of Coordinated Units	272
5.2.4.3.2.2 Instances of Subordinate Units	275
5.2.5 Summary of CLIL Arts & Crafts Writing Results	278
5.2.5.1 Summary of Writing Achievement Results	278
5.2.5.2 Summary of Writing Improvement Results	282
5.3 CLIL Science results vs CLIL Arts & Crafts results	286
5.3.1 Listening test results	286
5.3.2 Reading test results	289
5.3.2.1 Reading question 1	292
5.3.2.2 Reading question 2	293
5.3.2.3 Reading question 3	294
5.3.2.4 Reading question 4 (Comprehension)	295
5.3.2.5. Reading question 4 (Accuracy)	297
5.3.3 Summary of Listening and Reading Results	298
5.3.4 Writing	300
5.3.4.1 Fluency	300

5.3.4.1.1 Total Number of Words	300
5.3.4.1.2 Total Number of Words in English	301
5.3.4.1.3 Ratio: Total Number of Words in English/Total Number of Words	304
5.3.4.1.4 Total Number of Units	305
5.3.4.2. Accuracy	308
5.3.4.2.1 Total Number of Error Free Units	308
5.3.4.2.2 Ratio: Total Number of Error Free Units/Total Number of Units	311
5.3.4.3 Complexity	313
5.3.4.3.1 Lexical Complexity	313
5.3.4.3.1.1 Total Number of Lexical Verbs	314
5.3.4.3.1.2 Ratio: Total Number of Lexical Verbs/Total Number of Words in English	316
5.3.4.3.1.3 Total Number of Adjectives	319
5.3.4.3.1.4 Ratio: Total Number of Adjectives/Total Number of Words in English	321
5.3.4.3.2 Syntactic Complexity	324
5.3.4.3.2.1 Instances of Coordinated Units	324
5.3.4.3.2.2 Instances of Subordinate Units	327
5.3.5 Summary of CLIL Science vs. CLIL Arts & Crafts Achievement Results	329
CHAPTER 6 Discussion	335
6.1 Achievement and improvement in Listening skills	335
6.2 Achievement and improvement in Reading skills	344
6.3 Achievement and improvement in Writing: Complexity, Accuracy and Fluency	349
6.3.1 Achievement and improvement results in Fluency	350
6.3.2 Achievement and improvement results in Accuracy	357
6.3.3 Achievement and improvement results in Complexity	360
CHAPTER 7 Conclusions	365
7.1 Main findings and concluding remarks	365
7.2 Limitations and further research	377

REFERENCES	379
APPENDICES	405
Appendix A. Listening Test A1	407
Appendix B. Listening Test B1	408
Appendix C. Reading Test A1	409
Appendix D. Reading Test B1	410
Appendix E. Reading Test C	411
Appendix F. Reading Test D1	412
Appendix G. Reading Test A2	414
Appendix H. Reading Test B2.....	415
Appendix I. Reading Test D2	416
Appendix J. Students' Background Questionnaire	418
Appendix K. CLIL Teachers' Background Questionnaire.....	421
Appendix L. CLIL Teachers' Opinion Questionnaire	426
Appendix M. CLIL Class Observation Protocol	427
Appendix N. Sample of Writings from High and Low achievers.....	428

LIST OF TABLES

- Table 2.1 Summary of CLIL findings in Scandinavia
Table 2.2 Summary of CLIL findings in Austria, The Netherlands, Switzerland and Germany
Table 2.3 Summary of CLIL findings in Poland, Hungary, the Czech Republic and Italy
Table 4.1 Number of participants per group in each school
Table 4.2 Summary of participants by country of origin
Table 4.3 Summary of the languages spoken at home
Table 4.4 Number of hours of curricular exposure for the Control and CLIL groups
Table 4.5 Data collection times and hours of English instruction. School A: Science students and their Control Group
Table 4.6 Data collection times and hours of instruction. Schools B and C: Arts & Crafts students and their Control Groups
Table 4.7 Group comparisons: School A
Table 4.8 Group comparisons: Schools B and C
Table 4.9 Hours of previous exposure to English at school up to 5th grade
Table 4.10 Tests, skills and level
Table 4.11 Writing measures
CLIL Science students vs. Control Group
Table 5.1 Listening achievement results
Table 5.2 Percentage of improvement in Listening
Table 5.3 Listening achievement results: Group/High achievers interaction
Table 5.4 Listening achievement results: Group/Low achievers interaction
Table 5.5 Percentage of improvement in Listening. Group/Proficiency level interaction
Table 5.6 Reading achievement results
Table 5.7 Percentage of improvement in Reading
Table 5.8 Reading achievement results. Group/High achievers interaction
Table 5.9 Reading achievement results. Group/Low achievers interaction
Table 5.10 Percentage of improvement in Reading. Group/Proficiency level interaction
Table 5.11 Reading achievement results: Reading Question1
Table 5.12 Percentage of improvement: Reading Question1
Table 5.13 Reading achievement results: Reading Question 2
Table 5.14 Percentage of improvement: Reading Question 2
Table 5.15 Reading achievement results: Reading Question 3
Table 5.16 Percentage of improvement: Reading Question 3
Table 5.17 Reading achievement results: Reading Question 4 (Comprehension)
Table 5.18 Percentage of improvement: Reading Question 4 (Comprehension)
Table 5.19 Reading achievement results: Reading Question 4 (Accuracy)
Table 5.20 Percentage of improvement: Reading Question 4 (Accuracy)
Table 5.21 Summary of Listening and Reading achievement results
Table 5.22 Summary of Reading achievement results: individual reading questions
Table 5.23 Summary of Listening achievement results. Group/Proficiency level interaction
Table 5.24 Summary of Reading achievement results. Group/Proficiency level interaction
Table 5.25 Summary of Listening and Reading improvement results
Table 5.26 Summary of Reading improvement results: Reading Questions 1, 2 and 3
Table 5.27 Summary of Reading improvement results: Reading Question 4
Table 5.28 Summary of Listening improvement results. Group/Proficiency level interaction

Table 5.29 Summary of Reading improvement results. Group/Proficiency level interaction.

Table 5.30 Fluency achievement results: TNW

Table 5.31 Percentage of improvement in Fluency: TNW

Table 5.32 Fluency achievement results: TNWE

Table 5.33 Percentage of improvement in Fluency: TNWE

Table 5.34 Fluency achievement results: TNWE. Group/High achievers interaction

Table 5.35 Fluency achievement results: TNWE. Group/Low achievers interaction

Table 5.36 Percentage of improvement in Fluency: TNWE. Group/Proficiency level interaction

Table 5.37 Fluency achievement results: % TNWE/TNW

Table 5.38 Percentage of improvement in Fluency: % TNWE/TNW

Table 5.39 Fluency achievement results: TNU

Table 5.40 Percentage of improvement in Fluency: TNU

Table 5.41 Fluency achievement results: TNU. Group/High achievers interaction

Table 5.42 Fluency achievement results: TNWE. Group/Low achievers interaction

Table 5.43 Percentage of improvement in Fluency: TNU. Group/Proficiency level interaction

Table 5.44 Accuracy achievement results: TNEFU

Table 5.45 Percentage of improvement in accuracy: TNEFU

Table 5.46 Accuracy improvement results: TNEFU. Group/High achievers interaction

Table 5.47 Accuracy improvement results: TNEFU. Group/Low achievers interaction

Table 5.48 Percentage of improvement results: TNEFU. Group /proficiency level interaction

Table 5.49 Accuracy achievement results: % TNWE/TNW

Table 5.50 Percentage of improvement in accuracy: % TNEFU/TNU

Table 5.51 Accuracy achievement results: %TNEFU/TNU. Group/High achievers interaction

Table 5.52 Accuracy achievement results: %TNEFU/TNU. Group/Low achievers interaction

Table 5.53 Percentage of improvement in accuracy: %TNEFU/TNU. Group/Proficiency level interaction

Table 5.54 Lexical Complexity achievement results: TNLV

Table 5.55 Percentage of improvement in Lexical Complexity: TNLV. Group/Proficiency level interaction

Table 5.56 Lexical Complexity achievement results: TNLV. Group/High achievers interaction

Table 5.57 Lexical Complexity achievement results: TNLV. Group/Low achievers interaction

Table 5.58 Percentage of improvement in Lexical Complexity: TNLV. Group/Proficiency level interaction

Table 5.59 Lexical Complexity achievement results: % TNLV/TNWE

Table 5.60 Percentage of improvement in Lexical Complexity: %TNLV/TNWE.

Table 5.61 Lexical Complexity achievement results: TNAdj

Table 5.62 Percentage of improvement in Lexical Complexity: TNA

Table 5.63 Lexical Complexity achievement results: TNAdj. Group/High achievers interaction

Table 5.64 Lexical Complexity achievement results: TNAdj. Group/Low achievers interaction

Table 5.65 Percentage of improvement in Lexical Complexity:TNAdj. Group/Proficiency level interaction

Table 5.66 Lexical Complexity Achievement results. %TNAdj/TNWE

Table 5.67 Percentage of improvement in Lexical Complexity: %TNAdj/TNWE

Table 5.68 Syntactic Complexity achievement results: ICU

Table 5.69 Percentage of improvement in Syntactic Complexity: ICU

Table 5.70 Syntactic Complexity achievement results. Group/High achievers interaction

Table 5.71 Syntactic Complexity achievement results. Group/Low achievers interaction

Table 5.72 Percentage of improvement in Syntactic Complexity. Group/Proficiency level interaction

Table 5.73 Syntactic Complexity achievement results: ISU

Table 5.74 Percentage of improvement in Syntactic Complexity: ISU

Table 5.75 Syntactic Complexity achievement results: ISU. Group/High achievers interaction

Table 5.76 Syntactic Complexity achievement results: ISU. Group/Low achievers interaction

Table 5.77 Percentage of Syntactic Complexity improvement: ISU. Group/Proficiency level interaction

Table 5.78 Summary of Fluency achievement results

Table 5.79 Summary of Fluency achievement results. Group/Proficiency level interaction

Table 5.80 Summary of Accuracy achievement results

Table 5.81 Summary of Accuracy achievement results: TNEFU. Interaction Group/ Proficiency level interaction

Table 5.82 Summary of Lexical Complexity achievement results.

Table 5.83 Summary of Lexical Complexity achievement results. Group/Proficiency level interaction

Table 5.84 Summary of Syntactic Complexity achievement results

Table 5.85 Summary of Syntactic Complexity achievement results. Group/Proficiency level interaction

Table 5.86 Summary of Fluency improvement results

Table 5.87 Summary of Fluency improvement results. Group/Proficiency Level interaction

Table 5.88 Summary of Accuracy improvement results

Table 5.89 Summary of Accuracy improvement results. Group/Proficiency level interaction

Table 5.90 Summary of Lexical Complexity improvement results

Table 5.91 Summary of Lexical Complexity improvement results. Group/ Proficiency level interaction

Table 5.92 Summary of Syntactic Complexity improvement results

Table 5.93 Summary of Syntactic Complexity improvement results. Group/ Proficiency level interaction

CLIL Arts & Crafts students vs. Control Group

Table 5.94 Listening achievement results

Table 5.95 Improvement percentages in listening

Table 5.96 Listening achievement results: Group/High achievers interaction

Table 5.97 Listening achievement results: Group/Low achievers interaction

Table 5.98 Percentage of improvement in listening. Group/Proficiency level interaction

Table 5.99 Reading achievement results

Table 5.100 Percentage of improvement in reading

Table 5.101 Reading achievement results: Group/High achievers interaction

Table 5.102 Reading achievement results: Group /Low achievers interaction

Table 5.103 Percentage of improvement in reading. Group/Proficiency Level interaction

Table 5.104 Reading achievement results: Reading Question 1

Table 5.105 Percentage of improvement. Reading Question1

Table 5.106 Reading achievement results: Reading Question 2

Table 5.107 Percentage of improvement. Reading Question 2

Table 5.108 Reading achievement results. Reading Question 3

Table 5.109 Percentage of improvement. Reading Question 3

Table 5.110 Reading achievement results. Reading Question 4 (Comprehension)

Table 5.111 Percentage of improvement. Reading Question 4 (Comprehension)

Table 5.112 Reading achievement results. Reading Question 4 (Accuracy)

Table 5.113 Percentage of improvement. Reading Question 4 (Accuracy)

Table 5.114 Summary of Listening and Reading achievement results

Table 5.115 Summary of Reading achievement results: individual reading questions

Table 5.116 Summary of Listening achievement results. Group/Proficiency level interaction

Table 5.117 Summary of Reading achievement results. Group/Proficiency level interaction

Table 5.118 Summary Listening and Reading Improvement Results

Table 5.119 Summary of Reading improvement results: reading questions 1, 2, and 3

Table 5.120 Summary of Reading improvement results: reading question 4

Table 5.121 Summary of the Listening improvement results. Group/Proficiency level interaction

Table 5.122 Summary of the Interaction Group/Proficiency. Reading improvement results

Table 5.123 Fluency achievement Results: TNW.

Table 5.124 Percentage of improvement in Fluency: TNW

Table 5.125 Fluency achievement results: TNW. Group/High achievers interaction

Table 5.126 Fluency achievement results: TNW. Group/Low achievers interaction

Table 5.127 Fluency achievement results: TNWE

Table 5.128 Percentage of improvement in Fluency: TNWE

Table 5.129 Fluency achievement results: TNWE. Group/High achievers

Table 5.130 Fluency achievement results: TNWE. Group/Low achievers

Table 5.131 Percentages of improvement in Fluency: TNWE. Group/Proficiency level interaction

Table 5.132 Fluency achievement results: %TNWE/TNW

Table 5.133 Percentage of improvement in Fluency: %TNWE/TNW

Table 5.134 Fluency achievement results: TNU

Table 5.135 Percentage of improvement in Fluency: TNU

Table 5.136 Fluency achievement results: TNU. Group/High achievers interaction

Table 5.137 Fluency achievement results: TNU. Group/Low achievers interaction

Table 5.138 Percentages of improvement in Fluency: TNU. Group/Proficiency level interaction

Table 5.139 Accuracy achievement results: TNEFU

Table 5.140 Percentage of improvement in Accuracy: TNEFU

Table 5.141 Accuracy achievement results: TNEFU. Group/High achievers interaction

Table 5.142 Accuracy achievement results: TNEFU. Group/Low achievers interaction

Table 5.143 Percentage of improvement in Accuracy: TNEFU. Group/Proficiency level interaction

Table 5.144 Accuracy improvement results: % TNEFU/TNU

Table 5.145 Percentage of improvement in Accuracy: %TNEFU/TNU

Table 5.146 Lexical Complexity achievement results: TNLV

Table 5.147 Percentage of improvement in Lexical Complexity: TNLV

Table 5.148 Lexical Complexity achievement results: TNLV. Group/High achievers interaction

Table 5.149 Lexical Complexity achievement results: TNLV. Group/Low achievers interaction

Table 5.150 Percentage of improvement: TNLV. Group/Proficiency level interaction

Table 5.151 Lexical Complexity achievement results: %TNLV/TNWE

Table 5.152 Percentage of improvement in Lexical Complexity: % TNLV/TNWE

Table 5.153 Lexical Complexity achievement results: TNAdj

Table 5.154 Percentage of improvement in Lexical Complexity: TNAdj

Table 5.155 Lexical Complexity achievement results: TNAdj. Group/High achievers interaction

Table 5.156 Lexical Complexity achievement results: TNAdj. Group/Low achievers interaction

- Table 5.157 Percentage of improvement in Lexical Complexity improvement results: TNAdj
Group/Proficiency level interaction
- Table 5.158 Lexical Complexity achievement results; %TNAdj/TNWE
- Table 5.159 Percentage of improvement in Lexical Complexity: % TNAdj/TNWE
- Table 5.160 Syntactic Complexity achievement results: ICU
- Table 5.161 Percentage of improvement in Syntactic Complexity: ICU
- Table 5.162 Syntactic Complexity achievement results: ICU. Group/High achievers
- Table 5.163 Syntactic Complexity achievement results: ICU. Group/Low achievers
- Table 5.164 Percentage of improvement results: ICU. Group/Proficiency level interaction
- Table 5.165 Syntactic Complexity achievement results: ISU
- Table 5.166 Percentage of improvement in syntactic complexity: ISU
- Table 5.167 Syntactic Complexity achievement results: ISU. Group/High achievers interaction
- Table 5.168 Syntactic Complexity achievement results: ISU. Group/Low achievers interaction
- Table 5.169 Percentage of improvement in Syntactic Complexity: ISU. Group/Proficiency level
interaction
- Table 5.170 Summary of Fluency achievement results
- Table 5.171 Summary of Fluency achievement results: TNWE, TNU. Group/Proficiency level
interaction
- Table 5.172 Summary of Accuracy achievement results.
- Table 5.173 Summary of Accuracy achievement results: TNEFU. Group/Proficiency level
interaction
- Table 5.174 Summary of Lexical Complexity achievement results
- Table 5.175 Summary of Lexical Complexity achievement results: TNLV, TNAdj
Group/Proficiency level interaction
- Table 5.176 Summary of Syntactic Complexity achievement results
- Table 5.177 Summary of Syntactic Complexity achievement results: ICU, ISU.
Group/Proficiency level interaction
- Table 5.178 Summary of Fluency improvement results
- Table 5.179 Summary of Fluency improvement results: TNWE, TNU. Group/Proficiency level
interaction
- Table 5.180 Summary of Accuracy improvement results
- Table 5.181 Summary of Accuracy improvement results: TNWE. Group/Proficiency level
interaction
- Table 5.182 Summary of Lexical Complexity improvement results
- Table 5.183 Summary of Lexical Complexity improvement results. Group/Proficiency level
interaction
- Table 5.184 Summary of Syntactic Complexity improvement results
- Table 5.185 Summary of Syntactic Complexity improvement results. Group/Proficiency level
interaction
- CLIL Science students vs. CLIL Arts & Crafts students**
- Table 5.186 Listening achievement results
- Table 5.187 Listening achievement results: Group/High achievers interaction
- Table 5.188 Listening achievement results: Group/Low achievers interaction
- Table 5.189 Reading achievement results
- Table 5.190 Reading achievement results: Group/High achievers interaction
- Table 5.191 Reading achievement results: Group/Low achievers interaction
- Table 5.192 Reading achievement results: Reading Question1
- Table 5.193 Reading achievement results: Reading Question 2

- Table 5.194 Reading achievement results: Reading Question 3
- Table 5.195 Reading achievement results: Reading Question 4 (Comprehension)
- Table 5.196 Reading achievement results: Reading Question 4 (Accuracy)
- Table 5.197 Summary of Listening and Reading achievement results
- Table 5.198 Summary of Reading achievement results: individual reading questions
- Table 5.199 Summary of Listening achievement results. Group/Proficiency level interaction
- Table 5.200 Summary of Reading achievement results. Group/Proficiency level interaction
- Table 5.201 Fluency achievement results: TNW
- Table 5.202 Fluency achievement results: TNWE
- Table 5.203 Fluency achievement results: TNWE. Group/High achievers interaction
- Table 5.204 Fluency achievement results: TNWE. Group/Low achievers interaction
- Table 5.205 Fluency achievement results: % TNWE/TNW.
- Table 5.206 Fluency achievement results: TNU
- Table 5.207 Fluency achievement results: TNU. Group/High achievers interaction
- Table 5.208 Fluency achievement results: TNUGroup/Low achievers interaction
- Table 5.209 Accuracy achievement results: TNEFU
- Table 5.210 Accuracy achievement results: TNEFU. Group/High achievers interaction
- Table 5.211 Accuracy achievement results: TNEFU. Group/Low achievers interaction
- Table 5.212 Accuracy achievement results: % TNWE/TNW
- Table 5.213 Accuracy achievement results: % TNWE/TNW. Group/High achievers interaction
- Table 5.214 Accuracy achievement results: % TNWE/TNW. Group/Low achievers interaction
- Table 5.215 Lexical Complexity achievement results: TNLV
- Table 5.216 Lexical Complexity achievement results: TNLV. Group/High achievers interaction
- Table 5.217 Lexical Complexity achievement results: TNLV. Group/Low achievers interaction
- Table 5.218 Lexical Complexity achievement results: %TNLV/TNWE
- Table 5.219 Lexical Complexity achievement results: %TNLV/TNWE. Group/High achievers interaction
- Table 5.220 Lexical Complexity achievement results: %TNLV/TNWE. Group/Low achievers interaction
- Table 5.221 Lexical Complexity achievement results: TNAdj
- Table 5.222 Lexical Complexity achievement results: TNAdj. Group/High achievers interaction
- Table 5.223 Lexical Complexity achievement results: TNAdj. Group/Low achievers interaction
- Table 5.224 Lexical Complexity achievement results: %TNAdj/TNWE
- Table 5.225 Lexical Complexity achievement results: %TNAdj/TNWE. Group/High achievers interaction
- Table 5.226 Lexical Complexity achievement results: %TNAdj/TNWE. Group/Low achievers interaction
- Table 5.227 Syntactic Complexity achievement results: ICU
- Table 5.228 Lexical Complexity achievement results: ICU. Group/High achievers interaction
- Table 5.229 Syntactic Complexity achievement results: ICU. Group/Low achievers interaction
- Table 5.230 Syntactic Complexity achievement results: ISU
- Table 5.231 Syntactic Complexity achievement results: ISU. Group/ High achievers interaction
- Table 5.232 Syntactic Complexity achievement results: ISU. Group/Low achievers interaction
- Table 5.233 Summary of Fluency achievement results. Science vs. Arts & Crafts
- Table 5.234 Summary of Fluency achievement results. Science vs. Arts & Crafts. Group Proficiency level interaction
- Table 5.235 Summary of Accuracy achievement results. Science vs. Arts & Crafts

Table 5.236 Summary of Accuracy achievement results TNEFU. Science vs. Arts & Crafts.
Group /Proficiency level interaction

Table 5.237 Summary of Lexical Complexity achievement results. Science vs. Arts & Crafts.

Table 5.238 Summary of Lexical Complexity achievement results: TNLV, TNAdj. Science vs.
Arts & Crafts Group/Proficiency level

Table 5.239 Summary of Syntactic Complexity achievement results

Table 5.240 Summary of Syntactic Complexity achievement results: ICU, ISU. Science vs. Arts
& Crafts Group/Proficiency level

LIST OF FIGURES

Figure 3.1 A Language-content continuum for teaching languages to young learners.

Figure 4.1 Country of origin. Control Group A

Figure 4.2 Languages spoken at home. Control Group A

Figure 4.3 Country of origin. CLIL Group A

Figure 4.4 Languages spoken at home. CLIL Group A

Figure 4.5 Country of origin. Control Group B

Figure 4.6 Languages spoken at home. Control Group B

Figure 4.7 Country of origin. CLIL Group B

Figure 4.8 Languages spoken at home. CLIL Group B

Figure 4.9 Country of origin. Control Group C

Figure 4.10 Languages spoken at home. Control Group C

Figure 4.11 Country of origin. CLIL Group C

Figure 4.12 Languages spoken at home. CLIL Group C

CLIL Science Students vs. Control Group

Figure 5.1 Listening improvement results

Figure 5.2 Listening improvement results: Group/Proficiency level interaction.

Figure 5.3: Reading improvement results

Figure 5.4: Reading improvement results: Group/Proficiency level interaction

Figure 5.5 Reading improvement results: Reading Question 1

Figure 5.6 Reading improvement results: Reading Question 2

Figure 5.7 Reading improvement results: Reading Question 3

Figure 5.8 Reading improvement results: Reading Question 4 (Comprehension)

Figure 5.9 Reading improvement results: Reading Question 4 (Accuracy)

Figure 5.10 Fluency improvement results: TNW

Figure 5.11 Fluency improvement results: TNWE

Figure 5.12 Fluency improvement results: TNWE. Group/Proficiency level interaction

Figure 5.13 Fluency improvement results: %TNWE/TNW

Figure 5.14 Fluency improvement results: TNU

Figure 5.15 Fluency improvement results: TNU. Group/Proficiency level interaction

Figure 5.16 Accuracy improvement results: TNEFU

Figure 5.17 Accuracy improvement results: TNEFU. Group/Proficiency level interaction

Figure 5.18 Accuracy improvement results: % TNEFU/TNU

Figure 5.19 Accuracy improvement results: %TNEFU/TNU. Group/Proficiency level interaction

Figure 5.20 Lexical Complexity improvement results: TNLV

Figure 5.21 Lexical Complexity improvement results: TNLV. Group/Proficiency level interaction

Figure 5.22 Lexical Complexity improvement results: %TNLV/TNWE

Figure 5.23 Lexical Complexity improvement results: TNAdj

Figure 5.24 Lexical Complexity improvement results: TNAdj. Group/Proficiency level interaction

Figure 5.25 Lexical Complexity improvement results: %TNAdj/TNWE

Figure 5.26 Syntactic Complexity improvement results: ICU

Figure 5.27 Syntactic Complexity improvement results. Group/Proficiency level interaction

Figure 5.28 Syntactic Complexity Improvement results: ISU

Figure 5.29 Syntactic Complexity improvement results: ISU. Group/Low achievers interaction

CLIL Arts & Crafts Students

Figure 5.30 Listening improvement results

Figure 5.31 Listening improvement results. Group/Proficiency level interaction

Figure 5.32 Reading improvement results

Figure 5.33 Reading improvement results: Group /Proficiency level interaction

Figure 5.34 Reading improvement results: Reading Question 1

Figure 5.35 Reading improvement results: Reading Question 2

Figure 5.36 Reading improvement results: Reading Question 3

Figure 5.37 Reading improvement results: Reading Question 4 (Comprehension)

Figure 5.38 Reading improvement results: Reading Question 4 (Accuracy)

Figure 5.39 Fluency improvement results: TNW

Figure 5.40 Fluency improvement results: TNWE.

Figure 5.41 Fluency improvement results: TNWE. Group/Proficiency level interaction

Figure 5.42 Fluency improvement results: % TNWE/TNW

Figure 5.43 Fluency improvement results: TNU

Figure 5.44 Fluency improvement results: TNU. Group/Proficiency level interaction

Figure 5.45 Accuracy improvement results: TNEFU

Figure 5.46 Accuracy improvement results: TNEFU. Group/Proficiency level interaction

Figure 5.47 Accuracy improvement results: %TNEFU/TNU

Figure 5.48 Lexical Complexity improvement results: TNLV

Figure 5.49 Lexical Complexity improvement results: TNLV. Group/ Proficiency level interaction

Figure 5.50 Lexical Complexity achievement results: %TNLV/TNWE

Figure 5.51 Lexical Complexity improvement results: TNAdj

Figure 5.52 Lexical Complexity improvement results:TNAdj. Group/Proficiency level interaction

Figure 5.53 Lexical Complexity improvement results: % TNAdj/TNWE

Figure 5.54 Syntactic Complexity improvement results: ICU

Figure 5.55 Syntactic Complexity improvement results: ICU. Group/Proficiency level interaction

Figure 5.56 Syntactic Complexity improvement results: ICU

Figure 5.57 Syntactic Complexity improvement results: ISU. Group/Proficiency level

CLIL Science vs. CLIL Arts & Crafts students

Figure 5.58 Listening achievement results

Figure 5.59 Listening achievement results: Group/Proficiency level interaction

Figure 5.60 Reading achievement results

Figure 5.61 Reading achievement results. Group/Proficiency level interaction

Figure 5.62 Reading achievement results: Reading Question1

Figure 5.63 Reading achievement results: Reading Question 2

Figure 5.64 Reading achievement results: Reading Question 3

Figure 5.65 Reading achievement results: Reading Question 4 (Comprehension)

Figure 5.66 Reading achievement results: Reading Question 4 (Accuracy)

Figure 5.67 Fluency Achievement results: TNW

Figure 5.68 Fluency achievement results: TNWE

- Figure 5.69 Fluency achievement results: TNWE. Group/Proficiency level interaction
- Figure 5.70 Fluency achievement results: % TNWE/TNW
- Figure 5.71 Fluency achievement results: TNU
- Figure 5.72 Fluency achievement results: TNU. Group/Proficiency level interaction
- Figure 5.73 Accuracy achievement results: TNEFU
- Figure 5.74 Accuracy achievement results: TNEFU. Group/Proficiency level interaction
- Figure 5.75 Accuracy achievement results: % TNEFU/TNU
- Figure 5.76 Accuracy achievement results: % TNEFU/TNWE. Group/Proficiency level interaction
- Figure 5.77 Lexical Complexity achievement results: TNLV
- Figure 5.78 Lexical Complexity achievement results: TNLV. Group/Proficiency level interaction
- Figure 5.79 Lexical Complexity achievement results: %TNLV/TNWE
- Figure 5.80 Lexical Complexity achievement results: %TNLV/TNWE. Group/Proficiency level interaction
- Figure 5.81 Lexical Complexity achievement results: TNAdj
- Figure 5.82 Lexical Complexity achievement results: TNAdj. Group/Proficiency level interaction
- Figure 5.83 Lexical Complexity achievement results: %TNAdj/TNWE
- Figure 5.84 Lexical Complexity achievement results: %TNAdj/TNWE. Group/Proficiency level interaction
- Figure 5.85 Syntactic Complexity achievement results: ICU
- Figure 5.86 Syntactic Complexity achievement results: ICU. Group/Proficiency level interaction
- Figure 5.87 Syntactic Complexity achievement results: ISU
- Figure 5.88 Syntactic Complexity achievement results: ISU. Group/ Proficiency level interaction

LIST OF ABBREVIATIONS

BE	Bilingual Education
BICS	Basic Interpersonal Communication Skills
CAF	Complexity, Accuracy and Fluency
CALP	Cognitive Academic Language Proficiency
CG	Control Group
CHP	Critical Period Hypothesis
CLIL	Content and Language Integrated Learning
ECML	European Council for Modern Foreign Languages
EFL	English as a Foreign Language
ESOL	English for Speakers of Other Languages
FL	Foreign Language
FLA	Foreign Language Aptitude
FLL	Foreign Language Learning
LAC	Languages across the Curriculum
LEP	Limited English Proficiency
LTM	Long Term Memory
PILE	Pla Integrat de Llengües Estrangeres
PISA	Programme for International Student Assessment
SCLILP	Successful CLIL Programmes
SIOP	Sheltered Instruction Observation Protocol
SLA	Second Language Acquisition
STM	Short Term Memory
VYL	Very Young Learners
WM	Working Memory
YL	Young Learners
YLE	Young Learner's Exam

YLL Young Language Learners

CHAPTER 1 Introduction and Research Questions

1.1 Content and Language Integrated Learning (CLIL)

Content and Language Integrated Learning, or CLIL, as it is called in English, refers to any educational approach in which a second/foreign language is used as a means of teaching the contents of a syllabus subject in the classroom. The use of a second/foreign language is intended to develop the necessary ability to be proficient in that language itself, while contents are being learned alongside. As Marsh (1994:27) defined it:

‘CLIL refers to a situation in which the subject matter or part of the subject matter is taught via a foreign language with a two-fold objective: the learning of those contents and the simultaneous learning of a foreign language.’

The fundamental idea behind this type of content-based teaching is that giving priority to content as well as the opportunity it offers the students to speak and think in an authentic, significant and relevant way in a second language, not only improves the competence of the students in this language, but also enriches their cognitive development and cultural growth.

Interest in CLIL in Europe has grown since the 1990s due to the actions of the continent’s own institutions and their commitment to a multilingual Europe. According to the *White Paper on Teaching and Learning* (1995), a document issued by the European Community, one of the main objectives of European policy in education is for European citizens to become proficient in three languages: their first language, a language of international communication, and a ‘personal adoptive language’. According to the Euridyce Report (2006), CLIL initiatives, to a greater or lesser degree, exist in almost all European countries. Their implementation and growth have taken different courses over the years, ranging from child education to university teaching, in

the various member states of the European Union. There has also been an enormous variation in the number of models used, as well as in the level of exposure to the language involved, which ranges from a much-reduced period of time to the implementation of whole subjects completely in a second/foreign language. All of this is due to the expansion of methods, materials and curriculum approaches resulting from different educational situations in different countries. CLIL is, therefore, seen as a ‘flexible system which responds to a very wide range of situational and contextual demands’ (Coyle, 2005: 23); it is an educational approach that integrates content and language, by learning a content subject through the medium of a foreign language. It is efficient as it integrates language and content into the broad curriculum, which is especially interesting in bilingual contexts, where two languages already need to be accommodated (Muñoz and Nussbaum, 1997). Research shows that CLIL also creates conditions for naturalistic language learning and, therefore, develops communicative competence (Met, 1998; Marsh and Langé, 1999; Marsh 2000). It also involves the learner in using the language *of* learning, *for* learning and *through* learning (Coyle, 2000) and it further increases motivation and interest levels in language learning (Grabe and Stoller, 1997; Pavesi, Bertochi, Hofmannová, Kazianka and Langé, 2001) by diversifying methods and forms of classroom teaching and learning

1.2 Justification for the Study

In the last few years, the need to speak English has emerged from advances in the speed of communication systems and mobility, which have spurred globalization and have led to an exponential growth of the use of English as a lingua franca. Despite the resolution of the European Parliament (2009), which stated that it would be wrong for the European Union to restrict itself to a single main foreign language, learning English has become a necessity within the European Union. European educational

systems have attached increasing importance to the learning of foreign languages, mainly English, and they spare no effort to ensure that the students' level in this foreign language improves all the time. The increasing demand to improve language skills has prompted some countries in Europe, including Spain, to opt for an early start in Foreign Language Learning (FLL) at school in the belief that lowering the starting age of the learners will make a significant difference in the long term. Due to this trend of bringing forward the starting age for foreign language learning at school, young language learners have recently received increasing attention (Rixon, 2000; Nikolov, 2009; Hasselgreen, Kalédaité, Martín, Pižorn, 2011). In some schools, this trend has coincided with the emergence of CLIL programmes at primary level in the belief that, by combining EFL and CLIL, the L2 competence of the learners would improve. CLIL has become commonplace in most European systems because it is thought to improve overall language competence in the target language. Nevertheless, the European Council for Modern Languages has voiced its concern about the fact that 'the implementation of CLIL is outpacing a measured debate about the impact of its implementation on students and teachers' (ECML, 2007:11). Research at the *micro level*, with a clear focus on CLIL participants (teachers and students) has not yet reached a high level of interest (Dalton-Puffer and Smith, 2007; Pérez-Cañado, 2012). There is a need to evaluate language competence in CLIL, as well as the learners' conceptual skills and cognitive skills development, and the way in which classroom interaction affects the outcome of the approach, in order to be able to compare CLIL with traditional EFL instruction so as to be able to support the benefits of this approach and, at the same time, to discover the strengths and weaknesses of the implementation of CLIL in each educational context. As Coyle (2007) points out, there is no single blueprint that can be applied in the same way in different countries; hence research in each particular context is needed.

In spite of all the general claims about the beneficial effects of CLIL on language acquisition, as CLIL is a relatively innovative approach in education, there is still a ‘time-lag between practice and research-based validation’ (Langé, 2007: 352). In recent years, a growing body of relevant CLIL literature has emerged across Europe trying to provide secure evidence in different contexts and across different levels for the purported benefits of this approach. However, despite the success of CLIL programmes in most European education systems, academic research on CLIL still remains embryonic (Navés and Victori, 2010). A plethora of authors have claimed the need for further research in different areas in order to paint a clearer, reliable picture of the effects of CLIL: Ruiz de Zarobe and Lasagabaster (2010) have claimed that longitudinal studies on the effects of CLIL on L1 and L2 acquisition should be given priority; other authors have pointed out the need to assess language and content knowledge (Hütner and Rieder-Bünemann, 2010; Lasagabaster and Ruiz de Zarobe, 2010). There is also a niche to be filled stating a possible correlation between methodological aspects used in the CLIL classes and the outcomes of the CLIL approach (Admiraal et al., 2006; Lasagabaster, 2008; Lasagabaster and Ruiz de Zarobe, 2010). Research should also take into account the teachers’ language command, their assessment procedures and collaboration and coordination strategies between CLIL teachers and other teachers in the school and with stakeholders. According to Pérez- Cañado (2012) and Bruton (2013), it is fundamental that the studies take into account different variables such as the motivation, the sociocultural status and the gender of the learners, the type of school and its location, the amount of exposure to the language and the pupils’ previous exposure to the target language. Two other variables should also be considered: the linguistic and the methodological competence of the teachers. By factoring in all these variables, the homogeneity of the sample would be guaranteed. Following this same

author, claims should be made on the basis of reliable statistical methods and studies should be based on both quantitative and qualitative data.

Catalonia, where the present study has taken place, is a bilingual community in which two official languages are spoken: Catalan and Spanish. Catalan is the language of tuition in the schools, and children are officially introduced to the study of a foreign language (mostly English) at the age of six, that is, in their first year of primary education, although many schools offer foreign language awareness programs to students earlier on. In recent years, due to the poor results reported by the PISA reports¹ on the students' level in English, and following the recommendations of the Catalan education authorities, a growing number of primary and secondary state and state -assisted schools have progressively adopted a content-based methodology in order to increase the number of contact hours with English. The objective for most schools is to improve the poor standards of foreign language competence achieved by the students.

The present study which has been carried out within the framework of the CLILSLA Research Project² seeks to address the effects of the CLIL approach, using English as the target foreign language, in three state-assisted Catalan schools which decided to provide an hour a week in English in the subjects of Natural Sciences (one school) and Arts & Crafts (two schools) to students in their final stage of primary education (5th and 6th years) for two consecutive years. The study has a two-fold objective: on the one hand, it aims to determine the extent to which CLIL may affect the listening, reading and writing abilities of YLs exposed to Natural Sciences and Arts &

¹ <http://www.oecd.org/pisa/>

² CLILSLA Project is a government-funded research project on the implementation of CLIL in primary school in Catalonia: *Los efectos del aprendizaje integrado de contenidos curriculares y lenguas extranjeras (AICLE) en la adquisición y desarrollo del inglés. Un estudio longitudinal* (FFI2010-19997) It is led Dr Elisabet Pladevall. Universitat Autònoma de Barcelona.

Crafts in English, compared to the results obtained by children in the same schools exposed only to EFL. The results obtained by 5th primary graders exposed only to EFL classes (Control Group) have been compared to those obtained by 5th graders exposed to exactly the same number of hours of English (EFL and CLIL hours combined) to determine their achievement at different times (T0, T1, T2 and T3), and their progress in the target language at different time periods (T0-T1, T1-T2, T2-T3, T0-T3) over two school years. For the results to be reliable, one intervening explanatory variable has been taken into account for statistical data analysis: the proficiency level in English of the students involved (High and Low achievers). Another variable, the number of hours of school exposure to English until the end of the 4th grade, has also been taken into consideration in the analysis. It is to be hoped that, the results obtained will throw some light on the effectiveness of CLIL programmes at primary level. The study may also allow us to see whether the students' listening, reading and writing skills benefit from the CLIL approach and it may clarify some aspects that may need to be considered when schools decide to implement future CLIL programmes in primary education.

1.3 Research questions

In order to achieve the objectives described above, and to determine the achievement and the progress of students exposed to CLIL compared to the achievement and progress of their peers exposed only to EFL, always keeping the number of hours of exposure constant, the following research questions were posed. Questions 1 and 2 refer to the results obtained by CLIL students exposed to Science and Arts & Crafts compared to those obtained by EFL students within the same schools. Question 3 refers exclusively to CLIL learners and compares the results obtained by CLIL Science students to those obtained by CLIL Arts & Crafts students.

Research question 1 is concerned with the achievement and progress in the listening and reading skills of CLIL students compared to EFL students.

RQ1 Keeping the number of hours of exposure to English the same for both groups, CLIL and EFL, do the CLIL students' listening and reading skills benefit from their exposure to the CLIL experience?

RQ1.1 Are there any differences in achievement between CLIL and EFL learners statistically significantly in favour of CLIL learners at different times (T0, T1, T2, T3)?

RQ1.2 Are there any differences in progress between CLIL and EFL learners significantly in favour of CLIL learners after one year (T0-T1) and two years (T2-T3) of CLIL implementation? Are there differences in favour of CLIL students when we consider their progress from T0-T3?

RQ 1.3 How does the initial level of English proficiency affect the students' performance in the CLIL and the EFL groups?

Research Question 2 is concerned with the achievement and the progress in writing of CLIL students compared to EFL students.

RQ 2 Keeping the number of hours of exposure the same for both groups, CLIL and EFL, do the CLIL students' writing skills measured in terms of Complexity, Accuracy and Fluency benefit from their exposure to the CLIL experience?

RQ 2.1 In terms of fluency, are there differences in achievement between CLIL and EFL learners significantly in favour of CLIL learners at different times (T0, T1, T2, T3)?

RQ 2.2 In terms of fluency, are there differences in progress between CLIL and EFL learners significantly in favour of CLIL learners after one year (T0-T1) and two years (T2-T3) of CLIL implementation? Are there differences in favour of CLIL students when we consider their progress from T0-T3?

RQ 2.3 In terms of accuracy, are there differences in achievement between CLIL and EFL learners significantly in favour of CLIL learners at different times (T0, T1, T2, T3)?

RQ 2.4 In terms of accuracy, are there differences in progress between CLIL and EFL learners significantly in favour of CLIL learners after one year (T0-T1) and two years (T2-T3) of CLIL implementation? Are there differences in favour of CLIL students when we consider their progress from T0-T3?

RQ 2.5 In terms of complexity (lexical and syntactic), are there differences in achievement between CLIL and EFL learners significantly in favour of CLIL learners at different times (T0, T1, T2, T3)?

RQ 2.6 In terms of complexity (lexical and syntactic), are there differences in progress between CLIL and EFL learners significantly in favour of CLIL learners after one year (T0-T1), two years (T2-T3) of CLIL implementation? Are there differences in favour of CLIL students when we consider their progress from T0-T3?

RQ 2.7 How does the initial level of English proficiency affect the students' performance in writing in the CLIL and the EFL groups?

Research Question 3 is concerned with the achievement in listening, reading and writing of CLIL students exposed to Natural Sciences in English compared to CLIL students exposed to Arts & Crafts in English.

RQ 3 Keeping the number of hours of exposure the same for both groups, CLIL Science and CLIL Arts & Crafts students, do the CLIL Science students' listening, reading and writing skills benefit more than the Arts & Crafts students from their exposure to the CLIL experience?

RQ 3.1 Are there differences in achievement in listening and reading between CLIL Science students and CLIL Arts & Crafts students at different times (T0, T1, T2, T3)?

RQ 3.2 In terms of writing measured as complexity accuracy and fluency, are there differences in achievement between CLIL Science and CLIL Arts & Crafts students at different times (T0, T1, T2, T3)?

RQ 3.3 How does the initial level of English proficiency affect the students' performance in writing in the CLIL Science and the CLIL Arts & Crafts group?

The results from studies carried out in the Canadian immersion programmes (Genesee, 1987, Swain and Lapkin, 1990, Wechse, 2002, Lyster, 2007) as well as studies in CLIL in the European context (Wode, 1999; Hellekjaer, 2004; Gassner and Maillat, 2006; Dalton Puffer, 2007; Lasagabaster, 2008; Loranc-Paszylk, 2009; Jiménez-Catalán and Ojeda, 2009; Ruiz de Zarobe and Jiménez-Catalán, 2009) among many others revealed advantages for CLIL students. However, other studies seem to provide contradicting results: Jimenez Catalan et al., (2009) in a study carried out with primary learners found some weaknesses on productive vocabulary tasks; Vollmer et al., (2006) and Whittaker et al., (2006) analysed the written language of secondary students and found some deficiencies in academic literacy and writing skills. However, other studies Navés and Victori (2010), Navés, (2011) pointed out significant advantages for CLIL students in fluency and complexity but not in accuracy.

Nevertheless, CLIL seems to provide learners with a richer, naturalistic environment that reinforces language acquisition and learning, and leads learners of all abilities to a greater proficiency (Lyster, 2007; Krashen, 1985; Lightbown and Spada, 2006). Recent research on CLIL in tertiary education has suggested that even low level students benefit from their exposure to the CLIL approach (Aguilar and Muñoz, 2013).

In the light of these findings the following hypotheses are raised:

Hypothesis1. Keeping the number of hours of exposure constant, considering the results of research in content-based learning in different contexts, primary students in their two final years of primary education, exposed to CLIL at school, will obtain better results in receptive skills but not necessarily in written productive skills than primary students only exposed to EFL classes.

H1.1 The students exposed to an EFL+CLIL programme at school will obtain significantly better results in receptive skills, listening and reading, than the students exposed to an EFL programme.

H1.2 The students exposed to an EFL+CLIL programme will obtain better results in written fluency and complexity than the students exposed to an EFL programme.

H1.3 The students exposed to an EFL+CLIL programme will not obtain better results in accuracy than the students exposed to an EFL programme

H1.4 High and Low achievers exposed to an EFL+CLIL programme will obtain better results in listening, reading and some aspects of writing (Fluency and Complexity) than their counterparts in the EFL groups.

The schools involved in this study decided to implement CLIL using two different subjects. Little emphasis has been placed so far in the relevant literature on the influence that the choice of subjects may have on the outcomes of research results. As Pladevall-Ballester (forthcoming) suggests, the choice of subject will determine, among many other aspects, the content knowledge, the teacher, the activities carried out, and very importantly, the type of language that needs to be addressed in CLIL classes. CLIL seems to foster cognitive development by realigning language and cognition (Dalton Puffer, 2008). As the cognitive effort needed to understand a Science class seems to be bigger than that of an Arts & Crafts session, mainly due to the challenging nature of the contents in Science, Hypothesis 2 was raised.

Hypothesis 2 Keeping the number of hours of exposure constant, students in an EFL+CLIL programme exposed to Science in English will obtain better results in listening, reading and writing than students in an EFL+CLIL programme exposed to Arts & Crafts in English.

1.4 Organisation of the Thesis

This section provides an overview of the organisation of the thesis which is divided into seven chapters. The first chapter introduces the topic, provides a justification for the study and poses the research questions. The next two chapters lay the foundation for the study: they describe the theoretical background and the relevant research carried out on the topics of CLIL and young learners which constitute the basis of this dissertation. The rest of the chapters concentrate on the study: the method, the results and its interpretation and the conclusions reached.

Chapter 2 provides an overview of the topic of CLIL. The chapter introduces and defines the concept of CLIL, and provides a brief insight into its background and

origins. It then refers to CLIL in the European context before going on to consider the basis of successful CLIL programmes and presents an overview of recent research in Europe and in Spain. The final part of the chapter deals with the implementation of CLIL programmes in Catalonia, where this study has been carried out.

As the study involved young primary learners, Chapter 3 turns to the characteristics of young learners in general and young language learners in particular. It also reviews some general psycholinguistic principles that may affect language learning. As the main objective of the present dissertation is to ascertain the different results obtained by CLIL and EFL learners, the final part of the chapter focuses on two different approaches to language learning in the school context: EFL and content-based approaches.

Chapter 4 explains the methodology adopted in this study to examine the progress and achievement in English of CLIL and EFL learners exposed to CLIL at primary school level. School contexts, students and teachers involved in the study are described. This is followed by a description of the data collection times and the instruments used. The chapter also refers to the instruments used to elicit qualitative data: the students' and CLIL teachers' background questionnaires, the teachers' opinion questionnaires, and the CLIL class observation protocols used to get an insight into CLIL practices in the relevant schools. A description of the measures used to analyse the writings collected from the children is also provided. Finally, the data collection protocols and the data analysis are explained, an explanation which includes the statistical analysis and procedures for both the quantitative as well as the qualitative data which has been used to back up the findings.

Chapter 5 presents the results of the study and the answers to the research questions posed. Results concerning the progress and achievement in the listening,

reading and writing skills obtained by students exposed to CLIL in their Natural Science and Arts & Crafts subjects are presented and compared to the results obtained by the students exposed to EFL. This is followed by a comparison of the results obtained by CLIL students exposed to Natural Sciences in English with the results obtained by CLIL students exposed to Arts & Crafts in English.

Chapter 6 analyses and interprets the results presented in the previous chapter in terms of the listening, reading and writing skills of the students exposed to CLIL and EFL first, and CLIL (Science vs. Arts & Crafts) in the second place.

Finally, Chapter 7 presents the conclusions of the study, acknowledges the limitations and introduces future research directions in CLIL.

CHAPTER 2 Content and Language Integrated Learning (CLIL)

2.1 CLIL as a European Construct

The term CLIL (Content and Language Integrated Learning) was coined in Europe and it refers to the educational practice in which an additional language is used to teach curriculum subjects. It was launched by UNICOM, the University of Jyväskylä in Finland and the European Platform for Dutch Education in the mid-1990s (Marsh and Wolff, 2007; Fortanet-Gómez and Ruiz-Garrido, 2009) and adopted by the European Network of Administrators, Researchers and Practitioners (EUROCLIC) to refer to a plurilingual educational approach to enhance plurilingualism (Nikula and Marsh, 1997). CLIL refers to all types of methodological approaches in which content subjects in the curriculum are taught in an additional language. It describes a range of learning and teaching contexts where second or foreign languages are learnt and used in alternative ways to the regular foreign language classes. It is important to notice here that this term not only emphasizes the language learning aims but also points to the fact that a content subject is expected to be dealt with, in all its complexity, in the classroom, and is not just being used as a means to learn a foreign language. According to Wolf (1998:26), ‘the construct of CLIL is characterized by an extension of the formula ‘language across the curriculum’³ into ‘languages across the curriculum’, which carries with it the idea of integrating the learning of any language with the learning of a specific subject or content.

For many years, European institutions have tried to promote linguistic diversity in education. Linguistic diversity has always been highly important when planning the

³ LAC (Language Across the Curriculum) followed the example of the WAC (Writing Across the Curriculum) movement which became very popular in the UK in the 80s. It was a proposal for native language education which focused on reading and writing in all subject areas in the curriculum: *Every Teacher an English Teacher*, as it was called in the UK (*The Bullock Report*, 1975).

successful construction of Europe, and languages and linguistic diversity have always been an asset worthy of preservation and promotion. Despite the fact that schools in which certain subjects are offered in a foreign, regional or minority language have been around in Europe for decades, in most cases restricted to linguistically distinctive countries and regions, it was not until the 1990s that the discussion of language learning led to the realization of the need to explore innovative teaching methods. This discussion was influenced by the positive results achieved by students in the Canadian immersion programmes and North American bilingual teaching models, which attested to the success of such programmes at linguistic, content, cognitive and attitudinal levels (Lambert and Tucker, 1972; Canale and Swain; 1980; Cummins and Swain, 1986; Genesee, 1987; Byalistok, 1991; Wesche, 2002). These studies pointed out that students in early immersion programmes through French acquired and developed remarkable proficiency in the language in which they were taught in academic areas such as listening and reading skills. Learners of different abilities achieved different levels of competence, but their range of achievement was similar across the bilingual and monolingual groups. However, the studies also pointed out that the students were not as competent as their peers in monolingual classes as far as their productive skills and their grammatical competence were concerned. Cummins (1991) called attention to the fact that the acquisition of Cognitive Academic Language Proficiency Skills (CALP) took comparatively longer than Basic Interpersonal Communication Skills (BICS), and that form-focused language was needed to develop the students' linguistic accuracy. The outcomes of partial immersion and late immersion were not as successful as those of students in total immersion settings and they were attributed to the threshold level of language required for students to be functional in such contexts (Johnson et al., 1997). Although most studies focused on second language development, a few concentrated on

other aspects. In terms of content, students assimilated knowledge at almost the same level as monolingual students in the same context, as well as showing very positive attitudes towards the L2. De Jabrun (1997) even concluded that immersion students were more efficient mathematical learners than their mainstream counterparts. Senesac (2002) investigated students in a US bilingual programme and determined that students in partial immersion contexts performed better in language, mathematics, science and social sciences than students in monolingual settings.

The Canadian and bilingual models were highly successful experiments that influenced bilingual education in Europe and beyond. The outcome of the models was decisive for the design and the implementation of CLIL in Europe (Pérez-Vidal, 2007). The models were adopted and adapted for use in certain regional and national European contexts: in Catalonia and in the Basque country in Spain, and with Gaelic speakers in Scotland, and with the Welsh and the Irish among many others in Europe. The experience was also valuable in stimulating research in this area, as well as encouraging the development of a wide variety of educational experiences. According to Coyle, Hood and Marsh (2010), there are two types of reasons that underpin an interest in CLIL in a particular context: *reactive* and *proactive* reasons. In the case of Catalonia, for example, CLIL was pushed into Europe by *reactive reasons*, which were a response to the need to strengthen some minority languages in some specific European contexts: a language was adopted in the school systems to act as a language of national unity. Nevertheless, the European idea of promoting multilingualism was also developed for *proactive* reasons, which contributed to enhancing foreign language learning and other forms of educational, social and personal development. Although the Canadian and North American experiences are not directly transferable to Europe, because they are highly context-specific and bear little resemblance to the study of foreign languages

through CLIL programmes in Europe, the research outcomes of such programmes provided valuable lessons that were used by European institutions in order to set up general basic guidelines for foreign language learning in Europe through the medium of content and language integrated learning.

From the 1990s, CLIL became a priority as a major educational approach within the European Union. One of the first pieces of legislation launched in Europe was the Resolution of the European Council in 1995 on improving additional language learning and teaching within the education systems of the European Union. The resolution referred to the need to teach disciplines in an additional language, providing in this way a form of bilingual or multilingual teaching. The same resolution also proposed improving the quality of training for language teachers. In the same year, the White Paper on Education and Training⁴ (Teaching and Learning - Towards the Learning Society, 1995) emphasized the importance of employing innovative ideas to help EU citizens become proficient in three European Languages. The document stressed the idea of the formula 2+1, in other words, all European citizens should be able to use their own language plus two others. Through this document a whole strategy towards multilingualism was framed. Two factors were strongly emphasized: interdisciplinarity and intensity of exposure. They were both considered important elements to be incorporated into the educational policies of member states.

It was within this framework that a number of initiatives were established within the European Union, at different levels, to support the idea of a multilingual Europe. New programmes were introduced, while already existing ones received further backing: the BILD project (Bilingual Integration of Language and Disciplines) and the

⁴ http://ec.europa.eu/languages/documents/doc409_en.pdf

DIESeLL projects (Distance Inset for Enhanced Second Language Learning, the Thematic Network Project (TNP) in the area of bilingual education, the CLIL Compendium, the ALPME project (Advanced Level Programme in Multilingual Education), the TIE-CLIL (a European Cooperation Socrates Project for writing CLIL teacher education materials), the TICCAL as well as two networks: the CLIL Consortium Cascade Network (CCN) and the MOLAN Network Project which fosters the exchange of good teaching practices. Socrates, Comenius and Erasmus programmes were launched to provide the development of CLIL-type provisions and to support mobility and teacher training at all levels of education (Nikula and Marsh, 1997; Dalton-Puffer, 2008; Pérez-Vidal, 2008; Ruiz de Zarobe and Lasagabaster, 2010). In the words of Pérez-Vidal (2008), ‘these programmes have resulted in a European construct around which shared knowledge and expertise are already an asset to European language acquisition and language pedagogy research’ (p.41).

Many other steps have also been taken: for example, 2001 and 2008 were declared European Year of Languages in order to promote language learning and linguistic diversity. In 2001, the Council in Europe issued two documents for the promotion of language learning in Europe: the *Common European Framework of Reference for Languages: Learning, Teaching, Assessment*⁵ and the *European Language Portfolio*⁶. In 2002, the Barcelona European Council called on member states to make a substantial effort to ensure the teaching of at least two foreign languages from a very early age. The European Council reported the results of the

⁵ CEFR establishes a common base for curriculum and materials design which affects the assessment of linguistic competence of language learners

⁶ The European Language Portfolio was conceived as an informative language passport for European citizens and at the same time as a formative instrument to promote autonomy, self-assessment and awareness for language learners

symposium entitled *The Changing European Classroom: the Potential of Plurilingualism in Education*. Among many conclusions, there was the proposal that students should be involved in CLIL-type provisions at different levels of school education, and teachers should receive special training in CLIL.

Countries in Europe have responded to the calls for CLIL in many different ways and perspectives. However, in the words of Dalton–Puffer, Nikula and Smit (2010: 5), ‘the juncture between policy declaration and policy implementation is rather diffuse’. Although a general policy has been formulated at a European level, the legislation on CLIL has been dependant on the educational policies of each of the member states, and this has created a wide range of situations in the various states. Based on the Eurydice Report, Content and Language Integrated learning (CLIL) at School in Europe⁷ (2006), which analyzed the implementation of CLIL in Europe, CLIL-type provision is, in some form or another, part of mainstream school education in the majority of European countries.

Following recommendations from the European institutions, CLIL provision is offered in most countries at primary and lower and upper secondary levels of education. Several countries even provide pre-primary activities in another language, either in a minority language or in a foreign language. Although in some countries CLIL approaches are not widespread in mainstream education, in most European education systems CLIL provision has become an important tool for language learning. Most CLIL models in Europe aim at increasing the presence of a foreign language in the curriculum by incorporating different subjects taught through the foreign language (Van de Craen and Pérez Vidal, 2001; Lasagabaster, 2008; Lasagabaster and Ruiz de Zarobe,

⁷ <http://bookshop.europa.eu/en/content-and-language-integrated-learning-clic-at-school-in-europe-pbNCX106001/>

2010; Coyle et al., 2010). However, CLIL is still far from being a consolidated and fully-articulated unique model in any of the European countries surveyed. A great deal more needs to be done, for instance, in order to consolidate the theoretical underpinnings of CLIL and create a conceptual framework that is both coherent and applicable to different local conditions (Dalton Puffer, 2008).

As for the status of the languages involved in CLIL-type provisions across Europe, the language patterns vary from combinations involving foreign languages, to those involving regional or minority languages. Some regions, such as Catalonia and the Basque Country in Spain, provide scope for trilingual CLIL provision, combining the official languages and with a foreign language. As Pérez-Vidal (2001) stressed, CLIL programmes in such countries benefit from the accumulated experience of highly successful programmes for the normalization of their official languages.

As far as the choice of subjects taught in the CLIL-target languages is concerned, Maljiers, Marsh and Wolff (2007) provided an insight into CLIL in 20 European countries. They concluded that there are no significant variations between primary and secondary levels. Very few countries offer legislation on what subjects should be taught. Any subject can be taught as long as it is part of a national curriculum. Finally, according to the Eurydice Report (2006), the time allocated for teaching CLIL varies from one country to another and from one region to another. The time allocated depends mainly, on the status of the language in the CLIL provision. Hence, schools are largely free to determine the nature and scale of their own CLIL-based subjects or activities. A more recent European document, Key data on Teaching Languages at School in Europe⁸ (2008), backed up the idea of giving priority to the development of linguistic abilities in very young children. The document emphasizes the importance of

⁸ <http://bookshop.europa.eu/en/key-data-on-teaching-languages-at-school-in-europe-pbEC8108375/>

an early start to second language acquisition, the intensive teaching of a second language at primary school, and the transdisciplinarity nature of teaching among many other aspects.

As CLIL is deep-seated within the European educational systems and strongly European-oriented, it is no longer considered ‘the European label for bilingual education’ (Lorenzo, 2007: 28), nor an ‘offshot of other types of bilingual programmes’ (Pérez-Cañado, 2012: 319). It is perceived by most European countries as an increasingly acknowledged trend in foreign language teaching, a form of learning a foreign language without changing the existing curriculums. Its main aim is not to achieve native-like competence as in the case of bilingual programmes, but to acquire a functional level of the language without altering the contents of the national curricula (Dalton-Puffer et al, 2010; Pérez-Cañado, 2012). The language of instruction in many CLIL programmes is one that students will only encounter at school, as it is not a language widely used in their societies. In most countries, the language chosen for CLIL-type programmes is English, due to its role as a lingua franca for international communication.

Spain is one of the European countries that provides numerous examples of CLIL policies and practices. The political structure of Spain comprises 19 autonomous regions that were granted political and administrative power by the 1978 democratic Constitution. This regulates the legislative frameworks that guide the Spanish education system and, for this reason, Spain has as many models as it has regions. Some of the regions are monolingual communities where Spanish is the language of schooling together with one or two foreign languages, while others, such as Catalonia, provide scope for bilingual education. This has led to a variety of CLIL experiences which have been implemented under varied socioeconomic, political and linguistic models and

which are at different stages of development. Nevertheless, they all share the same main objective: to achieve communicative competence in second and/or foreign language across the curriculum.

In order to talk about CLIL in Catalonia, where this study was carried out, the next section will refer briefly to the historical and political contexts for language learning that have occurred in Catalonia as well as the emergence of CLIL programmes.

2.2 The Language Situation in Catalonia. The Emergence of CLIL

For years after the civil war (1936-1939), Spanish was the only official language in Spain, and multilingualism was forbidden. None of the other regional languages (Galician, Basque or Catalan) could be used in public or in public documents. However, languages other than Spanish were used in many homes and clandestine activities (Pérez-Vidal, 2000). During the 1970s, Catalonia received an influx of workers from other parts of Spain attracted by the industrialization of the region and the prospect of work. For all of these workers, Spanish was their first and only language. When democracy arrived in 1978, and the Statutes of Autonomy law was passed (*Boletín Oficial del Estado*, BOE, 1981, 101), a number of autonomous communities were granted political rights and administrative powers, including Catalonia. At the linguistic level, Catalan was granted official status and was declared to be the official language of Catalonia, together with Spanish (Pérez-Vidal, 2000).

At the beginning of the 1980s, several immersion language programmes were set up in Catalonia, propelled by the new socio-political situation. In 1983, the laws of linguistic normalisation were passed, and control over the education system was regained from the state by the regional government (Serra, 1997). A model of maximum normalisation was progressively set up for both primary and secondary education.

Pupils were taught in Catalan except for the subject of Spanish. Catalan slowly became the vehicle of instruction in schools for both Catalan speakers and Spanish speakers.

Up until 1986, when Spain joined the EEC (European Economic Community), French had been the most popular foreign language taught in schools. The teaching approach for French had been a traditional one, focusing on grammatical structures. In the early 80s, due to new social demands, there was a rapid change from French to English in most schools. A great number of teachers had to switch from teaching French to teaching English in a brief period of time. Their poor methodological and linguistic training, and their poor communicative skills, together with the fact that the classrooms were overcrowded, led most of them to put the emphasis on grammar and written skills rather than taking a communicative approach. In 1990, the government reformed the educational system through the *Ley Orgánica General del Sistema Educativo (LOGSE)*, which introduced new regulations to the teaching of foreign languages (Navés et al., 1999). It brought forward the introduction of foreign languages from the age of eleven to eight in primary education, and offered the possibility of learning a second foreign language at secondary level. It also increased the number of hours of foreign language instruction in order to promote communicative objectives that had been lacking in foreign language classes up to that moment. The new law strongly advised schools to take a languages-across-the curriculum approach. There were also strong recommendations for an inter-disciplinary approach to the teaching of foreign languages. In 2005, a new education bill LOE (*Ley Orgánica de Educación*), which brought forward the starting age for foreign language learning even further to the 1st grade of primary education, was passed. Several research projects on the advantages of starting to learn English at an early age were then carried out in Catalonia and the Basque country (Cenoz, 2002; Muñoz, 2006). The BAF project (*Barcelona Age Factor*)

(Muñoz, 2006), which had started back in 1995, studied the effect of age on the process of learning English. It also compared the results of English language acquisition of children who had started learning English at eleven with those who, under the law passed in 1990, had started learning English at the age of eight. Research results suggested that children's language proficiency did not benefit from an early start, if this early start did not imply a substantial increase in the number of hours of meaningful exposure to the language. The results also suggested that a content-based approach could guarantee better opportunities to use the language meaningfully.

In the last ten years, a movement in favour of implementing CLIL in different forms has slowly been gaining ground within the Catalan school system, following the recommendations of the *LOE*, which strongly recommended that the teaching of foreign languages be organised within a task-based/project work methodology. Interest in the learning of foreign languages, mainly English, has led many primary schools to introduce the teaching of English through different forms of CLIL, such as whole subjects, parts of subjects or even through occasional project work. Over the years, CLIL has slowly found its place in the school curriculum in Catalonia (Navés and Victori, 2010). During the last decade, the *Departament d'Ensenyament (Catalan Education Authority)* has promoted and funded a number of CLIL projects (*PELE projects*, in Catalan, *Programes Experimentals de Llengües Estrangeres* and, more recently, *PILE projects*, *Projectes d'Integració de Llengües Estrangeres*) in the state school system. Most of these projects last three years, and the teachers involved receive training in order to have a theoretical basis of CLIL as well as allowing them to create their own materials to put their projects into practice in primary or secondary education. However, apart from some pilot projects, supported by the Catalan Education Authority, CLIL is still a voluntary, teacher-led approach in most schools.

CIREL⁹ (*Centre de Suport a la Innovació i Recerca Educativa en Llengües*) has been supporting research and innovation projects on foreign language learning since 1999 through its '*In-service Teacher Training Unit*'. It organises teacher training courses for foreign language specialists interested in training as CLIL teachers in order to increase their theoretical knowledge of CLIL and to improve their linguistic confidence and competence. It also offers training and monitoring for state and state-assisted schools and those teachers who have voluntarily decided to implement pilot CLIL projects. It also provides advice and training in order to improve certain methodological aspects of CLIL and to help teachers to design and develop teaching materials, lesson plans and guidelines, which are later, uploaded onto the CIREL website. This has not only created a valuable collection of materials which can be consulted by teachers interested in the CLIL approach, but has also helped teachers to feel that they belong to a community that continuously contributes to CLIL with new materials and experiences.

A new initiative called *Train the CLIL Trainer Program* was launched in 2008. It was an initiative formed by a team of expert CLIL teachers from both primary and secondary education with the objective of designing training portfolios to be used in the CLIL teacher training courses that they have provided since 2009. Despite such initiatives, CLIL teacher training programmes are as of yet still quite scarce. At the present moment, primary teachers' initial university training degrees still do not include specific CLIL training qualifications. Only recently have some universities started to offer some modules in master's degrees to prepare teachers for this new way of learning. Furthermore, because CLIL is a relatively recent approach to teaching, in most Catalan schools in which a CLIL programme is implemented, especially in primary

⁹ http://phobos.xtec.cat/cirel/cirel/index.php?option=com_content&view=article&id=180&Itemid=1

education, it is usually the English specialist, quite often a non-native speaker of English with a limited command of the language, who teaches CLIL sessions and who takes part in the in-service CLIL training programmes offered by the Departament d'Ensenyament. At ground level, in secondary education, the provision of CLIL teachers varies, but mainly it is the content teacher who teaches CLIL sessions with the support of the foreign language specialist. As Navés and Muñoz (1999) had already mentioned, the fact that the teacher training system in Spain does not include dual qualifications makes it very difficult to find teachers to implement CLIL. The authors also mention the fact that teachers do not have enough basic theoretical information and that they do not always feel familiar enough with the rationale for CLIL. The successful implementation of a CLIL programme at school requires certain very important conditions to be fulfilled.

This section in this dissertation has analyzed the origins of CLIL and has provided some background information on the implementation of CLIL programmes, first in Europe and then in Catalonia, where the present study was carried out. The following section aims to provide an overview of the characteristics of successful CLIL programmes.

2.3 Successful CLIL Programmes. Characteristics and Outcomes

2.3.1 Insights into CLIL success factors

The reasons for setting up a CLIL programme vary across the different educational European contexts due to the variety of circumstances that surround language teaching and learning around the continent (Lasagabaster, 2008). However, *The CLIL Compendium Research Project*¹⁰ identified 5 dimensions of CLIL based on issues relating to culture, environment, language, content and learning, that define the

¹⁰ www.clilcompendium.com

basis of all CLIL programmes. These dimensions are heavily interrelated in CLIL practice, and provide a general base within which to frame some of the reasons why schools decide to implement CLIL. These dimensions may not all apply at the same time when a school decides to implement a CLIL programme and, as Marsh, Maljiers and Hartiala (2001:17) pointed out, 'the reason for the implementation may change over time'.

As suggested by Marsh et al., (2001), if one considers the socio- cultural dimension, CLIL can be seen as a broad educational practice and even a political way of achieving bilingualism, multilingualism and even multiculturalism in Europe. A CLIL programme prepares students for internationalization, and at the same time it brings an understanding of other cultures. It helps learners go beyond their local context and introduces them to a wider cultural ambit, hence developing their intercultural communication skills. This dimension is embedded in the already popular motto: 'Europe will be multilingual or it will not be.'

As for the language dimension, according to the *CLIL Compendium*, a CLIL programme focuses on improving overall target language competence and helping learners improve their attitudes towards both their mother tongue and the target language. This helps them to prepare for future studies and/or their working lives.

Focusing on the content dimension, CLIL programmes afford opportunities to study content subjects from different perspectives, and provide the possibility of accessing specific terminology in the target language (Pérez-Vidal, 2001). Finally, the learning dimension encompasses individual forms of learning in which languages carry out a special role alongside the learning of a specific subject or content: this dimension diversifies methods and forms of classroom practice and increases learners' motivation.

Following a trend of globalization, CLIL has embedded itself into mainstream education and it has gradually become an established teaching approach in most European countries. For numerous authors (Muñoz and Nussbaum, 1997; Coyle et al, 2010), this type of content-based programmes is seen in most cases as a form of supplementing the already existing hours of instrumental foreign language instruction on the curriculum, and also as a way of compensating for the limited exposure to the target language that most learners get from traditional instruction. The integration of language and content thus provides substantial exposure to the target language in a meaningful context.

Despite the heterogeneous implementation of CLIL in Europe, certain common characteristics can be identified. After considering the extensive research carried out on content-based, bilingual, immersion and CLIL programmes, Navés (2009) grouped and summarized the characteristics of successful CLIL programmes (SCLILP) under a number of headings. In her opinion, SCLILP support the learners' home language and require long-term, stable teaching staff, parental involvement and the joint effort of all parties. She also refers to the relevance of the teachers' profile and training, the materials and the methodology as determining factors for their success.

Second Language Acquisition research (Bialystock, 1991; Collier, 1989, 1992; Genessee, 1987; Thomas and Collier, 1997; Crawford and Krashen, 2007) has shown that the level of proficiency in the first language has a direct influence on the development of proficiency in the second/foreign language. Cummins (2003: 63) stated that 'the level of development of the children's mother tongue is a strong predictor of their second language development', while for Krashen (1997), literacy developed in the first language transfers to the second language. Therefore, teachers involved in CLIL should take into consideration the students' L1 literacy as well as their second

language proficiency. CLIL programmes should allow learners to use their L1 in classroom interaction, especially in the early stages, when the linguistic competence of the learners is too limited for them to be able to communicate in the target language. Novotná, Hadj-Mousová and Hofmannová, (2001) explain that code switching, that is the alternative use of two languages, is a natural communication strategy in CLIL classes. Learners, whose attention is focused on the non-linguistic content, need to have access to spontaneous speech, preferably in an interactive context. In two studies carried out by Adler (1998) and Setati (1998) with secondary students learning mathematics in a CLIL context, code-switching was proved to be a valuable educational resource and a means to foster secondary learners' mathematical understanding.

Darn (2007) has already pointed out that CLIL identifies a 'transition' stage at which learners become fully functional in both languages. It is at this stage that translation of key terms is considered acceptable in order to promote full understanding of the subject. Nevertheless, the input, that is, listening and reading, according to this author, should always be in the target language, and the output (speaking) may be accepted in the L1 of the learners. Lessons in contexts of partial instruction may involve the use of the CLIL language and the vehicular language in some specific types of activities (Coyle, 2010). This author refers to this way of using both languages as 'translanguaging', that is, a systematic shift from one language to another for specific reasons and in circumstances in which the use of the students' first language may be considered necessary, for example, to make sure that the students understand key concepts before performing an activity or continuing with a lesson.

Teacher quality is considered as fundamental in order to ensure school effectiveness and student achievement (Navés, 2009). CLIL programmes require linguistically and methodologically-competent, trained teachers who are proficient in

the target language, have some knowledge of the principles of second language acquisition, and are able to adapt their pedagogical skills to the age of the learners (Van de Craen and Pérez-Vidal, 2001). The present pedagogical approaches to teaching English to Young Learners suggest clear implications for teacher proficiency. These approaches imply an activity-based, interactive, learner-orientated plan of action, focusing on both meaning and communication (Moon, 2009). The teachers' command of English affects both the model of language that can be provided and the type of methodology that can be adopted. A teacher who lacks confidence and fluency in the language is unlikely to be able to set up the occasions for genuine interaction (Rixon, 2000: 3-4). Therefore, a CLIL teacher should have enough communicative competence in the target language to be able to use it to teach a content subject. Likewise, as mentioned before, a good knowledge of the learners' first language is advantageous, as teachers should be able to appreciate and provide solutions to the learners' language difficulties. Navés (1999: 93) points out that 'teachers should always show an understanding of learner's first language by responding appropriately and rephrasing learners' remarks made in their L1'. De Graaff (2007) described five indicators for effective language teaching performance. Three of them were directly related to the teachers' competence in the language: exposure to English, form-focus processing and output production. Hence, teachers should be able to analyze structural and lexical features of the target language in order to anticipate problems and to be able to establish the necessary language scaffolding to make learning profitable and successful. CLIL is unlikely to be implemented successfully if it simply consists of a translation of lesson contents into another language. The quality of the outcome will not be enhanced by simply changing a linguistic code (Coyle, 2006). In a CLIL setting, it is important to maximize the amount of comprehensible input and purposeful use of the target language

in the classroom. It is also vital to guarantee not only the quantity but also the quality of the input. Nevertheless, the target language instruction in CLIL is contextualized and integrated. Language in CLIL is used as a means to learn something else, not as an end in itself. In the words of Navés (2002: 93), 'target language instruction is not structured or of a pull-out nature but rather contextualized, integrated and sheltered'. Students in CLIL classes are 'sheltered' in the sense that they are considered LEP (Limited English Proficiency) students who are being taught in ways to make subject-area content comprehensible. Sheltered programmes in America have proven successful because they concentrate on the simultaneous development of content-area and ESL proficiency (Echevarria, Vogt and Short, 2007). Pavessi et al., (2001) suggested that, in CLIL, teachers are also expected to be experts in the subject area and to have a deep knowledge of the cognitive, psychological and cultural elements involved in language learning. As CLIL is not related to one specific methodology, teachers on a CLIL programme need to receive special training, that allows them to use appropriate strategies and to work on the consistent integration of cognitively-demanding content and target language. As Wolff (2007: 22) pointed out: 'CLIL teachers learn to understand the importance of content in the language learning process'. Frígols, Marsh and Naysmith, (2007) produced a list of 'idealised competences' required of a CLIL teacher. Among many considerations, they also mention the need for methodological training in order to exploit methodologies and resources that facilitate the understanding of meaning and the identification of linguistic difficulties, as well as exploiting strategies for modelling good language use in context. For Wolff (2007), it is important for teachers implementing CLIL projects to establish a partnership with non-language subject teachers in order to create careful co-ordination to enrich the outcomes of the CLIL programme in the long term. Along these lines, Coyle (2010) also mentions the

challenge that CLIL represents for teachers. Successful CLIL practice requires alternative ways of planning and teaching in order to address fundamental aspects of the effective integration of content and language.

However, even though it is widely recognized that teaching in a CLIL context requires a good command of the language, as well as solid methodological training, there is still no consensus between experts as to what the minimum proficiency for the successful application of the CLIL approach might be. Marsh (2000) claims that it is unnecessary for teachers to have native or near-native competence in the target language, although they do naturally need a high level of fluency. Few educational authorities in Europe have laid down formal requirements for those teachers who want to embark on CLIL. Different standards for CLIL teachers' foreign language proficiency apply in different countries: B2 level to C1 in Poland and Hungary, for example, whilst in Finland, for example, the Ministry for Education proposes a C2 level of proficiency (Eurydice Education Unit, 2006: 43). Nevertheless, as CLIL moves increasingly into mainstream education, the call for a more defined level of target language proficiency of CLIL teachers is increasing, as is the need for specialized training and methodological qualifications. As Takala (2002: 26) states: 'one crucial aspect of CLIL should be spelled out: how good should CLIL teachers' proficiency in the language of instruction be and how could that level be reliably checked?'. It is widely accepted that proficiency in the target language, the ability to use appropriate methodologies, and knowledge of children's cognitive and linguistic development are all fundamental aspects of teaching young learners (Blondin, Candelier, Edelenbos, Johnstone, Kubanek and Traute, 1998).

Since the introduction of CLIL in Catalonia, where the study was carried out, the Catalan education authorities have been running CLIL methodology courses for in-

service teachers (see section 2.2 in this chapter) and they have strongly recommended, at least in Primary education, that the teachers have a minimum level of English equivalent to B2 in the CEFR¹¹.

In the case of CLIL teachers, as has already been pointed out, a good knowledge of the content subject and its specific terminology should also be considered key points. Content education involves, among many other aspects, a good knowledge of terminology, phrases, expressions and concepts that are unique to the subject matter. Academic language has been defined in different ways by several educational researchers (Chamot and O'Malley, 1994; Flynt and Brozzo, 2010). In spite of the differences in their definitions, researchers agree that academic language is more than just content vocabulary words related to content-specific subject knowledge. It represents the whole range of language used in academic settings and differs considerably from the language used in conversational situations. In various analyses of the language used by teachers in assessment, Bailey, Butler and Sato (2007) described two types of academic language: content-specific language and general academic language that is useful across diverse academic settings. According to these authors, learning academic language is much more challenging, especially for those students who are acquiring English as a new language.

Motivation and positive attitudes are also key elements in planning the development and implementation of a successful CLIL programme. Coordination and cooperation between instrumental English teachers and CLIL teachers are fundamental for the implementation of CLIL models in schools. De Corte (2000) said that the teacher's task is to enable the students to develop not only their individual process of knowledge-building and meaning construction but also positive attitudes and

¹¹ CEFR : Common European Framework of Reference

motivation. He goes on to say that a qualified CLIL teacher should be able to maintain the learners' interest and motivation through relevant learning opportunities. Donato, Tucker and Igarashi (2000) found that the attitudes of young learners towards modern languages in primary schools in different countries and from different backgrounds were largely positive. Cook (2001) pointed out that motivation works in two directions: high motivation encourages successful learning, while successful learning encourages high motivation.

It is a common belief that languages are difficult subjects. Therefore, in order to help the learners succeed, it is of the utmost importance for the teacher to examine and analyze possible barriers that might have a negative impact on learning. The CLIL teacher should be able to suggest ways in which these could be minimized, and use a variety of effective teaching strategies that would help overcome individual learning difficulties. In the words of Coyle (2005: 12):

'Motivated teachers 'breed' motivated learners by enhancing learners' values and attitudes related to the foreign language, increasing learner expectations; making the content more relevant for learners (in terms of the subject matter and the cognitive level at which learners operate – which is not dependent on linguistic level)'

Even though the teachers, their training and their motivation can be considered essential elements in the final outcomes of a CLIL programme, the success of such a programme also feeds on the joint effort of all parties involved. Mehisto (2007) in his summary of the implementation of CLIL in Estonia refers to the implementation of CLIL programmes as a 'complex task' that needs to be supported by the school management, the school council, the teaching and non-teaching staff and the parents. The authorities need to provide expert teams to advise and provide general guidelines, as well as offering support of all types, including financial support on the implementation of such programmes. Contextual factors also play an important role.

The coordinators in the schools need to provide the necessary teaching conditions under which school teachers can work, coordinate with other content teachers, and plan and design their materials. Any school where CLIL is implemented needs to plan language teaching carefully, taking into account the competence in the language the students have already acquired. Snow (1989) talked about the importance of the coordination between language and content teachers in order to determine the language objectives: content-obligatory language (language essential to an understanding of content material) and content-compatible language (language that can be taught naturally within the context of a particular subject matter). Wolff (2007: 23) pointed out that in a CLIL setting, ‘teachers who are also content subject teachers will easily recognize the importance of language for the learning processes’. At the same time, coordinators should ensure that there is the necessary continuity across the different academic years to make a CLIL programme successful in the long run. One of the key aspects in the success of CLIL programmes is the continuity of the programme and the stability of the teaching staff. According to Navés and Muñoz (1999), implementing a CLIL programme is not an easy task. Along these lines, Coyle (2010) also claims that the support of the educational structures, which include school management and the administrative structures and even examination boards, is fundamental to the success of CLIL in primary, especially if CLIL is seen as a model to prepare the learners for in-depth long-term achievements.

For Navés (2009), parental support and involvement are also fundamental aspects for success in CLIL. It is fundamentally important that parents understand the nature and benefits of the CLIL programme for their children. As Mehisto (2007) suggests, parents must be given an overview of the programme’s goals and plans, as well as CLIL research, in order to address their fears and concerns. Parents should

maintain high expectations for their children's achievements in the long term rather than in the short run. Likewise, in order to implement a CLIL programme, part of the school policy should be to inform the parents of the fundamental aspects involved in learning a language and to make clear what they can expect from their children while they are exposed to the CLIL approach. Language learning and acquisition is a slow process, and sometimes parents are not aware of this. Quezada, Wiley and Ramírez (2000) claim that children in Bilingual Educational (BE) programmes can take up to seven years to develop their academic language proficiency to an adequate level.

In SCLILP the curriculum has been designed taking into account the L1 speakers' needs and goals, as well as the educational system. A CLIL approach should never allow lower standards in terms of curricular contents than those established for the L1. Coyle (2005) pointed out that CLIL is not about teaching what students already know in a different code. In fact, results derived from the implementation in some schools have shown that students at primary level achieve similar results in CLIL subjects to those obtained by their L1 counterparts. Van de Craen, Ceuleers and Mondt (2007), for example, concluded in a study about teaching maths in a CLIL context in Belgium that there are no indications that subject matter knowledge is worse in CLIL classrooms than in non-CLIL contexts.

SCLILP should promote linguistic awareness and cultural diversity. By using the languages of cultures new to the learner in order to teach content, SCLILP should promote and enhance awareness of languages and cultural diversity. The integration of culture is one of the key aspects mentioned in *The CLIL Compendium*. As suggested by the 4Cs Framework (Coyle, 2001), effective CLIL takes place through progression in knowledge, skills and understanding of the content. The framework also considers of fundamental importance the cognitive processing, interaction in the communicative

context, the development of appropriate language knowledge and skills as well as intercultural awareness whose rightful place is at the core of CLIL (Coyle, 2001).

Young CLIL learners should be made aware of the uses they can make of the foreign language they are learning. Cameron (2001) pointed out that one of the problems with young learners is that quite often they do not realize the use they can make of the language outside their instrumental English sessions. SCLILP, in contrast, bring real-life situations into the classroom and provide real reason for using the language in real contexts. Therefore, they improve not only students' competence in the target language, but also they may contribute to create positive attitudes to other languages and cultures.

In summary, although the integration of content and language is not new, the process underlying the implementation of such a programme is complex and requires certain considerations for it to be successful: administrative and parental support, effective in-service staff development and, very importantly, qualified teachers with a deep knowledge of both the subject content area and the elements involved in foreign language learning.

This section has discussed the general characteristics of successful CLIL programmes, and the structural demands and inherent limitations of the implementation of such programmes. The next section will throw some light on the claims of CLIL by reviewing the latest research in Europe.

2.3.2 CLIL Research in Europe

As has already been pointed out in section 2.1 of this chapter, education authorities throughout Europe are giving increasing importance to the ways in which to improve foreign language competence at school level. The integration of content and language is implemented in most countries in the belief that this kind of approach is the

best way to increase students' foreign language proficiency. Several claims have been made on different grounds about the purported benefits of CLIL. According to Coyle (2007), CLIL leads to greater linguistic proficiency as it triggers high levels of communication between teachers and learners and among the learners themselves. At the affective level, it generates positive attitudes, as it boosts motivation for language learning and is considered beneficial for learners of all abilities. It is believed to foster implicit and incidental learning by focusing on meaning rather than on form, and it generates greater intercultural awareness among students. From the late 1990s until the present day, there has been a noticeable growth of research all over Europe in order to provide insights to help CLIL develop as an educational practice. Individual papers in international journals, edited volumes on CLIL, in-depth explorations on specific aspects have all contributed to providing valuable insights to help CLIL develop as an educational practice (Dalton-Puffer et al., 2010). A variety of these studies in CLIL contexts and at different educational levels have demonstrated that CLIL, in certain contexts and under certain circumstances, can and does enhance many aspects of the learners' linguistic competence. The fact that interest in CLIL has mainly come from the field of Second Language Acquisition, and that the majority of researchers are language rather than content experts, has contributed to a growing body of research articulated around the effects of CLIL on the acquisition of foreign languages. Nevertheless, some research has also been conducted on other aspects such as content, motivation, attitudes and perceptions, as well as the effects of CLIL on L1 acquisition (Lasagabaster, 2008; Pérez-Cañado, 2012).

One of the European areas in which CLIL has been extensively implemented and researched is Scandinavia. In Finland, David Marsh, has set himself up as a key figure from a theoretical perspective, in developing the foundations of CLIL extolling

the potential benefits of the CLIL approach. He has contributed to the establishment of the CLIL Consortium, as well as to the development of materials and the organization of several European conferences (Fortanet-Gómez and Ruiz-Garrido, 2009). In terms of research, the studies carried out in this area have mostly dealt with the effects of CLIL on second language development, the students' attitudes and various aspects of content learning at different levels of education. Järvinen (1999, 2005) analysed the acquisition of subordination and relativization in students from 1st to 6th grades, and concluded that CLIL students wrote more complex and accurate sentences than their monolingual counterparts. Merisuo Storm (2006), in a longitudinal study with 1st and 2nd primary graders, studied the L1 literacy skills of CLIL and non-CLIL learners. She took gender and school readiness as variables. She found more positive attitudes towards language learning among CLIL learners than among their mainstream counterparts. However, she did not find any statistical differences between the groups when the variable gender was factored in: the positive attitudes levelled out and no differences were found between boys and girls. Jäppinen (2006) focused on the effects of CLIL environments on cognitive aspects and subject matter acquisition and determined that CLIL seems to have a positive effect on content learning. From a qualitative perspective, Romu and Sjöberg-Heino (1999) and Södergard (2006) concluded that primary learners exposed to CLIL showed satisfaction, positive attitudes and increased confidence towards the target language.

At pre-university level in Sweden, Airey (2004) refers to two studies: Washburn (1990) and Knight (1990), which investigated the language competence of the students and found no significant differences between CLIL and non-CLIL learners even though the students were matched for different variables such as motivation, intelligence and sociocultural status. Still in Scandinavia, Norway has also provided several studies

focusing on CLIL at different levels of education: Hellekjaer (2004) concentrated on whether upper secondary students in CLIL and non-CLIL classes developed their reading skills sufficiently to succeed in University CLIL courses. CLIL students scored significantly higher than their non-CLIL counterparts. However, the author acknowledged the need to replicate the study in order to concentrate on listening, speaking and writing skills as well. More recently, Airey and Linder (2006) in Sweden, and Hellekjaer (2010) in Norway, tried to determine the nature of problems in lecture comprehension in English–medium instruction of Swedish university students and Norwegian pre-university ones. The main issues focused on the students' difficulties in note-taking and their lack of confidence in asking and answering questions as well as unclear pronunciation on the part of the teachers, together with unfamiliar vocabulary and clear lines of thought in their lectures. According to the authors, the results of both studies provided revealing insights into aspects that affect language skills and lecturing behaviour on the part of the teachers.

As for the central European countries, the Netherlands, Germany, Switzerland and Austria have all contributed with solid research to the body of literature studies on CLIL. Admiraal et al. (2006) carried out a longitudinal study with Secondary Education students who had received four years of CLIL education through English in five Dutch schools. The authors measured receptive vocabulary knowledge, reading comprehension and oral proficiency, and they considered gender, entry ability level, home language, language contact outside school, and motivation as covariates. No differences were found for receptive word knowledge. However, differences emerged for the oral and reading components in favour of CLIL students. De Graaff, Koopman and Westhoff (2007) carried out a qualitative investigation which led to the design of an observation tool aimed at determining effective CLIL pedagogy. The observation tool comprised

five basic assumptions related to effective language teaching performance: the teacher facilitates input at a challenging level, meaning-focused and form-focused processing, opportunities for output production and strategy use. After examining the results of several studies in German-speaking countries, Dalton-Puffer (2008) concluded that the outcomes of CLIL are satisfactory because CLIL students attain a level of competence in the foreign language which is well above that of students enrolled in regular classes. According to her, receptive skills, vocabulary, morphology, fluency and affective outcomes are the aspects that benefit most from the CLIL approach. Wode (1999) reported significant gains for vocabulary in CLIL groups, and Vázquez (2007) observed general gains in the students' general communicative competence. Zytadieß (2007) tested 180 16-year-old students on grammatical, lexical, and communicative competences, as well as subject-matter literacy. The results of the study attested to a significantly higher overall language competence in favour of CLIL students. There was a substantial difference in lexical and grammatical aspects, accuracy and syntactic maturity.

The studies carried out in Switzerland have mainly focused on the effects of CLIL on oral competence, and they have been based on the analysis of classroom excerpts, classroom observation and the analysis of narratives. The results of a study carried out in the canton of Zurich with primary students, by Stotz and Meuter (2003), revealed that CLIL students outperformed their non-CLIL counterparts in terms of receptive skills. However, the outcome for language production was more inconclusive, which was attributed to a lack of opportunities for classroom interaction due to pedagogical aspects that resembled those of frontal classroom pedagogy. Gassner and Maillat (2006), however, reported an improvement in the productive skills of 11th graders in a French CLIL programme in Geneva. Still in Switzerland, Serra (2007) conducted a longitudinal study on the acquisition of subject content knowledge with

pupils from 1st to 6th grades. The results obtained showed that CLIL and non-CLIL students in three public Swiss schools performed equally well in the Italian and Romansch languages. However, in terms of mathematics, the CLIL stream outperformed their non-CLIL counterparts.

According to Pérez-Cañado (2012), most of the research carried out in Austria has concentrated on narrative competence and lexical proficiency and has been carried at secondary level and above. Ackerl (2007) found that pre-university CLIL students produced more complex sentences in their writings, as well as using a greater variety of tenses and more diversified vocabulary. Similarly, Hüttner and Rieder-Bünemann (2007), after studying the effects of CLIL on 7th graders in the Austrian system through the use of a picture story, revealed that CLIL had a positive effect on linguistic and coherent micro-level features, as well as on thematic coherence in their narrative task.

Along the same lines, Seregély's study (2008) agreed with Ackerl's study that there is an improvement in lexical competence: 11th graders in control and experimental groups showed a wider and more complex English vocabulary than traditional EFL students. Extracurricular exposure and time spent in English speaking countries have a significant impact on lexical competence. Jexenflicker and Dalton-Puffer (2010) examined the effects of CLIL in upper-secondary engineering schools in Austria. Positive effects of CLIL were reported on accuracy, vocabulary range, spelling and task completion, but were less clearly stated when organization and structure were analyzed.

Following the global trend of CLIL popularity, a significant number of schools in Poland have started to teach content subjects through the medium of a foreign language. The implementation of CLIL practice in education in Poland has taken place under the name of bilingual education. For a school to be called bilingual, it needs to offer at least two content subjects taught through a foreign language. The most popular

content subjects, depending on the target language, are mathematics, physics, geography, history, biology and chemistry, and less popular are citizenship, music, physical education and ICT. Czura, Papaja, and Urbaniak (2009) report on the outcome of a qualitative project coordinated by the National Centre for Teacher Training and the British Council, known as The Profile Report¹², whose aim was to provide an overview of practice, and to support the development of CLIL procedures throughout the schools. The results revealed that teachers considered CLIL to be a challenging practice and a source of professional satisfaction. Students' considered CLIL as a way of enhancing learning conditions. However, they complained about the use of traditional methodologies, the unsystematic use of code-switching in the classes, and the lower standards of content subjects compared to mainstream education. The report findings allowed the formulation of recommendations for improvement in CLIL practice and policy implementation. Luczywek (2009) provided a qualitative account aimed at identifying and providing the best CLIL practices to integrate the teaching of English as a foreign language through history and literature in Polish secondary schools.

At pre-university level, Loranc-Paszyk (2009) explored the potential of integrated reading and writing activities from the perspective of the students' linguistic achievements within the CLIL classroom of undergraduate students of International Relations, who were offered CLIL classes as an alternative to a traditional EFL course. These CLIL students were learning *History of European Integration* through English as an instructional medium. The analysis of the outcome of the research showed that the progress made by the control group was clearly lower, in spite of having the same time of exposure and a comparable amount of written production.

¹² <http://www.icpj.eu/?id=22>

Novotná and Hofmannová (2007) in the Czech Republic and Kovács (2005) in Hungary also reported on the situation of CLIL in their countries. In Hungary, Varkuti (2010) explored the English language achievement of CLIL secondary school students and those of non-CLIL intensive foreign language learners acting as a control group. The students took the same tests designed to measure conversational and academic language use (BICS and CALP). Data analysis proved that the social and academic language competence of the CLIL students was of a higher level than that of the control group. She concluded that using English as a medium for learning various subjects is a more efficient way of providing functional language proficiency than traditional foreign language learning.

CLIL has not spread all over Italy in a uniform way. There are some regions, especially in Northern Italy, which have been developing CLIL programmes for a decade, and other regions which have only recently become acquainted with this approach (Infante, Benvenuto and Lastruci, 2008). As in the case of Poland, research carried out in Italy has also followed a qualitative trend, mainly focusing on teachers' perceptions (Coonan, 2007) and CLIL practice development. The results obtained by Coonan after analyzing the answers of 33 secondary teachers involved in CLIL suggested that CLIL positively affects the way students learn content, their motivation, and their degree of attention in lessons. The participants in the study carried out by Infante, Benvenuto and Lastruci. (2009) also suggested positive outcomes related to CLIL practice. In spite of the increased workload, and the lack of materials, they consider CLIL to be a way of improving teaching practice which allows the development of thinking skills and the use of activities that foster fluency rather than accuracy. Table 2.1, table 2.2 and table 2.3 below provide a summary of the main research studies and findings within the European context

Table 2.1 Summary of findings in Scandinavia

PLACE	AUTHOR& YEAR	LEVEL/ GRADES	OBJECTIVE	RELEVANT FINDINGS	COMMENTS
SCANDINAVIA	Järvinen , 2005	Primary 1 st to 6 th grades	L2 development and syntax	Bilingual group: + complex and accurate sentences	Homogeneity of groups not guaranteed
	Merisuo-Storm , 2007	Primary 1 st and 2 nd grades	L1 literacy skills of CLIL and non-CLIL learners	+ positive attitudes in favour of CLIL learners No differences between boys and girls	Longitudinal study Variables included: School readiness and gender
	Jäppinen ,2006	Secondary	Effects of CLIL on cognitive aspects and subject matter acquisition	CLIL = positive effect on content learning	
	Romu and Sjöberg-Heino ,1999; Södergard ,2006	Primary learners	Learners attitudes in CLIL classes	CLIL learners= +satisfaction, +positive attitudes, increased confidence	Qualitative study
	Washburn ,1990 Knight ,1990	Pre-university learners	Language competence of CLIL/ non-CLIL students	No significant differences	Variables included: Motivation, intelligence and sociocultural status
	Hellekjaer ,2004	Upper – secondary students	The development of Reading skills in CLIL	CLIL students scored significantly higher	
	Airey and Linder, 2006 Hellekjaer, 2010	University students	Lecture comprehension in CLIL sessions	Difficulties in note-taking, Answering questions, coping with teachers’ mispronunciations.	Revealing information on the language competence and methodology of teachers

Table 2.2 Summary of findings in Austria, The Netherlands, Switzerland and Germany

PLACE	AUTHOR& YEAR	LEVEL/ GRADES	OBJECTIVE	RELEVANT FINDINGS	COMMENTS
Austria, The Netherlands, Switzerland and Germany	Admiraal et al.,2006	Secondary	Measure receptive vocabulary knowledge, reading comprehension and oral proficiency	CLIL: receptive vocabulary knowledge, no differences.+ reading comprehension and oral proficiency	Covariates: gender, entry ability level, home language, language contact outside school, and motivation
	De Graaff et al, 2007	Secondary	Designing an observation tool for effective classroom pedagogy		Based on CLIL class observations
	Dalton-Puffer, 2007	Secondary	An overview of research studies in German speaking countries	CLIL= + receptive skills, vocabulary, morphology, fluency, affective outcomes	
	Wode ,1999	Secondary	Measure vocabulary	CLIL = + gains in vocabulary	
	Vazquez, 2007	Secondary	Measure general communicative Competence	CLIL = + communicative competence	
	Zytadiß ,2007	Secondary	Grammatical, lexical and communicative competence + subject-matter literacy	CLIL= + overall language competence	
	Stotz and Meuter, 2003	Primary	Testing productive and receptive skills	CLIL= + receptive skills	
	Gassner and Maillat, 2006	Secondary	Testing productive skills	CLIL= + productive skills	
	Serra, 2007	Primary	Language competence and mathematics knowledge	No difference in language proficiency CLIL= +at mathematics	Longitudinal study
	Ackerl,2007	Secondary	Writing	CLIL= +tenses, vocabulary, complex sentences	
	Hüttner and Rieder-Bünemann 2007, 2010	Secondary	Coherence in oral narrative tasks	CLIL= + Coherence in oral narrative tasks	
	Seregély, 2008	Secondary	Vocabulary	CLIL= + complex vocabulary, especially male learners	Covariates: gender, extracurricular exposure, study abroad
	Jexenflieger and Dalton-Puffer, 2010	Upper-secondary	Measuring general competence and task completion	CLIL= + accuracy, vocabulary and task completion	

Table 2.3 Summary of findings in Poland, Hungary, the Czech Republic and Italy

PLACE	AUTHOR& YEAR	LEVEL/ GRADES	OBJECTIVE	RELEVANT FINDINGS	OTHER COMMENTS
POLAND, HUNGARY AND THE CZECH REPUBLIC AND ITALY	Czura, Papaja, and Urbaniak, 2009	Secondary	Provide an overview of practice/Support development of CLIL	Teachers= challenging Students= a way of enhancing language learning conditions	Students' complaints= traditional methodologies, lower standards of content subjects
	Luczywek , 2009; Novotná and Hofmannova, 2007; Kovacs, 2005	Secondary	Identify and provide the best CLIL practices		Qualitative descriptive accounts
	Loranc-Paszyk, 2009	Secondary	Analyse the potential of reading and writing in relation to ss linguistic achievements	+ CLIL	CLIL vs. non-CLIL= the same exposure and comparable amount of writing
	Varkuti, 2010	Secondary	Measure BICS and CALP	CLIL= a better way of providing functional language proficiency	
	Coonan, 2007	Secondary	Teachers' perceptions	CLIL= positive impact in practice. Teacher a key figure in engaging the learner	

This section has reviewed the scope of CLIL research in several European countries. The following section will deal with the research carried out in Spain. Due to the scope of this dissertation, as with the previous section, a general overview will be provided.

2.3.3 CLIL in Spain. An overview of research

The implementation of CLIL programmes in Spain has recently experienced an exponential growth. As Coyle points out, ‘Spain is rapidly becoming one of the European leaders in CLIL practice and research’ (Coyle, 2010: viii). In the last decade, the number of CLIL programmes has flourished in the different autonomous regions. The Spanish education system is decentralized, and the different educational and linguistic policies depend on the different autonomous regional governments. The gap between EU policy and CLIL grassroots actions in Spain is bridged via regional rather than national initiatives and, therefore, the country has no single blueprint for CLIL (Dalton-Puffer, 2008).

Because of the heterogeneous linguistic situation of the country, Spanish bilingual education can be considered a many-sided issue; it is important to differentiate between those CLIL experiences in monolingual Spanish-speaking regions in which the target language is a foreign language, as in for example, Madrid and La Rioja, and those bilingual regions in which the other co-official language is also used to teach content subjects. These particular regions, for example Catalonia and the Basque country provide scope for trilingual education. Thus, ‘drawing an uncomplicated, homogeneous picture of CLIL policy in Spain is almost an impossibility’ (Lasagabaster and Ruiz de Zarobe, 2010: 284). However, considering the multiple possibilities that the diverse CLIL spectrum offers, Spain could well serve as a model for other countries seeking to implement this type of programmes (Coyle, 2010; Ruiz de Zarobe and Lasagabaster,

2010). In spite of all the differences, English seems to be the hegemonic foreign language in CLIL programmes which are no longer elitist programmes, as they are part of mainstream education in the state education system (Pérez-Cañado, 2012).

Following Fernández Fontecha (2009), different autonomous regions in Spain are fostering multilingualism using various approaches and models. In Madrid, the Balearic Islands, Cantabria, Castilla y León, Castilla la Mancha, Ceuta, Extremadura, Murcia and Navarra, the Spanish Ministry of Education, Culture and Sports and the British Council signed the ‘MEC/ British Council Agreement’, *The Bilingual and Bicultural Project* in 1995, with the aim of providing an official bilingual and bicultural curriculum. This project aims to raise the English language levels of children in state schools from the age of three. Andalucía, in turn, has implemented the *Plan de Fomento del Plurilinguismo* (Plurilingualism Promotion Plan), while in La Rioja different projects have been set up: *Proyectos de Innovación Lingüística en Centros* (School Language Innovation Projects) and Bilingual Sections. The Educational Authority in Extremadura (*Consejería de Educación. Dirección de Calidad y Equidad Educativa*) has been promoting the so-called *Proyectos de Sección Bilingüe* (Bilingual Sections Projects) in order to set up CLIL experiences in Primary and Secondary schools. Due to the scope of this dissertation, only some of the most relevant approaches and research-related studies will be dealt with in detail. Firstly, I will refer to CLIL practices and research in some monolingual communities and then I will briefly refer to the research carried out in two bilingual communities: the Basque country and Catalonia. This last one will receive its own subsection.

In general, there is a ‘shortage of research on CLIL and related practices in Spanish monolingual communities’ (Fernández Fontecha, 2009: 15). Madrid, however, stands out as one of the most active monolingual areas in terms of research. Most of the

research has been carried through different research groups based at local universities: a team led by Dr Anna Halbach at the University of Alcalá which has mainly focused on teacher-related issues (Fernández, Pena, García and Halbach, 2005; Halbach, 2009; Halbach, 2010; Pena and Port., 2008) and the UAM-CLIL Project at the Universidad Autónoma de Madrid, led by Llinares and Whittaker (cf. Llinares and Whittaker, 2006, Llinares and Whittaker, 2010); at university level, the UCM-CLUE, directed by Emma Dafouz Milne from the Universidad Complutense de Madrid has conducted the project 'Content and Language in University Education' (cf. Dafouz Milne, 2006, 2011).

Two main CLIL programmes operate at school level in Madrid: the aforementioned MEC/British Council Project and the CAM Bilingual Project (Comunidad Autónoma de Madrid Bilingual Programme) launched in 2004. According to Llinares and Dafouz (2010), the MEC/British Council Project aims at teaching curricular subject areas from a very early age using authentic materials, thus exposing children to real communicative contexts. Halbach (2009) defined the outcomes of this project as follows: the students show high concentration skills and develop higher order-thinking skills. They show collaborative working skills, greater confidence and more awareness of cultural differences. In The CAM Bilingual Project a minimum of 30% of the syllabus needs to be taught in English, and a maximum of 50%. In the academic years 2010-11, this programme reached secondary education. Academically, students within the CAM Bilingual Project obtain higher results in foreign language competence, especially in the receptive skills (listening and reading), although evidence is still not so clear for non-linguistic areas (Llinares and Dafouz, 2010). At secondary level, Whittaker and Llinares (2009) analysed data collected over a four-year period from secondary students within the MEC/British council project. Using Systemic-Functional Linguistics as the framework for the analysis, the main objective was to

describe the features of the language of CLIL students in relation to the language needs of the discipline they were studying. The authors also compared the degree of lexicogrammatical development of English in the students, and the functional realisations in their spoken and written discourse. Then, they compared the CLIL students with parallel students doing the same subject in the L1. These studies showed little difference between CLIL and non-CLIL students and revealed that the CLIL students used appropriate lexis to express content-specific ideas. When compared to their peers' performance on the same topic in the L1, some differences, such as CLIL students' use of more clauses and fewer phrases to express circumstances, were noticed. McCabe, Llinares and Whittaker (2011) focused on the development of the complexity of students' noun phrases over a two-year period and concluded that CLIL students do not control the systems that signal given and new information.

The Plurilingualism Promotion Plan (*Plan de Fomento del Plurilingüismo*) was launched in Andalusia in 2005 to foster plurilingual education. One and a half years after its implementation, Lorenzo, Casal and Moore (2010) from Universidad Pablo de Olavide, conducted an evaluation based on the Andalusian Bilingual Sections Programme on behalf of the various local administrations. Participants were organized in line with three major variables: urban/rural, primary/secondary education and L2 (English, French and German). As for linguistic outcomes and competence levels, the study concluded that CLIL learners showed greater gains than their monolingual peers, and that later start learners demonstrated competences comparable with early start learners. In terms of educational effects, the results showed that there is widespread agreement among the teaching staff that CLIL is beneficial to the educational process in general, and this opinion was shared by parents and learners alike. The authors of this

report claimed that CLIL not only promotes the integration of content and language but also fosters interdepartmental collaboration:

‘the results suggest that CLIL is an approach that not only promotes the integration of content and language, but also fosters greater interdepartmental collaboration and conflates with other language development initiatives such as Language Across the Curriculum, the genre-based approach and multi-disciplinary curricula’ (Lorenzo, Casal and Moore, 2010: 19).

The Autonomous community of La Rioja started the *Proyectos de Innovación Lingüística en Centros (School Language Innovation Projects)* in the year 2004. Most of the research carried out by the GLAUR (*Grupo de Lingüística Aplicada de la Universidad de la Rioja*) led by Dr Jiménez Catalán, has been concerned with different aspects of lexical complexity in CLIL contexts, and has compared the results to those obtained by students in EFL contexts. They have conducted interesting joined research projects with some researchers in the Basque country (Agustín Llach 2009; Jiménez-Catalán, Ruiz De Zarobe and Cenoz Iragui, 2006; Jiménez-Catalán and Ruiz de Zarobe, 2009; Ojeda Alba, 2009). The results of several studies showed significant results in receptive vocabulary size in favour of the CLIL groups. However, the results were not so definite in terms of productive vocabulary. Jiménez-Catalán and Ojeda (2010) compared the lexical ability of CLIL and non-CLIL students in their final year of Primary Education. In a lexical ability test, the non-CLIL group scored higher than the students enrolled in CLIL. Agustín Llach (2009) compared the written production of CLIL and non-CLIL learners and concluded that the non-CLIL learners used fewer borrowings than their CLIL counterparts who, in turn, made much more use of calques, coinages and lexical inventions.

As Fernández Fontecha (2009) states, the Basque Autonomous Community (BAC) together with Catalonia represent two major exponents of multilingual education in Spain due to their long tradition in bilingual education and the large body of research

associated with this concept. Members of the REAL research group (*Research in English Applied Linguistics*), such as García Mayo, García Lecumberri, Cenoz Iragui, Lasagabaster, Sierra, and Ruiz de Zarobe have been responsible for a large body of research and have contributed to attesting the success of bilingualism in the Basque Country. Garcia Mayo and Lecumberri (2003) and Ruiz de Zarobe (2005) studied the age factor and its relation to foreign language learning. Lasagabaster (2008), Lasagabaster and Sierra, (2009, 2010) have also contributed with many studies on the attitudes towards the languages learned. More recent studies from this group have focused on assessing CLIL and non-CLIL learners' general proficiency (oral skills, pronunciation, receptive and productive vocabulary, written production, morphology and syntax) as well as on subject knowledge. The results have shown no substantial differences between CLIL and non-CLIL learners; however, when differences are found, these seem to be in favour of the CLIL groups in some of the aspects tested. Jiménez-Catalán, Ruiz de Zarobe, and Cenoz (2006) compared 130 CLIL and non-CLIL primary learners in order to assess their English knowledge and use. They concluded that CLIL has a more positive effect on reading comprehension and language level instruction than on productive vocabulary. Along the same lines, Ruiz de Zarobe and Jiménez-Catalán (2009) also provided empirical evidence that there is a mismatch between receptive and productive skills. Receptive skills seem to benefit much more from the CLIL approach. A group of 89 secondary learners were compared by Ruiz de Zarobe (2008) in a longitudinal study to test their oral and written production. The students were put into three groups: a non-CLIL group, a CLIL group with one curricular subject in English, and a CLIL group with two curricular subjects in English. In order to analyze their speech production, five categories were used: pronunciation, vocabulary, grammar, fluency and content. The results were significantly better for the

five categories analysed in the case of CLIL students with two subjects in English, followed by CLIL students with one subject in English. As for their written competence, significant differences in vocabulary were only found in favour of students with more hours of CLIL. These results in writing deficiencies are in line with some of the results found by Vollmer, Lena, Randi, Debbie and Verena (2006) and Llinares and Whitaker (2006).

The results obtained by Lasagabaster (2008) in a study conducted with 198 secondary students in the BAC showed that CLIL students outstripped non-CLIL students of the same grade in the four areas tested: listening, reading, writing and speaking. Moreover, in a comparison between students of different ages (the CLIL group was a year younger than the non-CLIL one), CLIL students outperformed the non-CLIL ones except in the listening comprehension task, an outcome that was also observed in Navés (2009). The latter concluded that CLIL had a clear impact on all language skills especially when students of the same grade were compared. In order to assess the impact of CLIL on content learning, a longitudinal study was carried out with 150 CLIL students enrolled in three different levels of secondary education and a control group for each experimental group of 10 students per group. The study evaluated on the one hand the linguistic competence in English of the students and, on the other hand the level of the contents taught in English in relation to the level of the non-CLIL control groups. The results showed that students in the experimental group obtained better marks than the control groups in all tests and that the differences increased after two years.

This section has dealt with the findings of the implementation of CLIL in Spain. The next section focuses on the description and analysis of the research based on CLIL

programmes in Catalonia, the autonomous community in which data for this study was collected.

2.3.4 CLIL Research in Catalonia

As has already been explained in section 2.2, the Catalan Department of Education launched The Plan of Action for the Promotion of Third Languages (*Pla d'impuls a les terceres llengües*) in 2005. One of its most prominent strands was the so-called Experimental Foreign Language Plan (*Pla Experimental de Llengües Estrangeres*, PELE). The Plan was launched with the idea of promoting and developing the communicative language competence of students in the Catalan school system. The Plan fostered the implementation of many CLIL programmes.

Since then, research conducted in Catalonia on the implementation of CLIL has been shared by a number of research groups: the consolidated group ALLENCAM (*Adquisició de Llengües des de la Catalunya Multilingüe*), coordinated by Dr. Carmen Pérez-Vidal at Pompeu Fabra University (UPF), the GRAL research group (*Grup de Recerca en Adquisició de Llengües*) within which the BAF project (*Barcelona Age Factor Project*) has been developed, among others coordinated by Carmen Muñoz of the University of Barcelona (UB); the GREIP research group (*Grup de Recerca en Ensenyament i Interacció Plurilingües*), coordinated by Lucila Nussbaum Capdevila at the Autònoma University of Barcelona (UAB) or the CLIL-SI research group (*Semi-immersió de llengua estrangera a l'aula inclusiva*) which have developed the *artICLE Project (Avaluació de Tasques col.laboratives i assoliment d'objectius d'aprenentatge en aules 'AICLE')* both coordinated by Cristina Escobar Urmeneta; the Barcelona SLA Research Team coordinated by Dr. Carmez Muñoz, and the AICLE-CLIL BCN European Project, coordinated by Dr. Teresa Navés at the University of Barcelona. Recently a new group (CLILSLA Group) has emerged in the Universitat Autònoma de

Barcelona. It is led by Dr. Elisabet Pladevall. The research carried out by the members of this group mainly focuses on the linguistic competence of primary learners exposed to EFL and CLIL.

Catalonia has provided pioneering contributions to the subject of CLIL. Some of the articles published have addressed important issues to take into account for the implementation of CLIL: Navés and Muñoz (1999); Pérez-Vidal (1999, 2000, 2001, 2007, 2008); Escobar (2008, 2010, 2012); Escobar and Pérez Vidal (2004); Pérez Vidal and Garau 2010, Navés, 2011; Navés and Victori, 2010).

The GREIP group (*Grup de Recerca en Ensenyament i Interacció Plurilingües*), led by Dr Cristina Escobar, has mainly focused their research on the analysis of oral and written students' production in relation to class interaction in CLIL settings (Escobar and Nussbaum, 2008). Evnitskaya and Aceros (2008) tackled the same issue from a conversational perspective and Escobar and Sánchez (2009) concentrated on language learning in inclusive CLIL classrooms in secondary education.

Pérez-Vidal has addressed issues related to the acquisition of L3 English by secondary and tertiary Catalan/Spanish bilinguals of two learning contexts: a Study Abroad (SA) context spent in the target-language country and a CLIL context of acquisition. Together with members of the Balearic Community, she has carried out the so-called SALA-Project (Pérez-Vidal and Garau, 2010). SALA results contrasting CLIL programmes and formal instruction in secondary learners' groups point to the presence of significant differences in several oral fluency measures between CLIL and FI students, to the advantage of the former, especially in terms of the rate of speech, as measured in words per minute produced (Juan-Garau, 2010). The advantage of CLIL learners tends to increase over time. Pérez-Vidal's research has also focused on the relation between input and interaction in CLIL lessons. Pérez-Vidal (2007) explored the

teacher's input in relation to the nature of teaching and the presence of focus-on-form in a sample of primary and secondary lessons. The results showed that focus-on form was inexistent in most lessons, which aimed at checking and understanding subject matter.

Navés and Victori (2010) conducted a study on writing proficiency. They analyzed the impact of CLIL on students from 5th to 9th grades. CLIL learners in all five grades surpassed their non-CLIL counterparts. Navés (2011) worked with 695 learners from 5th to 12th grade. The CLIL strand obtained statistically significant differences in its favour on fluency, syntactic and lexical complexity, and accuracy. The authors concluded that 7th and 9th grade CLIL learners obtained similar results to those obtained by foreign language students one or two grades ahead. At a micro level, several studies have been carried out that have demonstrated that CLIL, under certain circumstances works. Vallbona (2009) and Victori and Vallbona (2010) conducted a study in a semi-private primary and secondary school: 5th graders, who had never received CLIL instruction before were compared with students of the same level who had received one hour a week of CLIL instruction in the subject of natural science over a period of two years. Data was gathered on their productive and receptive skills by using a battery of instruments (oral test, listening comprehension test, a dictation, a cloze test, and a written composition in both their L1 and in English). According to the results, significant differences were identified among 6th CLIL graders who outstretched their non-CLIL counterparts in listening and reading as well as in lexical complexity, fluency and accuracy. When their written skills were compared, despite the fact that the CLIL students only received a total of 35 hours a year of additional exposure to English, CLIL students also obtained better results. Bret (2011) analyzed the progress made by a subsample of students involved in the study. She found statistically significant differences among 6th graders on fluency and syntactic complexity in the in the narrative

tests. However, as well as the positive results obtained for the language competence of the students, the study also uncovered a number of problems and challenges for the teachers, most of which related to a lack of specific CLIL training, inadequate teacher language proficiency, and lack of time and resources for the implementation of CLIL. At the secondary level, Miret (2009) reported positive effects on the linguistic competence of CLIL students. Moreover, he compared the results obtained by CLIL learners with those obtained by EFL learners who were one and three years older and concluded that CLIL lessons may enhance both, the students' receptive and productive skills.

2.4 The Implementation of CLIL in Primary Education in Catalonia

As the present study was carried out in Catalonia, it is important to understand the rationale and context that underlies foreign language teaching in primary schools in this bilingual community in order to understand some aspects of the implementation of CLIL.

Primary education in the Catalan Educational System is divided into three stages: initial, middle and superior. Children spend two years in each stage. As far as the learning of a foreign language is concerned, children officially start learning a foreign language (mainly English) in their first grade of primary education, and they continue all through the other grades of primary and then onto secondary school. Nevertheless, the latest trend is for most schools to introduce children to English in pre-primary education, although this is not compulsory by law. The law strongly emphasises that in the first two years of primary education (six to eight years of age), children should develop their oral skills (listening and speaking) and, then during the following two stages (eight until twelve years of age), the four skills (listening, speaking, reading and writing) should be gradually introduced and developed.

One of the main objectives of learning a foreign language in primary education is, according to the Catalan Education Curriculum¹³, to involve children in using the language communicatively. The Curriculum goes on to say that pupils need to use the language to fit the purpose of class interaction (for example, to answer the teacher's questions about the tasks proposed by the teacher, or to interact with their peers), and also to suit the context in which the tasks take place. Children should learn a foreign language in order to be able to communicate in their target language according to the needs of different contexts and the requirements of the curriculum. Their language learning should be closely integrated with real, meaningful communication.

Mckay (2006) points out that primary school children learning a foreign language in a formal school settings develop their capacity to understand the language used by the teacher, by another student and by somebody else (for example, in a listening exercise) as well as the capacity to understand the written language in a text or in a story, even though, at the beginning, children can only respond to directives with physical movements. The children learn to create their own utterances and form sentences based on structures and vocabulary that they have learned or heard, as well as to use the language appropriately in interactions where they are practising language routines. As a result of these processes, young learners of English, according to Mckay (2006) gradually develop their listening, reading, writing and speaking skills.

The Catalan Primary School Curriculum¹⁴ stresses the idea that learning English at primary level should imply the development of four important capacities: first of all,

¹³ Generalitat de Catalunya. Curriculum d'Educació Primària (1992) Departament d'Ensenyament <https://www.gencat.cat/diari/4915/07176074.htm>

¹⁴ http://www.xtec.cat/estudis/primaria/06_curriculum_2007/llengues_pri.pdf

it is fundamentally important to develop not only the pupils' interest in the learning of a foreign language, but also their creativity and imagination in order to make them able to communicate in the foreign language and, finally their confidence and motivation in the process of learning the language. Secondly, it is necessary to develop their capacity to understand oral messages related to familiar situations that can be easily recognised by the pupils, as well as developing their capacity to read different text types adequate for their level and which allow them to obtain general and specific information. Thirdly, it is necessary to develop the pupils' capacity to produce short simple texts in a foreign language following the correct rules of the written code, which includes spelling and grammar, among many other aspects. Finally, children should use their previous knowledge and experience with other languages to progressively develop their own strategies of autonomy in language learning.

According to the directives of the Primary School Curriculum, the learning of a foreign language should also be contextualized and meaningful. The methodology used should be active and participative so the students can see from the very beginning that what they are learning is useful. To make the learning clear and useful, it is necessary for the students to be motivated and for them to show a favourable attitude. The nature of what they learn has to be functional - in other words, the students should be able to use what they are learning when the situation requires that they do so, and the four basic skills (oral comprehension, oral expression, written comprehension and written expression) should be taught in an integrated way. Finally, it is also highly recommended that educational authorities should promote a relationship with other areas of the school curriculum in a way that enables the foreign language to be used as an instrument to express the contents of other subjects.

For all these reasons, the teaching of a foreign language in primary education has quite often, but not always, been related to content-based teaching, through topic-based and task-based units such as dealing with everyday routines and familiar topics. Hence, it is not surprising that CLIL has been perceived as an effective way of enhancing children's understanding, communication and motivation in a foreign language. CLIL programmes introduced at an early age may bring even more important benefits, most of which are supported by SLA findings.

This chapter has dealt with the main theoretical principles that underly the CLIL approach. It has also outlined a range of important characteristics that schools should take into account before implementing a CLIL programmes. An overview of research in CLIL in Europe as well as in Spain and Catalonia has also been presented. As the participants in the study were young learners in their final two years of primary education, a chapter on young learners is in order.

CHAPTER 3 Young Language Learners

In recent years there has been a growing global trend to lower the starting age of foreign language instruction at school, not only in Europe but in other continents as well. Because of globalisation and mobility, Europe is becoming increasingly multicultural and multilingual. In response to the changing social and linguistic reality, children have started learning a foreign language well before adolescence in many countries. In Catalonia, for example, where the study was carried out, foreign language instruction, officially begins at the onset of primary education, although most schools provide contact with this language to students from as young as three. Despite the recommendations of European institutions to teach different foreign languages, English is at present and by far the most widely learnt foreign language on the continent. The worldwide spread of English Language Learning programmes has been considered the outcome of English becoming the lingua franca for international communication (Graddol, 2006).

According to Nikolov and Djigunovic (2011: 95), ‘the number of young learners and their teachers has recently experienced an unprecedented increase’. This rapid growth has resulted in a considerable number of language policy documents and programmes, as well as a rise in the number of teachers’ handbooks and teaching materials adapted to the needs of this type of learner. It has also led to a growing body of empirical studies devoted to the topic of foreign language learning as seen from different perspectives and in different contexts.

Before concentrating on certain aspects that may affect the process of learning and acquiring a language on the part of young learners, an overview of the main characteristics that define young learners is in order.

3.1 Young Learners. Definition and Characteristics

Although the meaning of the word *young* varies in different contexts, within the European Union Member States the term Young Learners (YLS) is used to refer to primary students between approximately seven and twelve years of age. However, in some contexts, even 14-year-olds are considered to be within the YLS group (Nikolov and Djigunovic, 2011). Children below this age are often referred to as Very Young Learners (VYLs).

Assessing the language learning of YLS requires some knowledge of two fundamental aspects: knowledge of the general characteristics of young learners and knowledge of the characteristics and processes of language learning in general (Mckay,2006). During their primary years, children are in a state of constant cognitive, social, emotional and physical growth. They experience a progressive state of change in contrast to what Bialystok (2001) called the stable state of the adult mind. Their cognitive abilities associated with memory, reasoning, problem-solving and thinking, as well as their social and emotional skills, emerge in a continuous way throughout childhood. Following Piaget's (1965) stages of cognitive development, all children go through an initial period of time during which their knowledge of the world is limited to their sensory perceptions and motor activities. In the period between approximately two and six years of age, they begin to use their language; however, they still do not understand concrete logic and they cannot manipulate information. Between the ages of seven and eleven, when children move to upper-elementary grades in education, they go through the *concrete operational stage* during which they gain a better understanding of mental operations. They begin thinking logically about concrete events, but still have difficulty understanding some abstract or hypothetical concepts. It is from around the age of twelve that children start to develop the ability to think about abstract concepts.

Other skills such as logical thought, deductive reasoning and systematic planning also start to emerge around this age. Although Piaget's original ideas have been challenged and criticized, most developmental psychologists would agree on the existence of stages of development, although not necessarily as rigid as they were described and presented by Piaget himself. From a teaching-learning perspective, according to Pinter (2006) and Cameron (2001) what teachers should know about Piaget's theory is that YLs are active learners and thinkers who construct their own knowledge, and who are trying to make sense of the world by seeking out intentions and purposes in other people's actions and language. Language teachers dealing with young learners need to be aware that the needs and challenges of children change and evolve throughout their primary schooling, and that although children are constantly trying to make sense of the world around them, their own experience of the world is still very limited; this limitation should be reflected in the teachers' own practice, their classroom speech, their choice of materials and activities as they will inevitably affect the way the children respond to some tasks. Following Piaget's idea that 'children adapt through experiences with objects in their environment, Cameron (2001) suggests that the classroom and classroom activities should provide learners with plenty of varied opportunities for learning.

While Piaget concentrated on the biological basis of developmental progression in every child, the Russian psychologist Vygotsky (1978) emphasized the social aspects of development. Although he shared with Piaget the idea that children build knowledge by themselves, and that they participate in their own learning process, he focused on the importance of social interaction and described how culture and the social context shape the learning process. He was very interested in the learning potential of each individual because for him, all learners are unique individuals. Vygotsky considered learning to be one of the fundamental mechanisms of development. In the model of learning that he

provides, the context occupies a central position, and social interaction becomes the engine of development. He introduced the concept of Zone of Proximal Development (ZPD), which is the distance between the actual developmental level and the level of potential development. Learning and development are two processes that interact and they take place in a social context. The school needs to be consistent with the child's developmental level. Learning to do things and learning to think are both helped by interacting with adults and they occur more easily in collective situations. Children can do and understand much more with the help of adults than by themselves, because knowledge is not an object that is passed from one to another, but rather something that is built through operations and cognitive skills that are induced in social interaction. Vygotsky notes that the individual's intellectual development cannot be understood independently from the social environment in which the person is situated. For Vygotsky, the development of higher mental functions should be considered first on the social plane and then on the individual level.

Cameron (2001) suggests that Vygotsky's ideas are important for the construction of a theoretical framework for teaching foreign languages to young learners. In a classroom setting, the teacher is responsible for structuring interactions and developing instruction in small steps based on tasks. The instructor is also charged with providing support until the learner can move through all tasks independently. In order for teachers to guide learners through the tasks associated with learning a concept, they must "understand how the cognitive aspects of the tasks fit into the child's level of development. These tasks are called "scaffolds," and are tasks or levels on which the teacher builds to develop the learners' zones of proximal development. For Pinter (2006:129) 'the language used in interactions is fundamental as it is the vehicle through which understanding and learning takes place. Language, then, has a crucial role in

learning and, therefore, it has important implications for teacher-talk in all classes, including EFL classes. Listening to the teacher provides children not only with a model of pronunciation but also with an array of opportunities to understand input from context. However, children also need to interact in classes with both the teacher and the other learners. The teachers, then, need to create as many opportunities as possible for interaction in their classes by creating and using adequate tasks in order to elicit language from the children, as well as to scaffold their early language production so that they are able to use the language meaningfully both with the teachers themselves and with the other learners.

Bruner (1983) developed the concept of scaffolding which was then transferred and applied to the language classroom and teacher-talk. For children learning a language, it is of fundamental importance to be provided with a framework that allows both cognitive growth and language development. Based on Vygotsky's ideas, Wood (1998) suggested different ways in which the teachers can scaffold children's learning: attending to what is relevant, adopting useful strategies and reminding the children frequently during the lessons about the whole task and the goals. Wood's strategies are easily applicable to the language teaching context, because it is within these formats that children learn how to use the language: by pointing their attention to relevant aspects through strategies and useful tasks, the teachers are helping them do what they are not yet able to do by themselves. Young learners are easily distracted; their attention span is short and limited and they are distracted by their peers or by the context itself. Therefore, it is helpful for them to be reminded of the objectives and goals of the tasks they are involved in. As they grow older and language becomes more complex, if this language is within their ZPD, then the scaffolding process turns out to be a useful tool

that facilitates both language learning and helps them to make sense of new language forms in context.

Learning at primary level is associated with an active process in which students interact with other students, with the teachers and with their own environment. Despite that similarly-aged children share certain common characteristics, and that interaction with teachers may provide them with enriching learning experiences, every child at school is unique and the characteristics of each of them should be known to the teachers; Gardner (1983), in his Theory of *Multiple Intelligences*, suggested that intelligence manifested itself in various forms in different children. He defined eight types of intelligence: linguistic, musical, spatial, bodily kinaesthetic, interpersonal, intrapersonal, and naturalist. These different types of intelligences present a pluralistic panorama of learners' individuality and suggest the idea that early language learning needs to address the issue of individual learning differences and different learning styles.

This section has summarized some of the general characteristics and some of the cognitive development of young learners and the learning process in general. The next sections will address the characteristics of young foreign language learners and will also refer to three individual factors that are considered important in the process of learning a language: age, motivation and aptitude.

3.2 The Child as a Language Learner. Influential Factors

First language acquisition is a long process that continues well beyond childhood. Child learners who come into foreign language learning at a very early age are still in the process of fully developing their mother tongue. Therefore, they bring to the classes differently developed literacy skills and learning abilities in their first language. During the school years, their approach to language learning develops from a

holistic approach to a much more analytical one. Although at the beginning they can only understand the general meaning of the messages provided, in the final years of their primary schooling, they develop an interest in language as a more abstract system which allows them to start comparing patterns and linguistic forms in two or more languages. Their growing abilities in their mother tongue will contribute directly or indirectly to the process of learning a second/foreign language.

3.2.1 Age and the Critical Period Hypothesis

There is a folk belief in foreign language learning which has often been used to support the introduction of foreign language learning at a very young age: *the younger the better*. Because of their maturational growth and the development of their cognitive capacities and skills, young learners are perceived to acquire languages differently from adolescents and adult learners (Scovel, 2000; Nikolov, 2009). Age is, amongst others, one of the aspects that has attracted the attention of a great number of SLA researchers, and one that has also led to a great deal of controversy around the world, mainly because it has often been claimed that child L2 learners generally perform better in the long run than adult language learners (e.g. Patkowski, 1980; Krashen, Scarcell and Long, 1982; Felix, 1985; Johnson and Newport, 1989; Bley-Vroman, 1989; Slavoff and Johnson, 1995)

The Critical Period Hypothesis (CPH) (Lenneberg, 1967) is considered a relevant point in the discussion of young language learning. The major claim of this hypothesis is that natural language acquisition is available to young children due to the existence of a biologically determined period in life, which seems to finish around puberty, during which maximal conditions for language acquisition exist. Therefore, adolescents and adults would have limited or no access to these conditions. The most influential explanations of this hypothesis have come from different perspectives,

namely, the neurological and the developmental-cognitive ones (Muñoz, 2006). Neurological explanations have supported the idea of the reduction of brain plasticity which is attributed to a process which affects neurons and is linked to maturational stages (Long, 1990; Pülvermüller and Schumann, 1994). A recent view, supported by a number of cognitive psychologists has placed emphasis on brain structure and its organisation as the cause of L2 acquisition (Perani and Abulatevi, 2005). Other hypotheses have also been involved in the long-standing debate on whether age-related aspects constitute a crucial factor in L2 acquisition: the Competition Hypothesis (Felix, 1985), which claimed that post-puberty children can still access their in-born mechanism for L2 acquisition, and the Fundamental Difference Hypothesis (Bley-Vroman, 1989) for which the inborn mechanism that children have is no longer operative in adulthood and that adults rely much more on general problem-solving strategies. A detailed discussion of any of these hypotheses is beyond the scope of this dissertation.

According to Nikolov (2005), recent research into the CPH has mainly concentrated on two different perspectives: the rate of acquisition and ultimate attainment. Some researchers have interpreted research results in favour of the existence of CPH (DeKeyser, 2000, 2003; DeKeyser and Larson Hall, 2005), whereas others, have concluded against it (MacWhinney, 2005). Long (1990) argues in favour of the existence of sensitive periods governing language development, during which the acquisition of specific language abilities is more successful, and after which language acquisition is somehow incomplete. He emphasizes that the age-related loss of language learning ability is cumulative rather than a one-shot event affecting several linguistic domains successively; this loss is not limited to phonology which has always been claimed to be one of the areas most affected by the existence of the critical periods.

Overall, Long argues that even though there is no single critical age, the ability to learn a second language weakens with time and the possibility of reaching the native-like ultimate attainment decreases.

The results of several studies have confirmed that older age is an initial advantage in the rate of acquisition but a disadvantage in ultimate attainment, and that adults progress faster in areas of morphology and syntax than children do (Juffs and Harrington, 1995; Harley and Wang, 1997). These studies also support the claim that older learners progress faster than younger ones (Marinova-Todd et al., 2000) but child starters outperform adult starters in the long run (Singleton, 2005). In spite of these findings, it must be mentioned that most studies have relied on data from naturalistic second language acquisition, that is, in contexts of full immersion in the language community and have been generalized in instructed educational contexts in which the foreign language is just one of the school subjects and exposure to the language is limited to the classroom. These types of school contexts are referred to by Larson-Hall (2008) as ‘minimal input situation contexts’. It should be noted, however, that recent classroom-based research carried out in contexts with limited input has presented evidence for better performance of older learners (Cenoz, 2003; Muñoz, 2006).

Nevertheless, the possible age limits and the contextual factors constitute essential information to consider when deciding not only the age at which the teaching of a foreign language should start, but also, and equally important, the intensity required at different moments of the process and the type of programme that needs to be implemented. Although one of the most widely accepted views of the CPH is that it emphasizes the need to begin L2 acquisition before the end of the critical period hypothesis (Muñoz, 2006), there is no guarantee, in foreign language contexts, that all children will attain a high-level of proficiency because of the existence of many factors

contributing to ultimate attainment (Long 2005). Along the same lines, Lightbrown and Spada (2006) point out that achievement levels vary from programme to programme, and to a great extent, they depend on the goals of the programmes and their participants: if the goal is to achieve communicative ability in the language rather than native-like proficiency, the benefits of an early start are, for these authors, still not very clear. They suggest considering how individual factors may affect the process of language learning as well as the different contexts in which instruction takes place (see sections 3.2.3 and 3.4.1 for a detailed description of the importance of the context and a description of YL programmes).

3.2.2 Motivation and Aptitude in Young Language Learners

Ellis (1994) refers to language aptitude and motivation as two important individual aspects that may also affect language learning. Extensive research has been carried out with adult learners on the implications of these two aspects in language learning. From the perspective of young learners, motivation has recently been a popular topic among researchers as it has been considered a fundamental element in the success of foreign language programmes for young learners. As stated by Blondin et al. (1998), some of the main gains in early language learning lie in the development of positive attitudes and motivation towards the target language. Studies of an early start have shown that motivation and positive attitudes to a second language can readily be fostered in young learners and they also seem to suggest that young language learners show more positive attitudes and motivation than older learners do. However, in spite of early interest for language learning, this interest seems to wane in time (Burstall, 1975; Nikolov, 1999; Heining-Boyton and Haytema, 2007). The findings of a longitudinal study carried out by Mihaljevic Djigunovic (1998) revealed that, under favourable learning conditions, high motivation and positive attitudes can be maintained over

extended periods of time. Along the same lines, Nikolov (1999) followed three groups of young learners taught by the same teacher in an eight-year longitudinal study. She found that the learners' motivation could be maintained by intrinsically interesting and cognitive challenging tasks and that the attitudes of young learners could be shaped by what happened in the classes. In a similar study in Hungary, Hardi (2004) concluded that instrumental motivation was higher in those students studying English as an optional subject than in those students enrolled in compulsory EFL learning. Marschollek (2002) found that primary students maintained their motivation throughout elementary schooling. Other researchers, however, have investigated changes and fluctuations in motivation with students who started learning English at school at different ages and have found no age-related differences in motivation (Lasagabaster, 2003; Muñoz and Tragant, 2001). Cenoz (2004), however, in a study carried out in the Basque country with students learning English with a variety of starting ages, concluded that earlier starters were more motivated than late starters. Tragant (2006) suggests that, considering the general pattern that emerges in terms of motivation, there seems to be a decline in positive attitudes towards puberty, and that young learners' biological age may have a higher effect on motivational orientations than hours of instruction. Investigations carried out recently within the framework of a large-scale European project on Early Language Learning in Europe, the ElliE project, also found instances of declining motivation in the learners involved in this project. In spite of the high initial motivation that most of the learners showed in the questionnaires provided and that girls were more motivated than boys (Szpotowicz, Mihaljevic Djigunovic and Enever, 2009), a slow decrease in motivation was found towards the period of the end of the study (Lopriore, 2009; Mihaljevic Djigunovic, 2010). Mihaljevic Djigunovic and Krevelj (2010) concluded that some Croatian learners, admittedly under less than ideal

conditions (large groups, two lessons per week, unqualified teachers), which sometimes are common ground in some European contexts, started to develop negative attitudes towards the L2 as they got older.

Motivation is not the only factor that may affect the process of learning a language at a young age. Another individual factor that has been claimed to affect language acquisition and that has also been associated with proficient performance in language learning is aptitude. One of the pioneers in the area, Carroll (1991) defined aptitude as the ability to learn quickly. More recently, the notion of Foreign Language Aptitude (FLA) has been associated with the stable talent for learning a foreign language which seems to be between individuals' (Dörnyei and Skehan, 2003). Learners with high aptitude are considered to learn with greater ease and speed than other learners who may also be successful, however, if they persevere. Whether FLA is fixed/innate or amenable to training has been the departure point for most research in this area. It is worth mentioning, that over the last 15 years, the notion of 'aptitude' had developed from being seen as a stable unitary fixed trait to being considered a dynamic and multifaceted set of malleable abilities (Larsen-Freeman, 2001) that interact with other internal learner attributes and attitudes such as motivation and learning styles (Dörnyei, 2009), as well as with contextual factors.

In the context of young learners whose linguistic abilities are still developing, language aptitude cannot be seen anymore as a fixed set of specific abilities (an 'able' or 'less able child'). It seems more likely that YL aptitude for language learning is related to the learners' general cognitive abilities that may account for significant variation among learners. As cognitive skills develop with age, aptitude appears to improve. By the age of twelve, aptitude seems to account for a great deal of variation in language performance (Edelenbos and Johnstone, 2009). These authors go on to say that

the profiles and aptitudes of the students enable them to progress and learn more effectively if they are correctly identified and appropriately managed in the foreign language classes. Therefore, in terms of pedagogy, the variation in aptitude has clear implications for language instruction (Wesche, 2001; Erlam, 2005). Lightbrown and Spada (2006:54) support the idea that ‘knowing the strengths and weaknesses of learners can help teachers ensure that their teaching activities are sufficiently varied to accommodate learners with different aptitude profiles’.

One of the objectives of teaching foreign languages to young learners is to offer them a stable environment that allows them to move on from one educational stage to another, and to make progress within each educational stage. Children seem to progress through a sequence of stages in their internalized language development at differing rates. Their natural progression through the different stages does not seem to be unidirectional and smooth and seems to include periods of stagnation (plateau-effect or fossilization) or even periods of confusion (Edelenbos and Kubanek, 2007).

Taking as a starting point the idea that aptitude for general learning and for language learning in particular is not fixed from birth and can be developed from primary schooling (Enever, Moon and Raman, 2009), it is therefore important that foreign language programmes not only develop skills of communication in the target language but also the underlying qualities that go with aptitude in order to make the most of them. Blondin et al. (1998) pointed out that some children do not seem to be able to progress beyond the use of prefabricated utterances. However, he also suggests that, in order to help learners improve their command of the language beyond prefabricated utterances, language teachers dealing with young learners may need to combine activities involving the use of formulaic, learnt expressions with those that focus on accuracy and form. Supplementing communication activities with activities

that help children internalize concepts about language is strongly advisable in order to enhance the children's sensitivity to the language, to its underlying patterns .

In summary, age, motivation and aptitude are three aspects to be considered when dealing with young learners' foreign language programmes. However, as has already been mentioned at the beginning of this section, the outcomes of such programmes are also affected by other aspects including the context, the programme the learners are exposed to and the teaching practices. The next section will deal with the concept and the importance of context in language learning.

3.2.3 The learning context

Learning in general takes place in context and language learning is not an exception. Understanding the contextual dimension is essential in order to get to know how opportunities for learning are created and how the learners respond to these opportunities. Hymes (1972) observed that knowing what goes on outside the school setting is necessary to understanding what goes on inside. He noted further that 'the key to understanding language in context is to start not with language but with context ... and then to systematically relate the two' (p.9). According to Housen, Schoonjans, Janssens, Welcomme, Schoonheere and Pierrard (2011), two relatively recent coexisting lines of research have contributed to the overall picture that researchers and pedagogues have on SLA. On the one hand, scholars such as Long (1997) and VanPatten (1990) have emphasized the cognitive nature of the SLA process, and the importance of providing an understanding of the acquisition process in psycholinguistic terms relatively independent of external contextual factors (such as sociolinguistic variables or the particular methodology employed in a classroom); others contend that the best predictive models of SLA take into account the interaction of social activity and psycholinguistic elements (Collentine and Freed, 2004). Housen et al. (2011) consider

second language acquisition as a socio cognitive process opened to many external and contextual factors that must be taken into consideration. Traditionally, the study of the learning context has been characterized by dichotomies of the type *naturalistic vs classroom contexts*, *second language vs foreign language contexts*, *study abroad vs study at home*. The underlying assumption has been that naturalistic, second language and study abroad contexts result in more proficient learners as they provide them with more opportunities for language learning than classroom, foreign language or study at home contexts. Although some elements of this assumption have been backed up and validated by research, not all the studies have dealt with the same notion of context as a determining variable in SLA. The idea of context has been defined in many different ways. For some researchers, the learning context only includes the input and output opportunities that are available to learners, while, for others the notion of context goes further. Ellis (1994:197) talks about context as the ‘different settings in which L2 takes place’. In each setting, different social factors, such as status, sex and learner identity interact and influence the learning process and its linguistic outcomes (Ellis, 2008). Garcia (2009), in turn, sees the learning context as a dynamic construct which is influenced by physical, social, cultural, psychological, discursive, linguistic and cognitive factors where the learning takes place.

Housen et al. (2011), consider that the concept of context can be build up by three overlapping and intersecting levels: *the learners’ individual context*, *the educational or curricular context* and *the extra-curricular context*. The learner’s individual context includes the learners themselves and their needs, preferences, abilities, knowledge, and personal characteristics, all of which interact with external factors. The second level described by Housen refers to the educational or curricular context. This level goes beyond the individual and refers to aspects such as school

language policy and its implementation which, in turn depends on the structure of the education system, the design of the curriculum and the particular pedagogical practices which are in place. Classroom practices, pedagogical methods and resources for language teaching are considered to be fundamental elements within this level which have clear implications for the teachers and the way they plan and deliver their lessons, as well as ultimately, for the learners themselves. This is due to the fact that the teachers' practices, their methods and materials affect the input and output opportunities that are provided in the language classes. Within this level, Housen et al (2011) distinguish between *language-content classrooms* and *language-subject classrooms*. In the *language-content classrooms*, the L2 functions as the medium of instruction and communication and, secondarily, as the object of learning; whereas, in the second type, in language-subject classrooms, the L2 is primarily the object of learning. The third level of the learning context is the extra-curricular, context which is beyond the control of curricular intervention. It includes socio-demographic, cultural and institutional aspects both inside and outside the school. Two sublevels can also be distinguished within this level: the school level which refers to the opportunities for L2 exposure within the school or through extracurricular activities and the community level, that is, extramural exposure to the L2.

This section has dealt with the notion of context and its purported influence in the process of learning a second language. The next section will focus on the learner and some of the psycholinguistic principles that affect the process of learning a second language.

3.3 Review of some Psycholinguistic Principles in SLA

Although cognitive theorists acknowledge the importance of sociolinguistic factors in SLA, they stress the fundamental role of the cognitive processes in language

learning. Understanding cognitive language learning processes helps researchers, teachers and assessors to make judgements about the performance of the children and to act accordingly (Mckay,2006). Muñoz (2006) refers to four essential psycholinguistic components that should be taken into consideration: exposure to input, the processing of meaning, and the processing of form and language production. These principles correspond to the three different stages of information processing postulated by Skehan (1998): input, processing and output.

3.3.1 Exposure to Input. The Input Hypothesis

Language acquisition cannot occur without an input of some sort. According to Krashen's input Hypothesis (Krashen, 1985) exposure to abundant input is a fundamental requirement for the acquisition of a second language. Acquisition takes place when the learner is exposed to 'Comprehensible Input' that belongs to level 'i + 1', that is, when the learners are exposed to input that is slightly beyond their own level. Therefore, input is an essential ingredient for language acquisition to take place. However, as not all learners are at the same level of linguistic competence at the same time, Krashen suggested that *natural communicative input* is the key to designing a syllabus, ensuring in this way that each learner will receive some 'i + 1' input that is appropriate for his/her current stage of linguistic competence. His Acquisition-Learning distinction is the most fundamental of the entire hypothesis in his theory. Acquisition refers to the unconscious development of the target language system as a result of using the language for real communication. It requires meaningful interaction in the target language, that is, natural communication during which speakers are not concentrating on the form of their utterances, but rather in the communicative act itself. Learning is the product of formal instruction, and it comprises a conscious process which results in conscious knowledge about the language, for example knowledge of grammar rules. For

Krashen, due to the so-called 'non-interface position', there can be no interaction between these two independent knowledge systems: learning that has resulted from instruction cannot lead to acquisition.

Despite the notable impact that his theory had on language acquisition, it received a great deal of criticism. Although intuitively attractive, Krashen's model was extremely criticised as lacking theoretical or empirical foundations. Gass (1997) focused on the idea that incomprehensible input triggers learners' awareness of gaps in their knowledge. Long (1983) looked into the notion of comprehensible input and how input had to be negotiated in conversations between native speakers and non-native speakers. He concluded that there is no learning without comprehensible input. It is from this perspective that the hypothesis holds. The need for input in language learning has remained uncontroversial and has been defended from different angles (Dekeyser, 2000; Muñoz, 2006). According to this last researcher, apart from comprehensibility and quantity, input also needs to fulfill other conditions: it has to be authentic, contextualized and used in real communicative situations.

This highlights the importance of using the target language in the foreign language classroom. The main goal of any language program is for learners to be able to communicate effectively. By providing as much comprehensible input as possible, especially in situations when learners are not exposed to the target language outside the classroom, the teacher is able to create a more effective opportunity for language acquisition.

3.3.2 The processing of Meaning.

As Krashen pointed out, being exposed to comprehensible input is necessary but not sufficient for language learners because the learners need to be able to process the input they receive. In the words of Muñoz (2006:18), 'input is the necessary catalyst

through which language is processed and which results in changes in the learners' linguistic system'. However, this account of the importance of input does not take into account other components that also contribute to language processing, above all memory. Following Atkinson and Schiffrin (1968) two major memory stages are involved in the process of learning a language: the short-term memory (STM) which is considered to be limited in capacity and the long-term memory (LTM), that is, a vast store of knowledge and a record of prior events whose capacity seems to be much larger. The term 'short term memory' was substituted by the term 'working memory' (WM) (Gathercole and Baddeley, 2003) to introduce the idea of attention-related processes attached to it as well as to consider the contribution of the working memory in extracting, organizing and grouping input that is relevant to understanding. In the words of Skehan (1998: 45), the WM can be considered as 'a storage area while the different elements of a message are being orchestrated'. For learning to take place, information needs to be temporarily stored in the short term memory. This information can easily be retrieved because of its high activation; however, if it is not rehearsed, it will easily get lost. As the capacity of the STM is limited, the information that is properly rehearsed is moved to the LTM which registers information that is durable, although retrieving it is much more difficult and slower (Anderson, 1985).

The existence of a dual mode system to process the information that learners receive has been confirmed by recent cognitive research (Ullman, 2005). As Skehan (1998) suggested, the formulaic system (also called exemplar-based system) and the rule based system exist side by side. He argues that the central role in the formulaic system is occupied by a large system based on chunks, that is, words, groups of words or formulaic units that may have a structure but are unanalysed. Although limited, as the learners can only use the expressions learnt in the adequate context, the value of the

system is that it allows learners to communicate faster than they would do it using their knowledge of grammatical structures. By using formulaic language as a strategy, the learners, especially young ones, drive themselves to try to reach something that they have not yet understood, thus, pushing themselves to increased proficiency (Mckay, 2006). The learners also use the rule-based system, when they draw on rules of the language in order to construct sentences. However, this system in young learners is bound by their limited capacity to use knowledge about grammar. As Skehan (1998:89) suggests: ‘Clearly, neither the rule-based nor the exemplar-based system is ideal separately’. Language learning takes place when learners are engaged in meaningful communicative situations that activate both, their formulaic system and their rule-based system through focus-on form which, as they grow older, allows them to ‘notice’, that is, to become aware of the form of language. The next section will deal with the concept of noticing as stated by Schmidt (1990) in her output hypothesis.

3.3.3 The processing of Form. The Noticing Hypothesis

As has already been explained, several authors (Krashen, 1985; Long, 1985) have described input as a fundamental ingredient in language acquisition as it triggers acquisitional processes and drives interlanguage development. However, learning a second language has been proved to be only partly input dependent as it not only requires comprehensible quality input but also the processing of forms. Two main models of input processing have analysed the functioning of input in terms of attention processes: Van Patten’s model of processing (1996) and Schmidt’s noticing Hypothesis (1990).

According to Van Patten (1996), in comprehension-based models, as the one proposed by Krashen (1995), comprehension takes priority over form, whereas processing-based approaches focus much more on the conditions under which the

learners may attempt to make connections between forms in the input and meaning. He proposes three main principles for input processing. His first principle postulates that the learners process input for meaning before they process it for form, that is, they process content words before grammar words. His second principle states that prior to processing non-meaningful grammar forms, the learners need to be able to process communicative content at no cost to attention resources. His third principle suggests that learners must be taught strategies in order to help them to process the input. He argues that the processing approach is useful for learners to make the necessary links between form and meaning. For him, when the learners receive some input, they can deliberately attend to aspects of form which are then incorporated into a developing interlanguage system in order to relate the meaning to particular aspects of form. This input processing takes place at the input-intake stage. The focus of instruction at this stage aims at maximizing the efficiency of this stage so that acquisition processes take place effectively and contribute to output processes.

Schmidt (1990) argues that noticing is a necessary condition for L2 acquisition. His noticing hypothesis states that 'what learners notice in input is what becomes intake for learning' (Schmidt,1990:20); whether a learner deliberately attends to a linguistic form in the input or it is noticed unintentionally, if it is noticed, it becomes intake. For Schmidt, frequency and salience are two aspects that need to be considered in his hypothesis: the more frequent a form, the more likely it is to be noticed and integrated into the interlanguage system. Along the same lines, the more salient a form is, the more likely it will be noticed. He argues that forms are brought into awareness through instruction. Instruction makes aspects of input become noticeable and analyzable by providing structured, differentiated input that assists noticing by focusing attention on language features, and thus enhancing awareness (Skehan, 1998) The role of instruction

does not necessarily rely on the clarity of the explanations provided but on the way it channels attention to aspects that otherwise would probably be missed by students. A language feature may become frequent due to repeated instruction or through teacher talk. As such, when the item does appear more frequently in the input, the likelihood that an item will be noticed and integrated into the interlanguage system is increased (Schmidt, 1990). Noticing also impacts on memory and has a mediating role between input and memory systems. It is inside the short-term memory that noticing must in reality take place (Robinson, 1995). Some learners are better at processing input because they either have a larger working memory capacity or because they are better at analytical processing within the working memory (Skehan, 1998).

Apart from the aspects mentioned, Schmidt's hypothesis also involves other aspects directly related to the individual learner: skill level, task demands and comparisons. The noticing ability changes from learner to learner and their capacity to routinize previously-met structures varies. Their level and ability in the different skills will determine if they are ready to notice new forms in the input. Noticing will also be influenced by the demands of the task, that is, the way in which a task makes learners notice particular aspects that are necessary in order to carry it out (Schmidt, 1990). Ellis (1997) suggests that language features may be made intentionally prominent or the task may be designed to force learners to process the language. Skehan (1998) points out that noticing may be more or less likely depending on whether the level of processing that the task demands is low or high. However, noticing alone is not enough for input to become intake (Long, 1995). Learners need to be able to 'notice the gap' between their perceived input and their output production based on their current interlanguage system.

3.3.4 Language Production. The Output Hypothesis

The previous section has put forward the idea that input processing starts in the working memory and is then transferred to the long term memory in which it is assimilated into the second language system, causing the reorganization of the system. This section will deal with the last essential component in language learning: output production.

Following her research with Canadian students in content-based second-language French instruction, Swain (1995) and Swain and Lapkin (1995) formulated the Output Hypothesis in contrast to Krashen's input hypothesis. Swain argues that although input is necessary for language learning, it is insufficient, and that the learner also needs to produce language. Producing language constitutes part of the process of SLA.

Swain and her colleagues realised that students exposed to French-medium instruction for extended periods of time, achieved almost native speakers' level in comprehension abilities (reading and listening) whereas their productive ability lagged behind. These findings raised doubts about the validity of Krashen's input hypothesis and about the argument that input was the only element for SLA. Through class observations, it became obvious to the researchers that teachers did not 'push' students to talk in an accurate and sociolinguistically appropriate manner. Swain and her colleagues attributed it to the fact that the students' use of French in the classes mostly involved reading and listening to second language input, and that the teachers did not 'push' the students to speak or write in French at a high level. Swain concluded that 'simply getting one's message across can and does occur with grammatically deviant forms and sociolinguistically inappropriate language. Negotiating meaning needs to incorporate the notion of being pushed towards the delivery of a message that is not

only conveyed, but that is conveyed precisely, coherently and appropriately. Being ‘pushed’ in output is a concept parallel to that of $i+1$ of comprehensible input’ (Swain, 1985: 248-49).

The output hypothesis refers to three functions for output:

1. The notice/triggering function which allows learners to become aware of gaps in their interlanguage. It may bring their attention to something that they need to discover about their second language. With the use of this function, the learners realize there are some linguistic problems they need to manage. The notice function pushes the student to look for the adequate knowledge they require to complete the newly discovered gap (Swain, 1995; Doughty and Williams, 1998).
2. The hypothesis-testing function, which allows learners to try out language and reflect on their hypothesis of how to say or write their intents. This function suggests learners may use the method of “trial and error” to test their production expecting to receive feedback. This feedback can be applied in two ways: recasts and elicitation or clarification requests.
3. The metalinguistic function, which allows learners to discuss their hypotheses with others and reflect upon them

In summary, pushing learners to produce language puts them in a better position to notice the ‘gaps’ in their language knowledge and encourages them to upgrade their existing interlanguage system. Furthermore, as they are pushed to produce language in real time and thereby forced to automate low-level operations by incorporating them into higher-level routines, this may also contribute to the development of fluency.

According to Muñoz (2006), the output hypothesis also highlights the importance of error correction in the classroom. In classes in which attention is exclusively focused on meaning, the reformulations carried out by the teachers are not

necessarily processed by students from the perspective of form, as they may not notice them and, therefore, they do not contribute to the correction of their linguistic deficiencies. Lyster (2004) and Lyster and Ranta (1997) investigated the reaction of teachers to students' errors in communicative lessons and concluded that, in meaning-focused classes, the teachers responded with reformulations of the learners' incorrect productions rather than overtly corrected errors of form.

This chapter has so far provided a description of the theoretical background to young learners, their characteristics as learners and as language learners and has dealt with some the fundamental aspects that may affect their language acquisition process. In the following sections, the process of learning English by young learners will be developed from different perspectives: young learners programmes, the teacher factor, the teaching of English as a foreign language and content-based programmes.

3.4 Learning Foreign Languages at an Early Age.

In the last decade, an early start in the teaching of modern foreign languages has been one of the main priorities for the educational authorities in the majority of the European countries. This trend has been supported from the very beginning by the European Union Action Plan on Promoting Language Learning and Linguistic Diversity: An Action Plan 2004-2006 (2003)¹⁵. Educational authorities have spared no efforts in designing, developing and implementing language programmes for young learners (Nikolov and Djigunovi., 2006). Early Language Learning Programmes are, however, by no means uniform. They vary according to when they start, how much time they allocate to language learning, the curriculum they apply, the teachers involved and the programmes that are implemented. Their introduction entirely depends on the educational framework of a given country as well as on the status of the target language

¹⁵ http://ec.europa.eu/languages/eu-language-policy/action-plan-for-languages_en.htm

in each particular context. In spite of the differences, Pinter (2006) refers to the fundamental aims and objectives that most educational systems emphasize as the basis of the implementation of YLL school programmes. For her, the three most important objectives in the majority of the programmes are: to develop children's communicative and their cognitive skills and to raise their metalinguistic knowledge. Young Learners programmes also aim at encouraging and motivating children towards language learning and at developing 'learning to learn skills'. Nevertheless, the overall achievements of these types of programmes tend to be modest, as children are not expected to achieve native-levels of development.

3.4.1 Early Foreign Language Learning Programmes

Foreign language programs in primary education differ on the number of hours they allocate to teaching, the type of curriculum they apply, the teachers involved, the way they are implemented and the outcomes expected from them (Inbar-Lourie and Shohamy, 2009; Curtain, 2009; Nikolov and Djigunovic, 2011). In terms of time devoted to foreign language learning, they range from approximately one-hour a week, especially in awareness-raising programmes with very young learners, to approximately three hours a week in most FL programmes and several hours a week in some Content and Language Integrated Learning programmes.

Johnstone (2009) refers to four generally different models that can be found in the present panorama of Early Language Learning. The first model mentioned by Johnstone uses very general topics such as colours or parts of the body as the topics of study whereas, in the second model, topics are borrowed from other curricular areas, such as animals and their habitat. These two models are widely spread within the European context and they share several characteristics: the limited amount of time devoted to language instruction per week, and the teachers' proficiency, which often

falls short of native competence. The third model has as its main objective the idea to sensitize children to languages and raising their awareness towards language learning, while the fourth model reported by this author is immersion, in which an important part of the curriculum is taught through an additional language. In the same lines, Met (1999) had already pointed out that language programmes for young learners could be placed on a continuum which ranges from language awareness programmes to those that stress content. The models on the continuum differ on the proportion of language and content to be involved in them (see Figure 3.1).

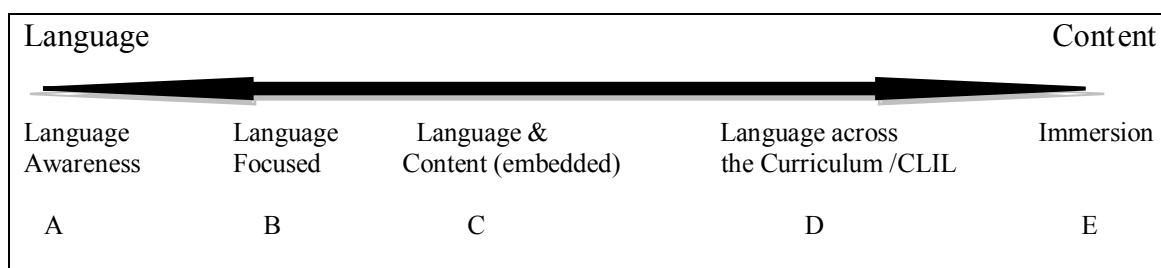


Figure 3.1 A Language-content continuum for teaching languages to young learners. Adapted from Met (1999)

On one of the ends of the continuum, language awareness programmes aim at developing familiarity with language systems, with little emphasis placed on developing language communicative competence in the target language (Martin, 2000). Further along the continuum, Language Focused and Language & Content programmes represent more traditional approaches to language learning. Their central goal is to develop the learners' language competence. Language focused programmes centre their attention on teaching lexicon and structures, and the topics used are selected from the learners' environment, for example, family, colours, parts of the body (Rixon, 2007). In Language & Content Programmes, language is embedded in content, and the topics used are derived from subjects taught in the school curriculum. This is perceived to have a positive impact on the link between the language learning experience and the

students' knowledge (Inbar-lourie and Shohamy, 2009). The next model in the continuum, Language across the Curriculum or CLIL (Content and Language Integrated Learning) rely much more on content than the previous models. In CLIL, the language is used as a means to acquire knowledge of different school subjects (see Chapter 2 for a detailed account of CLIL). The final part of the continuum (E) is referred to as immersion. In this type of programmes, in the words of Inbar-Lourie et al. (2009, p.85), 'language serves the content'. In spite of their differences, both models, Content and Language Integrated models and Immersion models are perceived as content-oriented models.

Considering the variation in programmes and approaches, providing a global scene for the teaching of English to young learners is not an easy task. The teaching of English varies considerably from programme to programme, and depends not only on the type of programme implemented but also on a number of variables that influence the outcomes of each programme.

The specific type of programme chosen may have important implications for the achievement targets although, as many authors have already stated, the overall achievements, except for the models of total immersion, tend to be modest: children are not expected to acquire native-levels of proficiency but rather a certain degree of communicative ability (Haenni Hoti, Heinsmann and Müller, 2009; Inbar-Loury et al, 2009).

Although there is no such thing as the 'typical Teaching of English to Young Learners view', several variables should be considered as common ground within the scene of young foreign language learners. Hugues (2011) refers to a number of variables that may contribute to defining the scene for teaching YLL and that may affect the amount and quality of the foreign language being learned by the children. He refers to

starting age, the level of expertise of their teachers as well as the amount and quality of teacher training, the syllabus, the amount of input provision, and the learning objectives outlined for each class. Although there is no such thing as a standard starting age, as has already been mentioned in this chapter, the general trend in many European systems has been to lower the initial age for language learning. It has long been hypothesized that young language learners are better students in terms of language learning than their older colleagues, and this belief has been used to support the early introduction of foreign language programmes, mainly English language learning programmes (for an account of the relation between age and the Critical Period Hypothesis, see Section 3.2.1). The importance of input has also been discussed in Section 3.3.1. Therefore, the next section will briefly refer to the teachers and their role in early language learning programmes.

3.4.2 The Teacher Factor in Early Language Learning

Teachers represent one of the key elements in the success of young language learner programmes: they are the main source of input for the learners, and they are an essential element for the motivation of the learners in the language classrooms. The successful introduction of English in the primary classroom depends a great deal on the training and methodological and linguistic qualifications of the teachers involved (Butler, 2004; Johnstone, 2009; Moon, 2009) as well as the way they implement programmes and deliver their lessons. In the words of Lundberg (2010), success lies in the hands of the person who performs the teaching act, puts into place the effects of language policies, interprets them, and spends hundreds and hours of schooling with his or her students.

Nikolov and Djigunovic (2011) refer to the three fundamental quality measures that teachers should meet. For them, teachers dealing with YLL should be proficient in

both the pupils' L1 and the L2 they are going to teach; they should be methodologically trained and they also need to understand the nature and principles of language learning in general and of foreign language learning in particular. Research on YLL teachers has shown that teachers often fall short of these criteria (Nikolov, 2008; Moon, 2009; Lugossy, 2007). Edelenbos, Johnstone and Kubanek (2006) suggest an important link between the teachers' ability to interact and use the language and the primary methodological approaches taken. From Vygotsky's perspective, the talk that takes place between teachers and students in the foreign language class not only provides a model of language but also information about the skills required to learn. The interaction between student and teacher is a tool for learning as it contributes to building up joint understanding between them.

In a study carried out by Nikolov and Djigunovic (2011), the authors refer to different types of teachers and point out how the nature of their training may affect the learning process of YLL. According to them, YLL teachers in most primary contexts in Europe are either generalist home room teachers or language specialists. In the case of home room teachers, these are teachers qualified to teach all or almost all the subjects in the school curriculum, including a foreign language; they are abreast of the curriculum and can thus embed the L2 in the content that children are familiar with; however, in some cases their proficiency may not be very good. Language specialists, by contrast, have been trained to teach their foreign language at this level of education; they usually have a higher level in the target language and may be methodologically more skilled in implementing age-appropriate methodology (Nikolov and Djigunovic 2011: 107). The ElliE study¹⁶ refers to another category of teacher, which, according to its authors, fully reflects the current situation in language learning in Europe: they refer to what they

¹⁶ <http://www.ellieresearch.eu/>

have called an ‘unqualified’ teacher, that is, a teacher who holds a university degree in the target language but who is neither a generalist nor a YL language specialist. In the case of Catalonia, where this study was carried out, most of the teachers involved with young learners fall into the category of generalist or specialist primary teachers. However, although the introduction of different models of Early Language Learning programmes has often been part of a long-term planned implementation scheme in some contexts, in other contexts there has been relatively little time to train teachers properly. This is the case, for example, of CLIL in primary education in Catalonia. Due to the fast trend in the implementation of CLIL programmes in primary, together with the lack of fully linguistically-competent trained primary teachers in some schools, teachers with university degrees in other subject knowledge areas but linguistically competent in English (or teachers with degrees in the target language), have been allocated in CLIL courses in primary education.

A good knowledge by teachers of the target language is the basis of success in the early stages of language learning. Primary English teachers need to have adequate knowledge of the target language as they are going to act as the main source of input for the learners. They will need to provide comprehensible input and natural exposure to the language in the classroom settings, as well as take decisions on the language to be addressed. Several studies, mostly based on classroom observations, have tried to define the role of teachers in YL classrooms. A study carried out in Hungary by Nikolov (2008), which combined classroom observations and interviews, concluded that Hungarian teachers not only lacked proficiency in the target language but also had poor methodological skills. The teachers claimed that teaching YL was very demanding and they considered that several strategies, such as the use of storytelling and games, were a waste of time. Butler (2004) concluded that early language learning teachers in Korea,

Taiwan and Japan felt their productive skills lagged behind their receptive skills. The teachers made explicit their feeling that they needed to improve their English to be able to implement their programmes in an adequate manner. In a similar context, Aline and Hosoda (2006) investigated the interaction between homeroom teachers and English native speaker teaching assistants in Japan. The latter's participation in the English classes affected the interaction in the classes in different ways. The authors observed four different patterns of behaviour among homeroom teachers: they acted as bystanders, translators, co-learners of English and co-teachers.

The issue of language proficiency in the target language, and its impact on children's development, has also been the focus of several research studies. Mihaljevic Djigunovic (2010) carried out a research project on the use of the L1 by teachers in Hungarian classes. They found that teachers perceived their use of L1 to be lower than that showed by observation data. When the authors introduced the learners' achievement into the equation, they discovered that those learners whose teachers had used the least L1 in their classes did not score the highest on receptive and productive tasks. After analysing the groups who had scored the highest, they concluded that it was not the amount of the L1 versus the L2 that made the difference, but rather the 'when' and 'why' the L1 was used. Inbar-Lourie (2010) studied the teachers' proficiency in the target language in the Israeli context. She explored the impact that the choice in the use of the L1 or the L2 in the classes had on the children's language development. Data was collected through observations and interviews. She concluded that the teachers' decisions on the use of the L1 or the L2 were based on the learners' needs. From a different perspective and in a different context, Nagy (2009) concluded that teachers working with fourth grade EFL groups felt that their practice was constrained by several

factors, including the children's aptitude and the teachers' own proficiency in English and their methodological preferences.

Research into the teaching of English to Young Learners in Norway has also contributed to the growing body of research on the role of teachers from different perspectives. Charboneau (2012) concentrated on aspects of reading instruction in EFL classes. Her conclusions offer useful insight into the teaching practices of EFL teachers at primary level in the Norwegian context. She concluded that the majority of teachers used a text-book based literacy approach and that they followed the same approach to reading as the teachers teaching in Norwegian. Drew and Pedersen (2012) carried out a qualitative study on the impact of readers' theatre in mainstream English classes. The study suggested that the use of readers' theatre has an enormous potential in language classes with learners of different abilities.

The studies carried out by Butler (2005), Moon (2009), Lugossy (2007) have provided a good insight on how teachers adapt their own practice. Butler found that teachers were really concerned about using a communicative methodology. A similar picture emerged from Moon (2009) in Vietnam. She demonstrated the influence that teachers' beliefs, course materials and language practices have on the daily classroom practices. Lugossy (2007) compared what teachers believed about communicative language teaching and how they actually implemented new ideas in their classes. She found that very few teachers changed their practices over time. In the Croatian context, Mihaljevic Djigunovic (2010) concentrated on the teachers' beliefs and how they might affect their own practice: all the teachers involved thought that children learn through games, songs and role-plays. Some considered the role of the teacher to be the leader in language learning, and none of them considered that doubling the number of hours of weekly exposure (from two to four) would make any difference to language learning.

They all thought the use of the L1 in the classroom necessary in order to ensure that the learners were following the classes. Matteoudakis, Dvorakova and Lang. (2007) investigated how pre-service teachers from the Czech Republic, Hungary and Greece applied theoretical principles they had learned about story-based language teaching in practical situations. The trainees were asked to adapt and present a story of their own choice to their peers. Results showed that the quality of work varied from country to country, and that the trainees' educational context partly influenced how they performed the task. In Sweden, Lundberg (2007) conducted a three-year project focusing on in-service teachers. The aim was to enhance the competence of language teachers of young learners in Sweden who lacked formal qualifications in teaching English. Before the training course started, Lundberg investigated several aspects: the content of typical language lessons, the use of the target language, and the pupils' and teachers' attitudes to the subject. She discovered that pupils were generally demotivated to learn English and the teachers were generally insecure and lacked confidence. Lessons were characterised by lack of variation, or tempo, boring materials, lack of planning, lack of continuity from previous lessons, insufficient challenges, lack of physical activities and lack of long-term progression.

A nationally-sponsored project in Israel, carried out by Shohamy and Inbar (2007), compared whether homeroom or EFL specialist teachers were best-suited to teach English. The findings were inconclusive about which category constituted the ideal first grade English teacher. Homeroom teachers integrated English with other subjects, whereas EFL specialists taught English in a more detached way. The EFL specialist teachers and school leaders were those who were most worried about the quality of teaching, and parents and school leaders generally preferred EFL specialists.

Evidence suggests that there is great variation in the quality and experience of the teachers who are actually implementing young learners' programmes. However, accepting the low exposure that young foreign language learners get from the environment, there seems to be a general consensus on the role of the teacher as one of the influential aspects to be taken into account because it affects not only the model of language provided, but also the type of methodology adopted. As Rixon (2000, 3-4) points out 'a teacher who lacks confidence and fluency in the language is unlikely to set up the occasions for genuine interaction. Without adequate opportunities to engage in genuine interaction with other users of the foreign language, another capacity of young learners will go to waste.' From a pedagogical point of view, teachers involved in young learners programmes need to understand the relevance of their classroom practice and the impact that it may have on their students' process of learning. The activities provided should be linked to young learners' developmental stages, they must be cognitive and linguistically challenging, and at the same time they must cater for individual differences among students.

3.4.3 Learning a Foreign Language in the School Context

The present study was carried out with students in their final two years of primary education. Under the legislation of the Catalan educational system, compulsory primary education lasts for 6 years, until the age of 12. It is provided by state, state-assisted schools and private schools and is divided into three cycles: the first cycle caters for children between the ages of 6 to 8 and comprises grades 1 and 2; the second cycle comprises grades 3 and 4, that is children between 8 and 10 years of age, and the third cycle, which includes grades 5 and 6, children aged between 9 and 12. By law, all primary schools provide English classes to children from their first year in primary education. The time allocated to English depends on the student's grade and the school,

ranging from 1 hour a week in the early stages to three to 4 hours per week in the final stages in primary.

The main aim of language programmes in primary education is to develop the learners' communicative competence. Learners study a foreign language to be able to use and practise it in real situations. Recent theorizing in the field of young learners with a focus on classroom practice (Cameron, 2001; Halliwell, 1992; Moon, 2000; Phillips, 1993; Pinter, 2006; Rixon, 1999; and Slattery and Willis, 2001) has emphasized that the primary focus when teaching young learners, especially in the first years, should be on meaning rather than on form mainly because of maturational constraints that do not allow young learners to take an analytical approach to language learning.

Numerous current EFL approaches for teaching foreign languages to young learners are related to theme-based courses of study. Theme-based teaching is an approach in which the courses are structured around certain themes or topics that are interesting for the learners. In such a course, different skills are incorporated into the theme which acts as a 'connecting' thread for pupils and teachers (Cameron, 2001). Theme-based instruction differs from traditional language instruction in that the language structures and vocabulary items to be covered in a syllabus are determined by the theme or topic. The emphasis is on exposing the students to the target language in a 'highly contextualized' environment by using different themes as the content of language learning (Wesche and Skehan, 2002). The rationale for this thematic approach is to 'avoid fragmentation and the use of unconnected skill exercises' (Berry and Mindes, 1993: p.6) as it allows the integration of a variety of activities around meaningful content. It is thought that teaching which is integrated around a theme suits the way children naturally learn languages. Therefore, theme-based teaching solves the

problem of what to teach in the foreign language classes, especially in young learners' classrooms, where a focus on language itself would not be appropriate. Such an approach provides learners in a foreign language class with motivating and meaningful uses of the language (Cameron, 2001).

The teaching of English at school level has always followed the trends that have affected foreign language learning in general. Over the last few decades, it has been common ground in applied linguistics to divide foreign language learning into the 'four skills': listening, reading, speaking and writing. Children who start learning a foreign language in a school setting at an early age do not encounter all four skills from the very beginning. It is through the spoken language that they first encounter, understand, practise and begin to learn the language (Cameron, 2001). The spoken form acts as the most important source for language learning. New language is introduced, practised and automatised orally. Young foreign language learners start learning the language with plenty of listening activities to encourage the students to develop this skill. Speaking fluently in a foreign language is a long process that requires a lot of practice. After the learners have been exposed to listening in the foreign language class, they soon want to communicate in the target language, but their capacity to communicate is still very limited. In a foreign language class with young learners, the teachers build up blocks that allow children to move from listening to speaking. In this way, the teachers provide 'unanalysed chunks' that allow children to participate in interactions. Chunks are picked up effortlessly by the children and they provide their basis for communication. Because the oral proficiency of the children is low, the reading and writing skills are introduced later in the traditional EFL class. As reading is a holistic process that involves many skills, it is first introduced at word level, using familiar language forms, and then it is gradually build up towards sentence and text level. Reading and writing are often taught

in parallel. Writing also begins at word level and, as the children become better writers in their first language, they write at sentence and text level (Pinter, 2006). Grammar and vocabulary are learnt in a holistic way in the first stages, until the children are old enough to understand and analyse the chunks of language provided by teacher or the textbook. This growing interest and capacity to understand language develops around puberty (see Section 3.1 in this chapter). It is then that children start breaking up the chunks of language, previously learnt as wholes, into their elements. Thus they start being able to recombine the parts to convey new meanings, separating vocabulary from grammar forms. Towards puberty, the teaching of grammar can take a more explicit analytical form. However, learning grammar is not a lineal process, and children do not learn patterns straightaway. They often use intermediate forms before they are actually ready to use forms that conform to the target language rules (Pinter, 2006). It is well-known in the field of SLA that children from different language backgrounds make similar mistakes when learning grammar until they are ready to use correct target language forms. Pieneman (1998) investigated the development stages of foreign language learners and stated that all learners of a foreign language go through the same developmental stages in exactly the same sequence. They attributed variation to individual and contextual factors. These authors contextualize progression as a series of steps on a ladder. Progressing through a foreign language is a complex and non-linear process with setbacks and multiple regression points (Mitchell, 2003). Peltzer-Karf and Zangl (1997) argue that children's utterances develop from short phrases to longer ones. During this process, grammar control improves and regresses until it stabilizes.

Most school programmes regularly use a textbook as the basis for learning. Textbooks provide organized units of work and have been designed following the guidelines and demands of the curriculum for foreign language learning. The activities

in each unit provide practice in the four skills, and language structures and vocabulary are in most cases organised around thematic- units.

As has already been explained in this dissertation, there has been in recent years a trend towards the implementation of content-based approaches, mainly CLIL programmes in primary school in the belief that this type of programme will help students increase their language competence in the target language. Therefore, a discussion on the topic of content -based programmes as well as the issue of language learning through content is in order.

3.4.4 Learning Languages through Content. Content-based Models

Content-based approaches come in different shapes and sizes. For Met (1998), content-based instructional settings range from language-driven content programmes to content-driven language programmes. Language driven-content programmes include language classes based on thematic units or with frequent use of content for language practice. As discussed in Chapter 2, a trend towards the use of content-driven approaches has spread across the educational systems in Europe. In Catalonia, where this study was carried out, schools are incorporating content-based approaches, mainly CLIL programmes, together with traditional foreign language ones, with the idea of improving the learners' foreign language competence. (For a detailed description of the situation in Catalonia, see Chapter 2, Section 2.4).

A range of theoretical perspectives support the use of content-based approaches to second language learning and teaching (Lyster, 2007; Echevarria and Graves, 1998). Many researchers have drawn on the principles of socio-cognitive theory to provide support and understanding for both language and content development in classroom settings (Bange, 2005; De Bot, 1996; Hulstjin, 1990; Lyster, 2007). Although there are many factors that affect language acquisition, practice in learning or skill acquisition

has always been considered highly relevant in language learning. Anderson (1985) described skill acquisition as a gradual change in knowledge from declarative to procedural mental representations. As far as language is concerned, declarative knowledge refers to knowledge of the language which is stored in long-term memory, whereas procedural knowledge involves the processing of language, in other words, knowledge on how to do things. It is seen as the learners' ability to understand and produce language through access to representations stored in memory (Lyster, 2007; (DeKeyser, 1998). The transition from declarative knowledge to procedural knowledge requires a transition from controlled processing to more automatic processes which result from repeated practice and feedback in transforming declarative representations into production rules, thus linking form and meaning. DeKeyser (2007: 288) argues that, 'a high degree of automaticity is the ultimate goal for most learners, because of its impact on the quality of the linguistic output and because of how it frees up resources for processing content instead of language'.

Another cognitive perspective is offered by Byalstok (1994), who asserts that learning does not proceed from explicit representations of declarative knowledge, but rather from representations of implicitly acquired and unanalyzed knowledge. Considering that young learners use a formulaic system to language learning in the initial stages and that their knowledge of the language is largely implicit and made up of unanalyzed forms, this author considers that they may benefit from instruction designed to increase analysis of the representations because this will lead to an increase in accessibility to knowledge. In contrast, 'knowledge of the language in a less unanalyzed form will limit the learner in terms of what can be achieved' (Bialystok, 1994:160). According to Lyster (2007: 20), 'skill acquisition theory may be useful in understanding interlanguage development but also apparent plateau effects in content-based classrooms'.

He refers to the lack of language instructional interventions in such contexts and also to the lack of effective feedback as two important elements for learning a language.

Johnson (1996) pointed out that naturalistic approaches, such as content-based ones, are designed to directly develop the procedural encodings of the target language. For him 'proceduralized' forms that come into the interlanguage 'quickly become highly automatized and are impermeable to change (p.99). 'The objective of teachers in content-based programmes is then double-fold: to help students develop declarative knowledge from procedural knowledge that they have acquired in a more or less naturalistic way and push students to develop new representations in the target language that compete with easily accessible interlanguage forms' (Lyster, 2007: 20). Second language classrooms have traditionally aimed at developing declarative knowledge and have not provided the learners with enough opportunities to proceduralize their declarative knowledge. Content-based approaches may encourage the students to use the target language not only as a communicative tool per se, but also as a tool that helps them to develop their cognitive abilities, as well as acquire new knowledge and be able to actively participate in the construction of knowledge. In order to maximise the potential of such programs, Lyster (2007) suggests considering Bruner's notion of scaffolding as a way to help the students to express what they would not be able to express by themselves within a meaningful context. In the words of Wood et al. (1976: 60), scaffolding appears to be encouraging as 'it enables learners to solve a problem, carry out a task or achieve a goal which would be beyond their unassisted efforts'. Lightbrown and Spada (2006), however, point out that building-up fluency changes in language is sometimes difficult to explain just through practice (Lightbrown and Spada, 2006). According to these authors, the restructuring of the learners' knowledge seems to account for what they call, 'bursts of progress': learners seem to put their knowledge together although apparently they have not received new

instruction or relevant exposure to the language. Lightbrown also refers to the fact that information is best retrieved in similar situations to those in which it was acquired. Knowledge that is acquired on a particular type of activity may easily be accessed on tests that resemble the learning activities. If the learners are occupied with a focus on meaning in communicative activities, retrieving specific language features may be more difficult (Gatbonton and Segalowitz, 2005).

Traditional language teaching in contexts where the target language is a foreign language which is not present in the wider environment quite often takes place in a classroom setting. Students receive limited input as learning only takes place in the classrooms and the classes are part of programmes that provide between one and four hours a week of contact with the language. The main focus is on primarily on language form and most of the materials that are used entail minimal discourse (textbooks, grammar workbooks, word lists). Muñoz (2007) argues that the input the learners receive is not always authentic as the language is treated as an object to be analysed and on many occasions memorized (e.g. vocabulary) rather than the medium of instruction. She goes on to say that the input the learner's receive is functionally restricted as it is either limited to that provided by textbooks or is related to the functions of the classroom (e.g. the teacher's instructions, the correction of exercises). This type of input is neither meaningful nor challenging to students, who are often not motivated because the input is neither real nor communicative for them. The linguistic production of the learners is, therefore, limited and it does not require deep processing.

Content-based programmes are perceived as having great potential not only for developing higher levels of language competence but also because they create the ideal conditions for cognitive development, given the optimal instruction practice that nurtures the relationship between content and language. Research carried out in the Canadian

context has demonstrated that students in immersion settings develop much higher levels of second language proficiency than students enrolled in courses who study the target language as a subject. In comparison to non-immersion students, they develop much higher levels in comprehension skills (listening and reading comprehension), as well as higher levels of fluency and confidence in using the second language. Their productive skills, however, fall short of native-like competence in terms of grammatical accuracy, lexical variety and sociolinguistic appropriateness (Genessee, 1992). As has already been pointed out, although the Canadian model is not comparable to the situation in Europe, and even less so to the CLIL models applied in the Catalan education system, it has provided a good inside into what happens when the students are flooded with real authentic input. It has also provided fundamental considerations on the type of input provided by the teachers as well as the teaching approach taken in terms of making the integration of content and language successful from two different perspectives, those of knowledge and language acquisition. However, as some authors have pointed out, content teaching and good language teaching do not necessarily come together. From class observations carried out in immersion settings, Swain (1998) concluded that some content teachers paid no or little attention to students' target language use and that content teaching did not invite much student production, and the production that did occur was quite often restricted to the functions of the language that appeared during the sessions. Certain uses of the language may not have been practised due to contextual limitations of place, topics and interlocutors. Along the same lines, Swain and Carroll (1987) had already observed that teachers tended to avoid language issues in content-based lessons, and they concluded that teachers in such settings relied to a great extent on 'incidental' language learning which has generally been defined as learning without the intent to learn. Incidental attention drawn to language in subject-matter instruction may be

insufficient in developing the learners' competence because of the lack of intentional and systematic focuses on language. In the words of Swain (1998) incidental language learning falls short of manipulating and complementing content teaching to maximize second language learning. Content-based instruction provides a considerable amount of exposure to contextualized language use and promotes vocabulary oriented learning. However, it does not ensure the learning of important morphosyntactic features of the target language while it also falls short in terms of encouraging the processing of form.

From the perspective of second language learners in minimal input situations such as the one dealt with in this dissertation, content-based approaches call for the need to adopt a more systematic, intentional instructional approach to language teaching in order to ensure the development of language growth. For Lyster (2007), content-based classrooms provide good conditions for teachers to effectively implement form-focused instruction. He argues that a systematic focus on language aspects which might not be noticed through exposure to content alone may contribute to create a 'discourse rich instructional setting' (p.58) which will contribute to the development of the learners' interlanguage.

CLIL programmes are set up in most educational contexts with a dual objective: the learning of content and at the same time the learning of the target language. These types of programmes provide real, relevant and motivating input beyond the limits of the input provided by traditional language classes. However, considering the experience of immersion settings in Canada, students, especially beginners who focusing exclusively on content in limited input situations, may not be able to notice salient characteristics of the target language. The limitations shown by implicit instruction in contexts flooded with input (DeKeyser, 2007), recommend considering the need to combine meaning and form-oriented instruction, especially in situations of limited amounts of input (Pérez-Vidal,

2007). Content teaching needs to be manipulated and complemented in ways that maximise second language learning (Skehan, 1998). CLIL classes seem to be a suitable setting to provide opportunities to integrate focus on meaning and focus on form (Muñoz, 2007). In a focus on form approach, meaning is drawn towards the language forms in order to develop linguistic awareness, which may result in intake for the learners. Although the need to focus on form is not a characteristic of CLIL teaching, it is, in the words of Muñoz, 'a desirable one of all communicative lessons including CLIL' (p.23). She goes on to say that the need to focus on form may arise from the need to complete a task or from the teachers' planning anticipating the need for several forms to deal with a particular theme. It may also arise from the knowledge of expected difficulties in a particular form. For learners to be able to transform input into intake and incorporate it into their interlanguage system, they need to be offered the possibility of noticing. Young learners in content-based classes may benefit from the inclusion of age-appropriate noticing activities that will enable them to draw on their linguistic knowledge for the kind of implicit analysis of naturalistic input that they need to engage in to drive their interlanguage development forward (Skehan, 1998). Because they rely heavily on the use of formulaic chunks in their early production, teachers may be able to exploit a progressive analysis of the formulae as a means of encouraging the development of the rule-based system. A content-based approach ensures extensive exposure to comprehensible input, however, as it has been discussed, in limited input situations, it does not guarantee a great deal of target language learning unless salient target features are made explicit by a focus on form approach.

CHAPTER 4 Method

4.1 Design of the study

As has already been explained in the introduction, this study is aimed at determining the effect of CLIL on the language competence of Young Learners of English, and to what extent their listening, reading and writing abilities were affected by exposing them to Natural Sciences (School A) and Arts & Crafts in English (Schools B and C). The results obtained by 5th and 6th primary graders exposed only to EFL classes, the Control Group (CG), were compared to those obtained by 5th and 6th graders exposed to exactly the same number of hours of English (EFL and CLIL hours combined), to determine the students' progress in the target language at different times (T0, T1, T2, T3) and time periods (T0-T1, T1-T2, T2-T3, T0-T3) over two school years. For the results to be reliable, several variables were taken into account in the statistical data analysis, including the number of hours of school exposure to English up to 5th grade and the initial level of proficiency in English (High or Low achievers).

4.2 School Contexts and Participants

Three different schools participated in the study. These schools offer tuition from early years up to and including compulsory secondary education, and the three of them are state assisted private co-educational schools¹⁷ located in three different towns in Catalonia. This type of school was chosen for the project because, although the present study focuses on 5th and 6th grades of primary education, these schools offer education from early years up to and including secondary education and they also

¹⁷ In the Catalan educational system, state-assisted private schools are privately-run schools which receive funding from the Government, and most of them have a religious foundation.

showed an interest in carrying out the study at the secondary level. Equally important was the fact that the CLIL teachers met the requirements for the project and accepted the conditions for data collection: they agreed to start CLIL classes one year after data collection had started for the Control Group (see section 4.3 for data collection details). In this way, we could have a control group and a CLIL group within the same school.

State schools and state-assisted schools in Catalonia offer the same study programme content, as they are governed by the same education law. Following the guidelines of the Catalan Educational Curriculum, the sample informants involved in this study were being instructed in Catalan in all the curricular areas except for two subjects, Spanish and English. Spanish was taught for 3 hours per week and the number of hours of English as a foreign language varied in the three schools: 3.5 hours per week in School A and 3 hours per week in schools B and C. For a more detailed account of the hours of English exposure, see section 4.3.1. In order to participate in this project, the head teachers of the schools agreed to teach one hour a week of a curricular subject (Natural Science or Arts & Crafts) in English to 5th graders for two consecutive years, thus providing exposure to CLIL for the experimental group.

The participants were primary school learners, N= 202, enrolled in their final two years of primary education (5th and 6th grades), whose ages ranged between 9 and 12 during the course of the study. In each school, the participants were divided into two different groups: the Control Group and the experimental CLIL Group. The CLIL group, which comprised the fifth and sixth grades in the school for the years 2011-2013, started one year after the Control Group which comprised the 5th and 6th grades for the years 2010-2012. (See Section 4.3 for data collection details)

SCHOOLS	Number of Participants	
	Control group	CLIL group
SCHOOL A	55	52
SCHOOL B	25	19
SCHOOL C	26	25
TOTAL	106	96

Table 4.1 Number of participants per group in each school

Although the informants were part of reasonably homogeneous groups with regard to age and gender, each individual group had distinguishing characteristics. In the next section the characteristics of each group and school participating in the study will be explained.

4.2.1 School A

School A is a state-assisted co-educational school which offers an educational programme from early years to secondary education. The majority of the children of its school population, approximately 700 students in the school roll is of Catalan origin; however, a few students per class belong to different ethnic groups. The school is located in a middle-sized town north of Barcelona. Each year is divided into two classes: in early years and in primary there are approximately 25 students per class. In secondary, this number rises to 30. The CLIL subject chosen by this school was Natural Sciences.

4.2.1.1 Control Group (School A)

The pupils in the control group numbered 55. The gender ratio was 32 females to 22 males and they were distributed in two classes. 85% of them were Catalan in origin

and the rest (15%) had different origins, mainly from South America, Morocco, Romania and Ghana (see Fig 4.1).

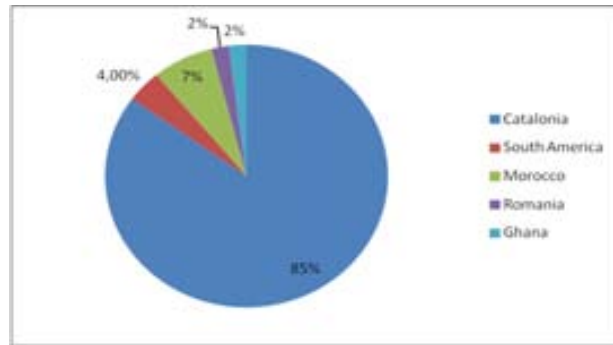


Figure 4.1 Country of origin. Control Group A

68% of the students in this group spoke Catalan as their first language, while the languages of the rest were made up as follows: 13% spoke Spanish and the rest Tamazight¹⁸ (11%), Twi¹⁹ and Romanian (4% each) (see Fig 4.2).

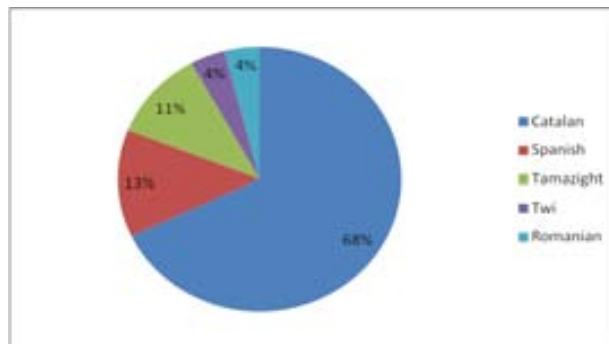


Fig 4.2 Languages spoken at home. Control Group A

In spite of the diversity of origin and first languages of the sample of students, all of them had attended classes in the school since they were three years of age and had started learning English at the same time as well. More than half the students in this group (64%) regularly attended extra-curricular English classes.

¹⁸ One of the languages spoken in Morocco

¹⁹ One of the languages spoken in Nigeria

New-comers to the school were not taken into consideration for the study as they did not meet the requirements of the other participants. This and the above mentioned fact that they had been together since they were 3, was also applied to the other two schools in the study in order to keep the initial requirements for the participants unchanged.

4.2.1.2 CLIL group (School A)

The pupils in the CLIL group in school A numbered 52 (22 females and 30 males) and, as in the case of the control group, were distributed in two classes (of 26 students each). 92% of the students were local Catalan students, while rest (8%) were of different origins, mainly from Morocco, Romania and Ghana (Figure 4.3). 81 % of the students in the control group spoke Catalan as their first language, while the languages of the rest were made up as follows: 15% spoke Tamazight, 2% Twi and 2% Romanian

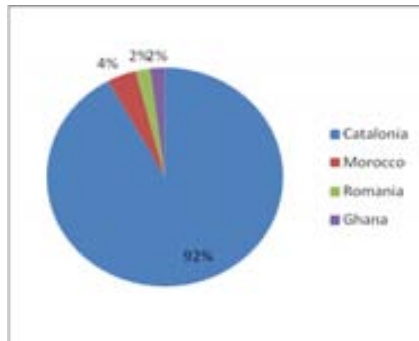


Figure 4.3 Country of origin.
CLIL Group A

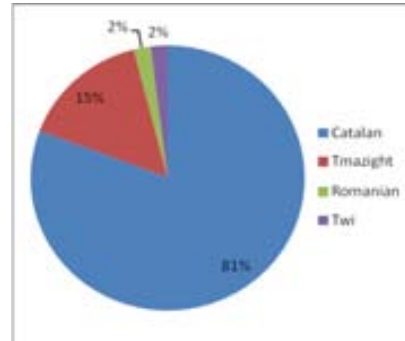


Figure 4.4 Languages spoken at home.
CLIL Group A

As in the case of the Control Group, all the students had attended classes in the school since they were three years of age, and none of them had learnt English abroad. At the time of the study, 73% of the students in this group regularly attended extra-curricular English classes.

4.2.2 School B

School B is also a co-educational state-assisted private school located in a small town in the north of the province of Barcelona. The school has 325 students in early years, primary and secondary education of a variety of nationalities. It is a smaller school than school A and it has only one class per year. The CLIL subject chosen by the school was Arts & Crafts.

At the time of the project, this school was participating in a PILE project (modality A) organized and financed by the Catalan Department of Education. PILE programs are included in the Pla Integrat de Llengües Estrangeres²⁰ (PILE, in English *Integrated Foreign Language Learning Scheme*). Schools and teachers involved in a PILE project receive training schemes as well as financial support designed to improve the quality of Foreign Language (FL) teaching and learning. CLIL Projects, such as the one received by this school, are included as an independent modality within the PILE general scheme with a double objective: to extend the number of hours of exposure to the first foreign language and the integration of FL language and curricular contents.

4.2.2.1. Control group (School B)

The Control Group in school B (26 students: 14 females and 12 males) was rather a homogenous group in terms of country of origin and the variety of languages spoken at home: 92% of the students were local Catalan students, and the rest (8%) came from Morocco. 76% spoke Catalan regularly with their families, 16% spoke Spanish at home and the rest, 8%, spoke Tamazight (Figures 4.5 and 4.6). As in school

²⁰ <http://phobos.xtec.cat/pluriling/pele.html>

A, all the students in this group had been together at school since they were three years old and none of them had repeated an academic year.

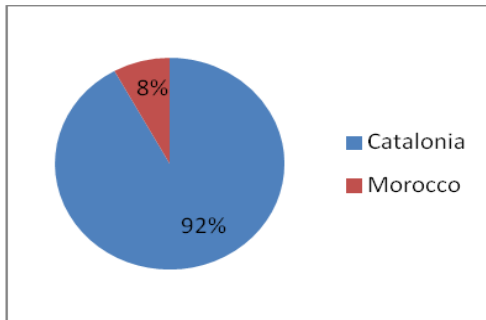


Figure 4.5 Country of origin. Control Group B

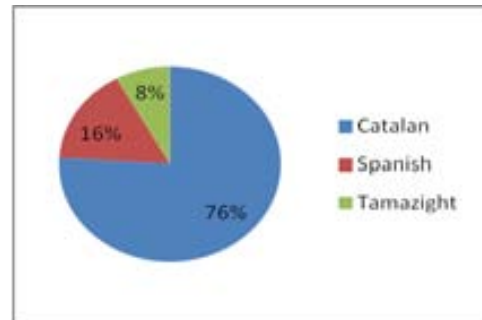


Figure 4.6 Languages spoken at home. Control Group B

More than half the students in this group (68%) regularly attended extra-curricular English classes and none of the pupils had spent time in an English speaking country.

4.2.2.2 CLIL group (School B)

The CLIL group in school B was slightly smaller than the other groups: 19 students (9 males and 10 females) made up the group. 94% of them were local Catalan students, and the rest (6%) came from Argentina (see Fig 4.7). 70% of the students had Catalan as their first language, 18% spoke Spanish at home and 12% claimed to speak both Catalan and Spanish.

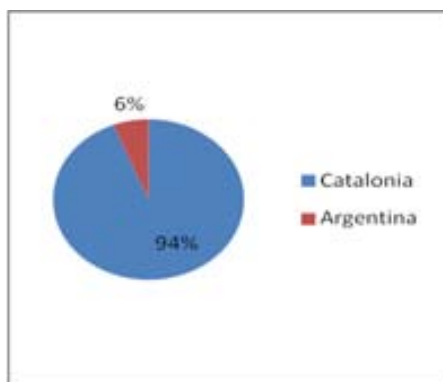


Figure 4.7: Country of origin. CLIL Group B

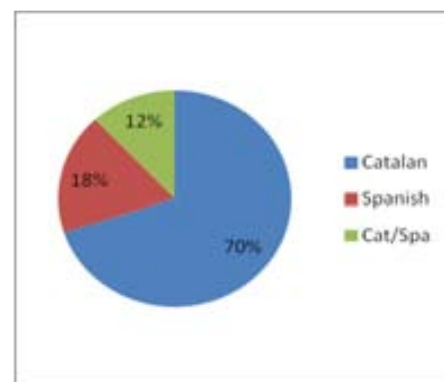


Figure 4.8 Languages spoken at home. CLIL Group B

All the students in this group had been together in the same school since they were three years of age. At the time of the study, 65% of the informants in this group regularly attended extra-curricular English classes. None of the students had spent time in an English Speaking country and they had all been to school since they were three years of age.

4.2.3. School C

School C is a co-educational, primary state-assisted school. It is situated in small town in a rural area in the province of Lleida. The school roll is around 300 students and the school offers tuition to students from 4 months to 16 years of age. There is only one group per year and the CLIL subject chosen was Arts & Crafts.

4.2.3.1 Control group (School C)

80% of the students in the control group are local Catalan students, and the rest come from The Ukraine (8%), Romania (8%) and Russia (4%) (Figure 4.9). 56% speak of them regularly speak Catalan with their families, while 16% speak Spanish at home, 8% speak both Catalan and Spanish, and the rest, 20% speak other languages (Ukrainian, Romanian or Russian (Figure 4.10). However, all the students had been in the school since they were three years of age. None of the students had repeated an academic year and none of them had spent time abroad. 68% of the students attended extra-curricular classes.

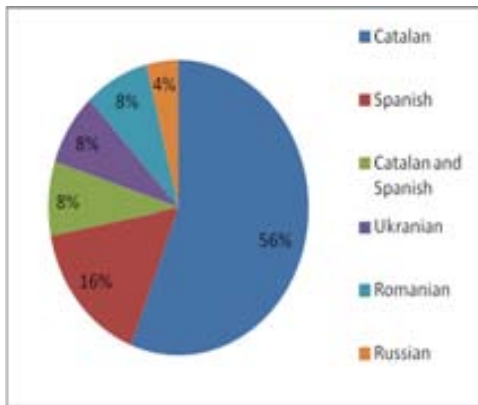


Figure 4.9 Country of origin. Control Group C

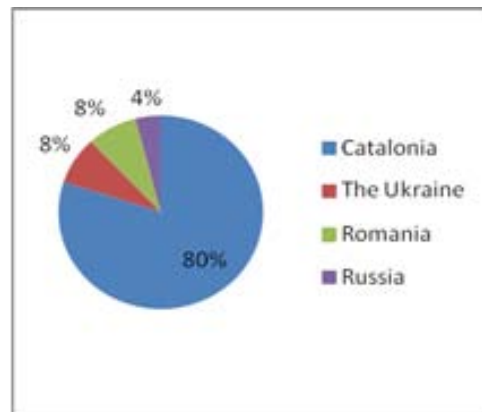


Figure 4.10 Languages spoken at home. Control Group C

4.2.3.2 CLIL group (School C)

Twenty-six students made up the CLIL Group in school C, 16 girls and 10 boys. 66% of them were local Catalan students. Romanian students accounted for 16% of the students. Ukrainian, Bulgarian, Colombian and Czech students accounted for 20% of the total (see Figure 4.11).

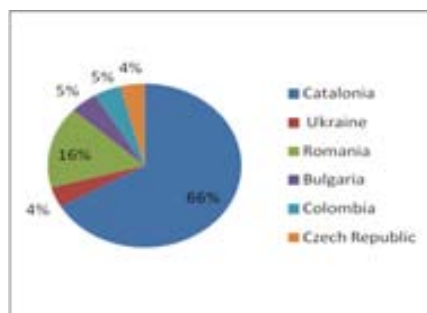


Figure 4.11 Country of origin. CLIL Group C

The diversity on the students' origins can also be seen in the number of different mother tongues spoken in this group. 60% of the children had Catalan as their first language, and 12% spoke Spanish at home. 16% speak Romanian as their mother tongue and the rest speak Bulgarian, Czech and Ukrainian (Figure 4.12). As in the case

of the other groups, the participants in the study had been together since they were three years of age and they had not been abroad to an English Speaking country. 70% of them attended extra-curricular English classes.

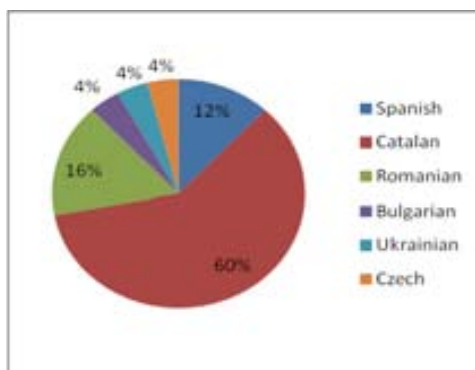


Figure 4.12 Languages spoken at home. CLIL group C

4.2.4 Summary of Participants

To summarize, there was great similarity between all the students who participated in the study. The great majority of pupils were local students from the school catchment areas, although each school had, to a certain extent, students from other countries but who had been educated in Catalan, the academic language of tuition in the schools in Catalonia, since they were three years old. Students who did not meet this requirement, were excluded from the study. In this way, it was easier to calculate and account for the previous hours of exposure to English that every group had because this was one of the variables considered for the analysis. None of the pupils had practised their English in an English-speaking country, and none of them had repeated an academic year. For all the participants English was a foreign language, to which they were devoting a limited number of hours per week in the schools. The table below provides a summary of the number of students per group and their country of origin.

SCHOOLS	GROUPS	NUMBER OF PARTICIPANTS	COUNTRY OF ORIGIN	%
School A	Control Group	56	Catalonia	85%
			Morocco	7%
			South America	4%
			Romania	2%
			Ghana	2%
	CLIL Group	52	Catalonia	92%
			Morocco	4%
			Romania	2%
Ghana			2%	
School B	Control group	26	Catalonia	92%
			Morocco	8%
	CLIL Group	19	Catalonia	94%
			Argentina	6%
School C	Control Group	25	Catalonia	80%
			The Ukraine	8%
			Romania	8%
			Russia	4%
	CLIL Group	26	Catalonia	66%
			Romania	16%
			Ukraine	4%
			Colombia	5%
			Bulgaria	5%
			Czech Republic	4%

Table 4.2 Summary of participants by country of origin

As already mentioned, although all the students had been educated through the medium of the Catalan language, not all of them had Catalan as their mother tongue. As we can see in the summary in Table 4.3, a high percentage in each group, spoke Catalan at home. However, some students, due to their different origins, had a variety of mother tongues. This variety, nevertheless, was similar in the schools, in both the control and the CLIL groups.

SCHOOLS	LANGUAGES SPOKEN AT HOME		%
A	Control Group	Catalan	68 %
		Spanish	13%
		Tamazight	11%
		Romanian	4%
		Twi	4%
	CLIL Group	Catalan	81%
		Tamazight	15%
		Romanian	2%
B	Control group	Catalan	76%
		Spanish	16%

		Tamazight	8%
	CLIL Group	Catalan	70%
		Spanish	18%
		Catalan/Spanish	12%
C	Control Group	Catalan	56%
		Spanish	16%
		Catalan/Spanish	8%
		Ukrainian	8%
		Romanian	8%
		Russian	4%
	CLIL Group	Catalan	60%
		Romanian	16%
		Spanish	12%
		Ukrainian	4%
		Bulgarian	4%
		Czech	4%

Table 4.3 Summary of the languages spoken at home

The composition of the groups involved in this study reflects the current sociolinguistic situation of the school population in areas that, for economic reasons, have experienced waves of immigration from different parts of the world (Simó and Telford, 2011).

4.2.5 CLIL Teachers

As of today, CLIL instruction in Catalonia has not yet been completely integrated into mainstream education. Yet the schools that want to implement CLIL have two options: they can apply to participate in innovation projects that the Catalan government has launched (known collectively as PILE) or they can implement CLIL on their own initiative using their existing resources. As CLIL is still a relatively new approach for primary schools and despite the fact that the Department of Education and other institutions provide in-service training schemes for teachers involved in CLIL PILE projects, not all the teachers in such projects receive the same training and have the same profile. The rapid expansion of CLIL programmes has forced schools to recruit the best possible candidates among those teachers who show a good command of the

language, together with the best possible methodological profile for the teaching of curricular subjects in English.

The role of the foreign language teacher is one of the factors that affects programme outcomes, especially in low-exposure school programmes (Moon, 2009). Several authors have highlighted the important link between primary school methodological approaches and the teachers' ability to interact and use the languages well as the link between primary teachers' language proficiency and the children's target language achievements (Edelenbos et al., 2006; Blondin et al., 1998). Teachers in a CLIL context need appropriate training, both methodological and linguistic, but they also need to be able to reflect it in their teaching (Halbach, 2009; Fernández and Halbach, 2011). Traditionally, in an EFL context, the English specialist covers the basic communication skills of the learners (BICS); however, the teaching of curricular subjects in English inevitably requires the use of more cognitively-challenging academic language (CALP) in order to explain the contents of the subject the teacher is giving. In a CLIL context, therefore, the two areas of discourse merge into one figure, the CLIL teacher, who must integrate two roles in one (Dafouz, 2011).

For the present study, the minimum required level of English of the teachers was either First Certificate in English from ESOL Cambridge University or Level 4 (Independent User) established by the European Framework of Reference for Languages (Council of Europe, 2001). As for the general academic and teaching background, as well as the training of the teachers involved in the project, there was great variation. Each of the teachers will be referred to individually.

The CLIL teacher in School A was a 32-year-old female teacher. She held a degree in History and was also a qualified general homeroom primary teacher who specialised in the teaching of English as a foreign language to young learners. At the

time of the study, she had 6 years' experience as an English teacher and she was teaching English in early years. She was also one of the English teachers in primary in her school. She said that she regularly had contact with English in her private life and apart from English, she had some knowledge of French. Her self-reported level in French was B1.

Prior to the beginning of the project, she had attended several general methodology in-service training courses (a total of 80 hours) offered by the Catalan Department of Education. The school where she was working offered her the possibility of participating in the present study. She voluntarily accepted and attended a 20-hour CLIL training course in order to be trained as a CLIL teacher. Apart from this course, she also attended the training and support sessions offered by the CLILSLA Project staff which took place a few months before the beginning of the project. During the project, she was assessed at all times by a member of the group who gave her advice on methodological aspects of her lesson plans. The CLIL subject she was teaching during the study was Natural and Social Sciences.

As for School B, the CLIL teacher was a 29-year-old female primary and secondary teacher. She had a degree in Environmental Sciences and she had more than two years' experience in the teaching of English as a foreign language in primary and secondary education. As a student she had been involved in an Erasmus programme for 6 months and as a teacher, she was involved in a Comenius project with several European schools. She said that she had daily contact with English in her private life. In her biodata questionnaire, she pointed out that she also speaks French and Polish at B1 level.

She had little experience as an English teacher, but she had attended two in-service methodology courses offered by the Department of Education. As her school

had been granted a PILE project, which involved several hours of CLIL teacher training, she was able to attend a one-year methodological course on CLIL before the CLIL project started. The CLIL subject she was teaching was Arts & crafts. As in the case of the teacher in school A, she also attended the training sessions offered by the CLILSLA Project staff. She was chosen by the school head teacher, with her agreement, as the CLIL teacher because of her competence and fluency in the language. Before and during the project, she was assessed methodologically by the class tutor and arts teacher, a homeroom teacher with 25 years teaching experience in several areas including art. They planned the lessons and designed the course materials together because, of all the art classes during the week, only one hour a week was given in English and the rest, in Catalan. They were both following an Arts & Crafts CLIL book as a reference book to prepare their classes.

The profile of the CLIL teacher in School C was slightly different from the other two. The teacher was a 48-year-old homeroom primary teacher as well as a qualified English teacher. She had 27 years' experience in primary education. She had two different certificates in English: First Certificate in English from Cambridge ESOL and level B2 issued by the *Escola Oficial d'Idiomes*²¹. Apart from her English qualifications, her level of French was approximately B1. As with the other two teachers involved in this study, she usually attends in-service training courses in English. When her school decided to implement a CLIL project teaching Arts & Crafts in English, she volunteered to be the CLIL teacher. Up to that moment, she had had no previous experience in CLIL, and she taught herself by reading methodology books on the new approach. She participated in the CLIL training sessions that the

²¹ EOI (in Catalan, Escola Oficial d'Idiomes) the Official Language School in Catalonia which depends on the regional Government.

CILSLA project staff offered before the beginning of the study. She was guided and assessed during the project by a member of the group in terms of methodology and her classroom practices, but she decided herself on the choice of topics and projects for the class. In the case of School C she was both the EFL and the CLIL teacher of the groups, as she also taught English as a subject in the final stage of primary education.

The next sections will deal with data collection times and the instruments used for collecting data, as well as the measures used for data analysis.

4. 3. Data Collection times and Instruments

4.3.1 Data Collection Times

Several studies have so far compared CLIL and non-CLIL students in many different aspects of language acquisition (Jimenez et al, 2009; Huttner and Rieder-Bünemann, 2007; Ackerl, 2007; Lasagabaster, 2008; Ruiz de Zarobe and Jiménez-Catalán, 2009; Olaizola and García Mayo, 2009; Naves and Victori, 2010; Lorenzo et al., 2010) among many others (see Chapter 2). However, many of the participants in these studies were students in secondary education and very few researchers have compared students in similar contexts to the ones in the present study (Serra, 2007; Agustín Llach , 2009; Alba, 2009; Vallbona, 2009, 2011; Bret, 2010). Most importantly, longitudinal studies in primary education in low-input situations are rare. Some voices (Bruton, 2011, 2013; Pérez-Cañado, 2012) have already been raised about the fact that not many of the studies published so far have been rigorous enough in the control of the variables that may affect the final results: the number of hours of exposure to the language and the number of hours of previous exposure to the target language at the time of data collection, the different proficiency levels of the participants within the groups as well as several methodological aspects.

As mentioned in the design of the study, each school had a Control Group and a CLIL group. Data was elicited from both of them for two consecutive years. The students were enrolled in their two final grades of primary education, 5th and 6th grades, in the three schools. At the time of the study, the Control Group in each school was being exposed to the obligatory curricular hours of English as a Foreign Language (EFL). The students in School A were attending 3.5 hours per week of English lessons, and those in schools B and C were exposed to English lessons 3 hours a week. At the end of the two years, the participants in school A had received 245 h of English, 122.5 per academic year, and the students in schools B and C had received a total of 210 h, 105 per year . The CLIL group was also receiving the mandatory EFL hours plus one hour a week of a curricular subject in English: Science for one of the schools (School A) and Arts & Crafts for the other two (Schools B and C). Therefore, at the end of the two years, the students in school A had received a total of 315 h of English classes (122.5h of EFL per year plus a total of 70 hours of CLIL) and those in schools B and C had received a total of 280 hours (210 h of EFL plus 70 hours of CLIL) (see Table 4.4 for Number of hours of exposure to English).

	CONTROL GROUP			CLIL GROUP				Hrs at the end of study
	EFL hrs per week	Hrs per academic year*	Hrs at the end of study	EFL hours per week	CLIL hours per week	Total number English hours/ week	Hrs per academic year* ²²	
SCHOOL A	3.5	122.5	245	3.5	1	4.5	157.5	315
SCHOOL B	3	105	210	3	1	4	140	280

²² Each academic year has 35 weeks

SCHOOL C	3	105	210	3	1	4	140	280
-------------	---	-----	-----	---	---	---	-----	-----

Table 4.4 Number of hours of curricular exposure for the Control and CLIL groups

This study aims at determining the development of proficiency of CLIL and non-CLIL students trying to be as rigorous as possible in the control of the variables. In order to obtain reliable results, it was decided that the results obtained by 5th primary graders exposed only to EFL classes (Control Group) would be compared to those obtained by 5th graders within the same school, exposed to exactly the same number of hours of English (EFL and CLIL hours combined) to determine the students' achievement and progress in the target language at different time times (T0, T1, T2, T3) and at different time periods (T0-T1, T1-T2, T2-T3, T0-T3) over two school years. As this was a two-year longitudinal study, the first step was to calculate the different time periods for data collection for each group. Time 0 (T0), was set at the beginning of the first academic year, that is, in September, for each Control and CLIL groups. Firstly, taking into account that the Catalan school year lasts 35 weeks, two tables were created for each of the schools, calculating, in the first place, the weekly number of hours of EFL for the Control Group and the weekly number of hours of EFL plus CLIL for the CLIL group. By accumulating the number of weekly hours, we obtained a detailed account of the accumulated monthly hours of English as well as the amount of hours of exposure to EFL and to EFL+CLIL at the end of each academic year.

T1 of data collection was set at the end of the first academic year (June) for the Control Group because the study was trying to find out the achievement and progress after one and two years of exposure to EFL and EFL+ CLIL. However, as the CLIL group had one extra hour of weekly exposure, and considering the initial objective of

keeping the number of hours of exposure constant at the time of data collection, T1 for the CLIL group came earlier in the first academic year (in April and not in June).

As it has already been mentioned, T2 data collection time for the second academic year occurred in September. This was only true for the CLIL group because, to be rigorous with the exact amount of exposure for data collection, T2 for the Control Group came halfway through the first term, when the Control Group had already done the same number of hours of English as the CLIL group had at the beginning of the second academic year. T3 of data collection happened at the end of the second academic year for the Control Group but, as it had already happened with T1, T3 data collection for the experimental CLIL group came earlier in the academic year because they received more weekly hours of English (see Tables 4.5 and 4.6)

	1 st year (5 th graders)										2 nd year (6 th graders)									
Months	Sep	Oct	Nov	Dec	Jan	Feb	Mar	Apr	May	Jun	Sep	Oct	Nov	Dec	Jan	Feb	Mar	Apr	May	Jun
CG*	T0 0h → T1 122.5h										T2 150h → T3 245 h									
CLIL*	T0 0h → T1 122.5										T2 150h ← T3 245h									

Table 4.5 Data collection times and hours of English. School A: Science students and their Control Group

* Control group: 3.5 h/w EFL

* CLIL group: 3.5h/w EFL+ 1h CLIL=4.5h/w

	5 th graders										6 th graders																																																	
Months	Sep	Oct	Nov	Dec	Jan	Feb	Mar	Ap	May	Jun	Sep	Oct	Nov	Dec	Jan	Feb	Mar	Ap	May	Jun																																								
CG*	T0					→					T1					105h					→					T2					→					T3					144h					→					T3					210 h				
CLIL*	T0					→					T1					→					T2					→					T3					144h					→					T3					210h									

Table 4.6 Data collection times and hours of instruction. Schools B and C: Arts & Crafts students and their Control Groups

* Control group: 3h/w EFL

* CLIL group: 3h/w EFL+ 1h CLIL= 4h/w

As has already been mentioned, the comparison had as its main objective to determine the achievement and progress of the students only exposed to EFL and compare it to the achievement and progress of the students exposed to EFL and CLIL(Science, in School A and Arts & Crafts in Schools B and C²³) at different moments keeping the hours of exposure constant. Data collection started with 5th graders in each school in 2010 for the Control Groups and went on for two consecutive school years. Data collection for the CLIL groups started one year later, in September 2011 and it also went on for two consecutive school years. The progress of the students in both groups was analysed over different time periods: first T0-T1, T1-T2 and T2-T3 and then T0-T3. This final comparison had to show the students' progression from the beginning to

²³ In order to obtain a big enough sample of students that would allow reliable statistical results, Schools B and C have been considered in this study as one group: CLIL Arts & Crafts students.

the end of the study after 245 hours of exposure, EFL only or EFL + CLIL, for the groups in school A and 210h for the groups in schools B and C.

	2010 - 2011	2011 - 2012	2012-13
CG	T0 → T1 0h → 122.5h	→ T2 → T3 (150h) → (245h)	
CLIL G		T0 → T1 → T2 (0h) → (122.5h) → (150h)	→ T3 (245h)

Table 4.7 Group comparisons: School A

	2010 - 2011	2011 - 2012	2012-13
CG	T0 → T1 0h → 105h	→ T2 → T3 (144h) → (210h)	
CLIL G		T0 → T1 → T2 (0h) → (105h) → (144h)	→ T3 (210h)

Table 4.8 Group Comparisons: Schools B and C

As has already been mentioned, apart from being rigorous at the time of collecting data by keeping the number of hours of exposure constant, one more variable was also taken into account for statistical data analysis: the previous number of hours of school exposure to English up to 5th grade.

For each school, the number of hours of previous exposure to English since the students were first exposed to the foreign language up to the end of 4th grade was calculated using the information provided by the head teachers in each school. This information was considered essential for the study as it showed that previous exposure to English was not the same in all the schools (see Table 4.7).

SCHOOLS	PREVIOUS EXPOSURE TO ENGLISH (up to 5 th grade)	STARTING AGE
A	420h	P3 ²⁴
B	332.5h	P3
C	455h	P3

Table 4.9 Hours of previous exposure to English at school up to 5th grade

Because of the different number of hours of previous exposure to English in each school and the fact that each school was devoting a different number of hours to English per week and, therefore, the accumulated annual number of English hours varied from school to school, statistical analyses in the study also incorporated previous exposure as a variable.

Apart from the previous number of hours of exposure, the different proficiency levels of the students were also taken into account for the statistical analysis. The groups in each of the Control and CLIL groups, were divided into two groups, High achievers and Low achievers. The division was carried out taking into account the results of the initial listening and reading tests. It must be said that the number of correct answers at this stage was low. In the case of the CLIL students exposed to Science as well as their Control Group, the students who scored between 0 and 4.5 were considered Low achievers and those who scored above 4.5 were classified as High achievers. In the case of those students in the CLIL Arts & crafts group and the students in their Control

²⁴ P3 refers to the first year in Early Years Education

Group, those who scored below 2.75 were considered Low achievers and those who scored above 2.75 were considered High achievers.

4.3.2 Data Collection Instruments and procedure

This section describes the different instruments used to elicit and measure the language competence of the young learners involved in the project, as well as the qualitative instruments used to collect bio data of the participants, the students and teachers' background questionnaires, and the instruments used for carrying out class observations and for analyzing class methodology. The section also includes the measures used to correct the tests and the written assignments.

4.3.2.1 Language Proficiency Tests

Current approaches to the assessment of young language learners promote the use of instruments that comply with the needs and special characteristics of young learners' cognitive levels, topics and tasks (Inbar-Louries et al., 2009). Children need to be tested through tasks that make use of the type of language or situations that make sense to them, that replicate real life situations, that are challenging and that they can relate to their already existing knowledge of the world (Donaldson, 1978; Mackay, 2006).

The main criteria in this study for selecting the tests was that they were adequate to the age of the learners, contextualised around topics of their interest and suitable for the language knowledge and skills to be assessed. These criteria emerged from the findings of previous studies (Vallbona, 2009, Victori and Vallbona, 2010) which concluded that some of the results obtained from CLIL students in a very similar context to the one in the study might have been affected by the nature of the tests used.

As the informants in the study were students in years 5 and 6 of primary education, Cambridge Young Learners' exams (YLE)²⁵ were administered to the groups. These tests are aimed at students between 7 and 12 years of age and cover the different language skills. According to the University of Cambridge ESOL examination board, the tests address relevant and meaningful language use for young learners and cover the four skills (listening, speaking, reading and writing). YLE²⁶ tests are organized in three different levels: *Starters* (aimed at children in years 2 and 3 in primary education), *Movers* (for learners in years 4 and 5, and *Flyers* (for year 6 learners). As this classification was only an approximate recommendation of the publishers of the tests, it was decided to carry out a small-scale pilot study covering the main skills which were going to be tested in the study. This would determine the approximate level of the participants and would avoid a possible ceiling effect in the results of the tests during the study.

Preliminary testing was conducted in two different schools: a primary state school and a state-assisted primary school in very similar contexts to those involved in the study. The tests were administered to young learners at the end of their 4th and 6th grades of primary education. The reason why 4th graders were chosen is because the present study had as its youngest participants, students at the beginning of their 5th grade, who had just obviously finished their 4th grade of primary. The learners in the state school had only been exposed to EFL curricular English; however, the learners in the other school had been exposed to EFL and CLIL. The main aim was to find out how similar learners performed with different level tests as well as to find out the highest

²⁵ <http://www.cambridgeesol.org/exams/yle/index.html>

²⁶ <http://www.cambridgeenglish.org/exams-and-qualifications/young-learners/>

score of the EFL students and the highest score of the EFL+CLIL students in order to determine the most adequate level tests and also to avoid a ceiling effect which would clearly reduce variability in the gathered data. The analysis of the scores in the tests contributed to the final decision on the type and the level of tests to be used for the study. As the schools in the pilot study were two form entry schools, the students in each class were given very similar tasks at different levels (*Movers and Flyers*) in order to see how they performed in the different skills. The results of the tests varied across the different skills and the different tasks. Results suggested that the highest level, *Flyers*, would be too difficult for some of the learners involved in the project in the reading comprehension and cloze tests, and the results that we would obtain might not be realistic. Nevertheless, students obtained good results in the listening and in some of the reading tests of the middle level, *Movers*. So, the final choice for the tests used in the study was the following:

TESTS	SKILL	LEVEL	TYPE OF TEST
A1	Listening	Flyers	Listening for specific information
B1	Listening	Flyers	Listening for specific information
A1/A2	Reading	Flyers	Reading and matching
B1/B2	Reading	Flyers	True/false reading comprehension task
C	Reading (Cloze type Test)	Movers	Reading for specific information
D1/D2	Reading	Movers	Reading for specific information

Table 4.10 Tests, skills and level

As this was a longitudinal study and there were several data collection times along the two years, some of the tests were changed during the second year in order to avoid a memory effect on the answers provided by the students. The listening tests were kept unchanged as was the cloze test. However, three of the reading tests were changed. The format of the exams was kept intact but the contents were changed. The texts were taken from the same set of YLE from Cambridge.

4.3.2.1.1 Listening tests

Listening Test A1 (Level: Flyers) was a Listening for specific information type task in which the students were asked to listen to 5 short conversations between an adult and a child. The child was talking about what she had done on the day of her birthday. For each of the 5 questions in the exercise, there were three different pictures representing possible answers, but only one of them was correct. The students had to tick the picture that best represented the right answer. One point was given to each correct answer (see Appendix A).

Listening Test B1 (Level: Flyers) was also a Listening for specific information task. The students were asked to listen for words, names and detailed information. In this exercise, there were two different sets of pictures: on the left hand side, there were six pictures (including an example) representing different school subjects. On the right hand side, there was a set of pictures with letters but no words. Students had to listen to a conversation between a teacher and a student who was asking the teacher what to take to school for the next day. They also had to match each of the pictures on the right with one of the subjects on the left. The names of the subjects were not in the order in which they appeared in the typescript. The total number of answers in the exercise was also 5. One point was given for each correct answer (see Appendix B).

4.3.2.1.2 Reading tests

In order to assess the students' progress in reading, four different tests from Cambridge Young Learners' Exams were chosen.

Reading Tests A1 and A2 (Level: Flyers) The test consisted of fifteen words and ten definitions (sentences that described or explained ten of the fifteen words). The students had to read the definitions and write the matching word beside the correct definition. The semantic areas covered in this exercise were the names of jobs, clothes, vehicles and places. There were 10 definitions excluding that given as the example. One point was given for each correct answer (see Appendix C).

Reading Tests B1 and B2 (Level: Flyers) In this exercise there was a big picture and seven sentences; some of these sentences described the picture correctly but some others did not. The child was asked to write *Yes* if the sentence was true and described something in the picture and *No* if the sentence was false. One point was given for each correct answer (see Appendix D).

Reading Test C (Level: Movers) This was a reading for specific information task in which there was a text with some missing words (gaps) in it. Each gap needed to be filled with a word (a noun, adjective or verb). Next to the text, the students had a box with a choice of three words per gap. They were asked to put the correct word into the gap. One point was given for each correct answer. Tested use of English (see Appendix E).

Reading Tests D1 and D2 (Level: Movers). This type of test was a story in three parts. Each part of the story had a matching picture. After each part, the students had to complete 10 sentences about the story using one, two or three words. This task

was scored under two different categories: accuracy and comprehension. In order to assess accuracy, following what had been done with the previous tests; one point was given for each correct answer. However, as the exercise required writing one, two or three words as the answers, and considering that in many answers it was clear that the student had understood the text and the question and was answering the question but did not write the answer correctly, the answer was considered valid from the comprehension category and was awarded one point (see Appendix F).

In order to avoid a memory effect on the answers provided by the learners which might have affected the reliability of the final results, during the second year of the study, three of the reading tests (A1, B1 and D1) were changed: the format of the new tests was kept intact but the contents varied. The cloze-type test was kept the same as well as the listening ones.

4.3.2.1.3 Written composition

A written composition was given to the students in their own classroom in order to gauge their writing skills. Although the Cambridge listening and reading tests administered to students already involved some kind of writing, this was only at the word/phrase level and involved mainly copying the answers into the exercises themselves. In order to assess and analyse their achievement and progress in the writing skill, it was decided to ask students to carry out a *free* writing. Writing is one of the four skills stated in the Catalan Curriculum for foreign language learning²⁷. The curriculum itself states a clear common aim for L2 learners in primary education: when children

27

http://www20.gencat.cat/docs/Educacio/Home/Departament/Publicacions/Col_leccions/Curriculum/educacio_primaria/curriculum_ep.pdf

finish their primary schooling, they should be able to write in a communicative and creative way, short texts attending to the type of text, the addressee and purpose of the written text.

In order to control time and topic constraints and to make the compositions comparable (Wolfe Quintero et al., 1998), the participants were given the same amount of time, 15 minutes, to write an essay on the topic: *My life*. Most of the writers did not use up all the time because of their limited language proficiency. As for the choice of topic, *My life* was considered a familiar and personal topic about which all children could write. Drew and Sørheim (2009: 88) point out that ‘children usually have stories to tell about themselves and the world they live in, which they are keen to share with others’ and ‘despite their linguistic shortcomings, are able, and quite often happy, to share episodes from their lives’.

Before they started writing, the students were given some prompts to help them get begin and they were instructed to write as much as they could, and to write in English as much as they were able to. In external and formal assessment tasks, prompts for writing tasks are considered important (Mckay, 2006) as they provide a guide and a starting point. The researcher made it clear to the students that the task was not considered an exam and that it would not be assessed by their own teacher. The aim was to make the writing conditions as anxiety-free as possible and the whole testing experience a positive one (Inbar-Lourie et al., 2009). As clear instructions on how to carry out a task are needed for assessing very young learners (Mckay, 2006), they were given in Catalan to avoid any misunderstandings and dictionaries were not allowed during writing time. The students were not helped at any moment by either the researcher or their own class teachers who did not take part in the process.

4.3.2.2 Students' Background Questionnaire

In order to elicit information about the learners themselves, their personal history of language learning, their exposure to English and their personal attitudes towards the studying of English, a background questionnaire was given to them in Catalan (see appendix J) which included two types of questions. The first type elicited information on the learners' personal history and their language learning history as well as any exposure to the English language outside the regular school periods. A second group of questions elicited behavioural information, such as how often, how and when they used English outside the regular English classes. The questions in this part were Lickert-type questions. The participants in the study completed the questionnaire once in the first session before the process of data collection actually started. The administrator tried to clarify all possible doubts that learners might have had. The questionnaire was adapted from the one used for the BAF project (Muñoz, 2006) because it had already been used in a context similar to the one in this study.

4.3.2.3 CLIL Teachers' Background Questionnaire

A CLIL teachers' background questionnaire (see Appendix K) was specifically designed by the members of the CLILSLA research group together with other members of the CLILSLA Project. The three CLIL teachers involved in the study responded to this questionnaire during the first semester of CLIL implementation in the schools. It provided information on their personal and professional background as well as on their experience in implementing CLIL. The questionnaire consisted of 41 items and was divided into three parts: the first part contained 11 open-and-close ended questions and elicited information on the teachers' personal details, their academic and professional background, their position in the school and the contact they kept with English in their

private lives. All the questions in this part were open-ended questions except those concerned with their regular contact with English that were Lickert-type ones. The second part in the questionnaire (9 open-ended and closed-questions) focussed specifically on their CLIL professional background: their previous CLIL experience and training, the reasons for being the CLIL teachers in their schools, and their perceptions on the challenges and difficulties of being a CLIL teacher (les seves expectatives). The third part of the questionnaire, 21 Lickert-type questions, concentrated on the implementation of CLIL in their schools. The teachers were asked about the way their coordination with the EFL teachers, the use of different types of CLIL materials, the activities and strategies used in the class, the use of the mother tongue in CLIL classes, the assessment criteria applied to their students and their general feelings as CLIL teachers. (Appendix K).

4.3.2.4 CLIL Teachers' Opinion Questionnaires

After each year of CLIL implementation the relevant CLIL teachers were asked to answer an opinion questionnaire on the implementation of CLIL in their schools. The 15-item questionnaire (also created by members of CLILSLA Project) was sent online to the teachers who, in turn, sent it back to the researcher. The survey allowed us to obtain a range of information on how the teachers viewed the implementation of CLIL in their schools and how equipped they felt in terms of their language proficiency, the methodology and the strategies used to implement the CLIL programme. The teachers also provided their opinion on the support received from the school and from the other teachers, and how they felt at the end of their CLIL experience.

4.3.2.5 Class Observations

The way in which teaching performance promotes language acquisition in content-based classrooms has extensively been studied in naturalistic learning contexts, such as, for example, the Canadian context (Lyster, 2004; Lyster and Ranta, 1997). According to De Graaf et al. (2007) the teaching performance for language acquisition in a CLIL context is made up of three important characteristics: functional communication, simultaneous attention to form and meaning and the type of corrective feedback given to students which needs to be considered within a broader framework for language acquisition: exposure, use and motivation. De Graaf et al. (2007) carried out a study in secondary schools in the Netherlands and concluded that effective teaching performance facilitates language development and proficiency in CLIL classes.

In order to find practical evidence of the relationship between methodological aspects and the results of the study, CLIL and EFL class observations were carried out in each school during the study: two in the first year of implementation of the study and two during the second year. The observations took place during the *Science* and *Arts & crafts* classes as well as during the EFL classes. The protocol established was a combination of observation and note-taking following a checklist observation tool based on the SIOP Model (Sheltered Instruction Observation Protocol) (Echevarria et al. 2007). The SIOP protocol was developed to check the techniques incorporated and used by teachers in the context of Sheltered Instruction in the USA in order to provide meaningful instruction in the content areas for Limited English Proficient (LEP) students.

The observer took a series of field notes on aspects such as methodology and strategies used in the class, the activities provided by the teacher, the teacher's input and

the students' output, the use of English and the use of the first language and any other aspect that seemed relevant. The field notes were used to write a report on the weaknesses and strengths of the session which were then transferred into a chart created by members of the CLILSLA project based on the model above indicated, the SIOP model and on Ballester's model (Ballester, 2010), who in her study, adapted different models and protocols for CLIL class observation to the Catalan School context. (See Appendix M). The chart was divided into six different categories: Preparation, Building background, Comprehensible input, Strategies, Interaction, Practice-application, Lesson delivery. Each category was rated in a scale from *Highly evident* to *Not applicable*. In order to clarify and discuss any aspects of the lesson that the observer could not obtain from the simple observation, several informal exchanges took place with the relevant teachers. The outcomes of the exchanges helped the observer to transfer information to the charts created and used for the analysis. The sessions were useful, in the sense that they allowed to see the variety of activities used by the teacher, as well as the methodologies and strategies applied in the class, and the use of the mother tongue in the sessions in certain situations, in order to evaluate aspects on both the teacher's and the students' side, which were valuable for the general interpretation of the results obtained in the present study. They were also useful to make recommendations to the CLIL teachers for them to be able to improve their future sessions.

In order to be able to compare methodological aspects between the EFL and CLIL classes, several class observations were also carried out in the EFL classes. The groups in all the school followed a textbook and most of the sessions observed had a similar structure: they began with a listening or speaking activity aimed at introducing the topic. Depending on the lesson, the students were asked to read a short text and then, several vocabulary and language-focused exercises were carried out. The book used

integrated the four skills and provided plenty of opportunities for the students to participate in the class through oral or written activities. In several informal exchanges, the EFL teachers explained that they followed the book and that, occasionally, they used songs, games and short films to provide the students with natural input to foster their motivation.

4.3.3 Writing Measures

This section will provide an explanation of the measures used to assess writing and the reasons why they were adopted. Measuring writing proficiency is frequently a more complicated issue than measuring language proficiency tests especially in contexts where ‘the participants are very young learners who have not yet fully developed the imaginative and organizational skills needed to produce extended writing’ (Taylor and Saville 2002: 324).

In general terms, Second Language writing has often been measured either holistically, using band scales, or analytically, through quantifiable measures. In holistic scoring a written text is evaluated for its overall quality whereas analytic measures are considered more reliable as they contribute to rate different aspects of the compositions separately and to provide separate scores to each feature (Celaya, Pérez-Vidal and Torras, 2001). These measures may be developed by different types of learners under different conditions, so the results can be generalized across different contexts.

Wolfe-Quintero, Inagaki and Kim (1998) thoroughly reviewed 39 studies in order to account for the most widely used analytic measures that better described writing development and to see how those measures correlated with different levels of proficiency. Complexity, Accuracy and Fluency (CAF) measures have been extensively used in assessing second language writing. Generally speaking, CAF measures have

been considered by researchers as descriptors that can adequately capture L2 performance and can be indicators of the learners' proficiency; they have also been used to measure progress in language learning (Skehan, 1998; Ellis and Barkhuizen, 2005). However, as many researchers have pointed out, although Complexity, Accuracy and Fluency have been thoroughly investigated, there are still controversial questions as to how the constructs can be defined, how these three constructs can be best-operationalised as components of L2 proficiency and how they can be reliably measured (Housen and Kuiken, 2009; Pallotti, 2009).

Accuracy is probably the least controversial of the three. It refers to the degree of deviance from a particular norm; it is concerned with how well language is produced in relation to the rule system of the target language. (Wolfe-Quintero et al., 1998). In the words of Housen and Kuiken (2009), 'fluency typically refers to a person's general language proficiency, particularly as characterized by perceptions of ease, eloquence, and smoothness of speech or writing' (p.3). Other authors refer to fluency in different terms: ease of retrieval of language items; the speed of language production and the capacity to use language in real time (Lennon, 1990; Skehan and Foster, 1999; Hulstijn, 2001). Complexity is the most complex and ambiguous of the three. It can refer to both, language and task complexity. Linguistic complexity focuses both, on grammatical and lexical complexity. According to Housen and Kiuken (2009), linguistic complexity has been interpreted as 'the size, elaborateness, richness and diversity of the learners' linguistic L2 system'(p. 4). The general underlying assumption behind CAF measures is that, as learners progress and become more proficient, they write more fluently, more accurately and their writings are more grammatically and lexically complex (Drew, 2010). Some authors, however, Skehan and Foster (1999), have considered these measures in competition with one another. For these authors, Fluency is seen as "the

capacity to use language in real time, to emphasize meanings, possibly drawing on more lexicalized systems” (1999: 96-97) and Complexity refers to “the capacity to use more advanced language, with the possibility that such language may not be controlled so effectively. According to them, when learners are faced with cognitively demanding tasks, they will attend to conveying meaning first, and to accuracy and linguistic complexity last.

This study wanted to analyse, among other aspects, the written level of development of YL interlanguage exposed to EFL and CLIL. The relevant literature has provided several studies on the writing development of EFL learners. However, most of the studies have been conducted with grade 6 students and above: Vigrestand (2006) compared the fluency and complexity of Dutch and Norwegian 7th to 9th graders; Drew (2003) compared the fluency and complexity of 7th graders writing; Tjerandsen (1999) investigated the use of subordinate clauses by 8th graders ; Hasselgreen (2010) assessed the writings of primary students in four different European countries using a scale based on the Common European Framework of Reference; Miret (2009) compared the proficiency of CLIL and non CLIL secondary students writing development; Celaya et al.(2001) analysed the short and mid-term effects of an earlier start on the written production of 6th graders; Ojeda (2009) analysed themes and vocabulary in CLIL instruction. Very little research, however, has been carried out so far on the development of the written production of very young EFL language learners (below grade 6) in low-input situations comparing very young EFL learners and CLIL learners.

The analysis of the development of young learners in their first stages of the acquisition of EFL through a corpus of written products entails taking decisions on how

to describe and how to measure their written linguistic change over time (Torras et al., 2006). The interlanguage of learners is dynamic and is in constant change; as young learners receive more input and they are able to revise and restructure their L2 system (Lightbrown and Spada, 2006). An initial analysis of the written compositions revealed several problems. In general, the compositions were very short. The average number of words per composition for the Control and the CLIL Groups was approximately 65 words at T0. In addition, most of the compositions lacked punctuation marks or these were used incorrectly. They contained quite a few borrowings from Catalan or Spanish (e.g. my mother is *advocat*) as well as lexical inventions (e.g. My favourite *assignatur*, from the Catalan word *assignatura* or my father is *en par*, in Spanish, *en paro*). As the pupils had received so little instruction in English, it was necessary to find adequate measures that would be suitable for learners at the beginning but also at the end of the study.

Following Wolfe-Quintero et al. (1998), the compositions that the students handed in were analyzed under three categories: complexity, fluency and accuracy and two types of calculations were used: simple frequency counts of a particular unit and ratio calculations which allowed us to relate units to other reference units. In order to analyse the students' fluency, Wolfe Quintero's definition of fluency was taken into account. He refers to fluency in the following way: 'fluency means that more words and more structures are accessed in a limited time, whereas a lack of fluency means that only a few words and structures are accessed' (p. 14).

For the purpose of this study, several measures were considered suitable indicators of the development of young learners' fluency: Total Number of Words (TNW), Total Number of Words in English (TNWE), the ratio Total Number of Words

in English/Total Number of Words (TNWE/TNW) and the Total Number of Units (TNU). In order to count the total number of words in English, a general word count for the total number of words was carried out, followed by a manual recount of the words that were genuinely English words. Words with spelling mistakes (and that clearly showed a resemblance with English words, for example, *swiming pol*, *cocrodil*, *granfader*) were not excluded from the word count of English words; however, words written in their L1, either Catalan or Spanish, or words that the students had invented and that objectively did not have any relation to real English words were not taken into consideration. Contracted forms were accounted as one word.

The number of T-units has been used as a measure of fluency for young learners instead of sentences (Drew, 2010). The concept of T-unit was developed by Hunt (1965: 49) who defined T-units as ‘one main clause plus the subordinate clauses attached or embedded within it’. T stands for ‘terminable’, and refers to units that may be grammatically terminated by sentence final punctuation. In the analysis of the writings in this study, T-units were not considered for the study because, as it has been pointed out, children did not always punctuate or they did not punctuate correctly and their writings were very much made up of formulaic chunks of language and there were very few instances of subordination. Instead of T-unit, writings were analyzed using the concept of *units*, created by the researcher, that is, ‘meaningful chunks of language that contain, at least, one finite or non-finite verb’. In this way, the number of units would not be affected by punctuation. The number of units was considered appropriate for the present study because, on the one hand, it would show to what extent the pupils were able to put their words together in meaningful chunks of language. On the other hand it allowed us to account for chunks that were not correct according to the rules of standard English and that reflected some of the common errors of young learners’ interlanguage

such as the lack of subjects and objects or the lack of agreement between subjects and verbs. Young children's writing contains many examples of what has been called *formulaic sequences* or *chunks of language*. Wray (2002: 9) defined them as 'a sequence, continuous or discontinuous, of words or other elements, which is, or appears to be prefabricated, that is, stored and retrieved whole, from memory at the time of use'. The value of the chunks is that 'very young learners can draw on them in moments of communicative pressure' (Mckay, 2006: 36). Chunks play an important role for further language learning because young learners tend to accumulate memory-based chunks which they combine to build language.

Accuracy was another of the aspects considered in the analysis of writing. The language of young learners is in constant evolution and their natural tendency is to attend to meaning rather than to form (Bialystok, 2001). According to Ellis (1996), accuracy is related to language automatization: the more the language is automatized the more accurate it becomes. However, this expectation may not always correspond to reality and as the language becomes more sophisticated, the learners may produce more errors as the language is not completely automatized. Young learners are likely to make a high number of mistakes (Drew, 2010). In spite of this, and because the language proficiency tests, except for one, had been approached and corrected from the accuracy point of view (see section 4.4), it was decided to incorporate two measures of accuracy in the analysis of the learners' written production to capture the general picture of accuracy among young learners: Total Number of Error Free Units (TNEFU) and the ratio between the Total Number of Error-Free units in relation to the Total Number of Units (TNEFU/TNU), that is, the percentage of units with no grammatical errors according to the rules of standard English. In order to carry out the analysis, the units

were classified into two different groups: error-free units, and rejected units. Totally error-free units were the only ones used for the analysis.

In the present study, complexity was analyzed in terms of syntax (Syntactic Complexity) and in terms of vocabulary (Lexical Complexity). As the informants were young learners, they were expected to write simple units with no coordination or subordination. In spite of the initial lack of syntactic complexity of the compositions analyzed but based on the assumption that as language learners progress, their grammar becomes more complex, it was decided to incorporate two measures of syntactic complexity: Instances of Coordinated Units (ICU) and the Instances of Subordinate Units (ISU), that is, the number of examples of coordinate and subordinate units found in the text, were calculated. Therefore units of the type *I went to the park and I played with my friends* were accounted for as one example of coordination and units such as *I like the school because it is funny* or *I prefer to go swimming* were accounted for as an example of subordination.

Variation and sophistication of language have been recognized of strong indicators of Lexical Complexity (Wolfe-Quintero et al., 1998: 101-104). Verb types have been studied as one of these indicators as it has often been assumed that children in the early stages of language learning will learn vocabulary more efficiently. The initial writings handed in were mostly written using formulaic units which involved the verb to be (*My name is, My favourite animal is, My school is...*) As their writings became more sophisticated, children increased the number and variety of lexical verbs as well as the use of adjectives. Therefore, the Number of Lexical Verbs and adjectives were considered as a measure to account for their lexical variation and sophistication of their writings (Linnarud, 1986).

The texts were therefore analyzed on the basis of the following measures:

LEXICAL COMPLEXITY	SYNTACTIC COMPLEXITY	ACCURACY	FLUENCY
TNLV			TNW
%TNLV/TNWE	ICU	TNEFU	TNWE
TNAdj	ISU	%TNEFU/TNU	%TNWE/TNW
%TNAdj/TNWE			TNU

Table 4.11 Writing measures

TNW: Total Number of Words

TNWE: Total Number of Words in English

%TNWE/TNW: Ratio Total Number of Words in English/Total Number of Words

TNU: Total Number of Units

TNEFU: Total Number of Error Free Units

%TNEFU/TNU: Ratio Total of Error Free Units/Total Number of Units

TNLV: Total Number of Lexical Verbs

%TNLV/TNWE: Ratio Total Number of Lexical Verbs/Total Number of Words in English

TNAdj: Total Number of Adjectives

%TNAdj/TNWE: Ratio Total Number of Adjectives/Total Number of Words in English

ICU: Instances of Coordinated Units

ISU: Instances of Subordinate Units

4.4 Data analysis and statistical procedures

This section describes the statistical procedures undertaken in order to analyse data obtained through the different tests administered to the participants. Although quantitative and qualitative data were collected in this study, the approach taken was mainly quantitative. Qualitative data was used to support the main findings of the statistical analysis.

First of all, the answers in the listening and reading tests were evaluated as either correct or incorrect. However, as has already been explained, Reading question 4 was analysed from two perspectives: accuracy and comprehension. Therefore, in this question, children were also rewarded for their level of comprehension. For example, a child who wrote '*bed*' instead of '*on the bed*' in answer to the question 'Where is the toy?' was evaluated positively in a separate column entitled Comprehension.

The two listening tests were analysed together. The percentage of correct answers was calculated based on the number of correct answers provided by the students in relation to the total number of answers in the two tests. Reading was also analysed in two different ways: as in the case of listening, in the first place, the percentage of correct answers was calculated taking into account the total number of answers in the four reading tests and then, the percentage of correct answers of each of the individual reading tests.

Compositions were first converted into computer-readable files, keeping their original form; spelling mistakes, punctuation, grammar and lexical errors were left untouched. In the analysis of the compositions, the CLAWS 5 Tagger (University of Lancaster) was used in the first place for the general word count and the grammatical tagging of the words, followed by a manual recount and revision of the outcomes of the tag set, due to the fact that, especially in the initial stages, children wrote many words in Catalan. This manual recount provided us with the Total Number of Words written in English (TNWE) as well as the Total Number of Lexical Verbs and the Total Number of Adjectives. Contractions such as *isn't* were counted as one word. When counting verb types, following the definition provided by Read (2000) of verb lexemes 'a *single lexical item which may consist of more than one-word form*' were counted. So if the students used *go* and *went* in the same text, these forms were accounted as one form.

Verbs forms which were used more than once were only accounted once. A careful manual qualitative analysis of the compositions was carried out with the purpose in mind of determining, in the first place, the Total Number of Units per composition (TNU). The process was divided into three stages: in the first stage, the texts were divided into units according to the definition provided in section 4.3.3 ('a meaningful chunk of language that contains at least one finite or non-finite verb written in English'). The division was followed by a manual recount of the number of units per text. In the second stage, units without a finite or non-finite verb form were rejected (*My favourite 'menjar' macaronis*) and excluded from the unit count, as well as units that contained verbs not written in English (*Today football, yesterday(vaig anar) patinatge*). A second unit recount was then carried out with one clear objective: to find out the number of Error-free Units (TNEFU) in each text. All the units that had grammar mistakes according to the rules of standard English were excluded from the second recount. A third manual recount of the compositions was also used to account for the number of Instances of Coordinated units (ICU) and the number of Instances of Subordinate units (ISU) (see section 4.3.3).

To ensure the reliability of the scoring of the written compositions, an independent rater was required. Inter-rater measures measured as percentage of agreement on a random sample of the writings collected at three different times was estimated at 95%.

As for data analysis, in order to ensure the reliability of the data, a validation of the internal consistency of the variables in the database was carried out. Bivariate tests of homogeneity were applied to make sure that the elements of the main explanatory variable (Group) were homogeneous in terms of the other two variables (Type of School and Initial Proficiency Level).

In order to examine the differences between the results obtained by the groups exposed just to EFL sessions, and the results obtained by the groups exposed to both EFL and CLIL classes in the listening, reading and writing tests, generalized linear mixed models were used. The covariates included in the models were time (T0, T1, T2 and T3), type of school (Science/Arts & Crafts), group (EFL/CLIL), initial proficiency level (High and Low) and their interactions with Time and Group. The school was considered a random effect in the comparison CLIL Science vs. CLIL Arts & Crafts. Adjusted means with 95% confidence intervals were estimated for each group at different times, for the interaction between group and type of school. Intergroup and intragroup comparisons were carried out in order to provide an overview of the results. When multiple comparisons were computed, p-values were adjusted using Tukey's correction. The analysis was carried out with software SAS v 9.2. and the significance level was set at $p < 0.05$.

CHAPTER 5 Results

In this chapter the findings related to each research question will be presented. First, the results of the listening and reading tests, as well as the results of the analysis of the written production of the EFL Control Group exposed to the mandatory hours of curricular English, will be compared to those obtained by CLIL students exposed to the mandatory curricular EFL hours plus one hour a week of the subject Natural and Social Sciences. As explained in Chapter 4 and in order to obtain reliable results, the number of hours of exposure to the language was kept constant. Secondly, the results in listening, reading and writing of the EFL Control Group compared to those students exposed to one hour a week of the subject Arts & Crafts in English, also keeping the number of hours of exposure constant, will also be reported. Finally, the results obtained by CLIL Science students and CLIL Arts & Crafts students will be compared.

As has been pointed out in the previous chapter, the analysis of the listening, reading and writing tests was performed using longitudinal logistic regression models. Adjusted means with 95% confidence intervals were estimated for each group at different times and time periods. Intergroup and intragroup comparisons are presented in order to provide an overview of the achievement and progression results. The first comparison, the intergroup comparison, is concerned with the achievement results obtained at different times (T0, T1, T2 and T3). The intragroup comparison presents the improvement results obtained by each of the groups after e different time periods: T0-T1, T1-T2, T2-T3 and T0-T3.

The results of the interaction between group and proficiency level (High and Low achievers) will be presented for the listening and reading tests, as well as for a the following measures of Complexity (TNAdj, TNLV), Accuracy (TNEFU) and Fluency (TNWE, TNU) in writing. The division between High and Low achievers was carried

out following statistical analysis of the results obtained at T0. In general terms, the results obtained in the language proficiency tests at T0 were very low for all the students. In order to get groups that allowed reliable statistical analysis, the groups were divided into two groups, High achievers and Low achievers, taking into account the students' performance at T0. As has already been explained in the previous chapter, in the case of the CLIL students exposed to Science as well as their Control Group, the students who scored between 0 and 4.5 were considered Low achievers and those who scored above 4.5 were classified as High achievers. In the case of those students in the CLIL Arts & Crafts group and the students in their Control Group, those who scored below 2.75 were considered Low achievers and those who scored above 2.75 were considered High achievers. As explained in the previous chapter, the analysis of the interaction Group/High and Low achievers was also carried out using longitudinal logistic regression models. When multiple comparisons were computed, p values were adjusted using Tukey's correction. Finally, a summary of the achievement and improvement results is also provided at the end of the listening and reading sections as well as at the end of the writing report.

5.1 CLIL Science results

In sections 5.1.1 and 5.1.2, a description of the results obtained by CLIL Science students compared to those in their Control Group will be presented in order to answer the following research question and sub questions: RQ1: Keeping the number of hours of exposure to English the same for both groups, CLIL and EFL, do the CLIL students' listening and reading skills benefit from their exposure to the CLIL experience?

RQ1.1 Are there any differences in achievement between CLIL and EFL learners statistically significantly in favour of CLIL learners at different times (T0, T1, T2, T3)?

RQ1.2 Are there any differences in progress between CLIL and EFL learners significantly in favour of CLIL learners after one year (T0-T1) and two years (T2-T3) of CLIL implementation? Are there differences in favour of CLIL students when we consider their progress from T0-T3? RQ 1.3 How does the initial level of English proficiency affect the students' performance in the CLIL and the EFL group?

5.1.1 Listening test results

The intergroup comparisons between the CLIL and Control Groups do not yield the same results for all the times. The comparison of the scores at T0 shows that, despite the slightly higher percentage obtained by CLIL students (43%), the differences between the CLIL and the Control Groups were not significant ($F= 0.28$, $p= 0.5979$); nor were the differences significant between the same groups at T1 ($F= 0.01$, $p= 0.9197$) and T2 ($F =0.00$, $p= 0.9995$). The mean results at T2 were the same for both groups. The results at T3, however, showed a statistically significant advantage in favour of the CLIL group²⁸ ($F= 4.81$, $p= 0.0289$).

	Group	Mean	StdErr	Lower	Upper	F value	P value
T0	CLIL	43.1%	3.6%	36.3%	50.2%	0.28	0.5979
	Control	40.7%	2.9%	35.1%	46.5%		
T1	CLIL	52.7%	3.8%	45.3%	60.0%	0.01	0.9197
	Control	53.2%	3.0%	47.3%	59.0%		
T2	CLIL	56.8%	3.5%	49.9%	63.5%	0.00	0.9995
	Control	56.8%	3.0%	50.9%	62.6%		
T3	CLIL	74.0%	2.9%	67.9%	79.3%	4.81	0.0289
	Control	64.2%	3.2%	57.7%	70.2%		

Table 5.1 Listening achievement results

²⁸ Significantly different results have been highlighted in yellow

An intragroup comparison was also carried out to determine the percentage in the students' improvement at different time periods. In general terms, the students' scores in the listening tests showed a linear increase for both the CLIL and the Control Groups.

The Control Group increased significantly from T0 to T1 ($p= 0.0028$), but there were no statistically significant differences for the CLIL group ($p= 0.1370$). Between T1 and T2, there were no significant differences for any of the groups. Nevertheless, between T2 and T3, there was almost no change for the Control Group, whereas the CLIL group increased significantly from 56.8% to 74% ($p<.0001$). The improvement from T0 to T3 was significant for both groups, the CLIL ($p<.0001$) and the Control Groups ($p<.0001$). As has already been reported in the achievement section, the CLIL Group significantly outperformed the Control group at T3. (see Figure 5.1 and Table 5.2).

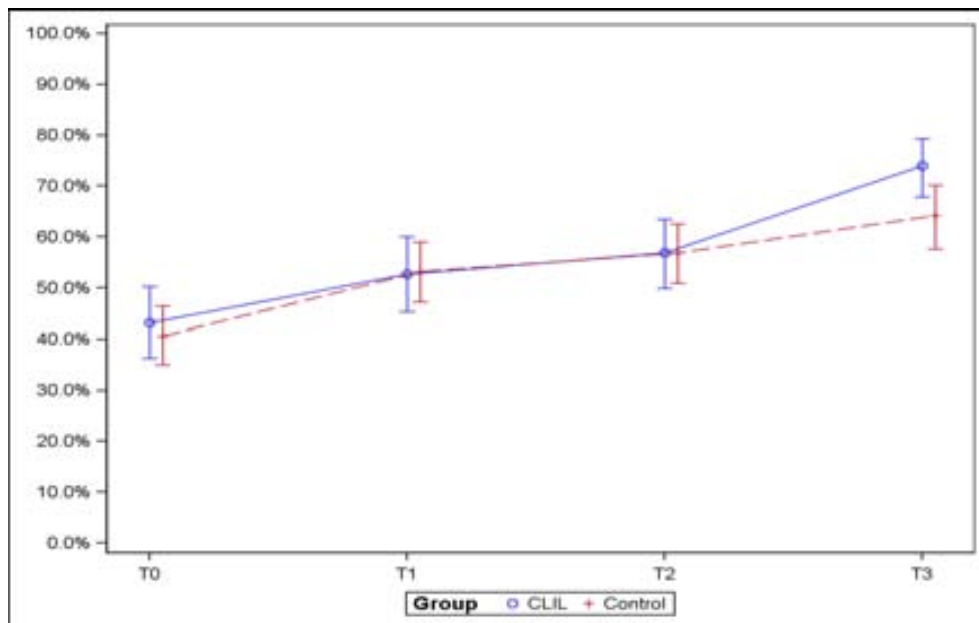


Figure 5.1 Listening improvement results

	I0_1	I1_2	I2_3	I0-3
CLIL	22.1%	7.9%	30.2%	71.5%
Control	30.7%	6.9%	13.0%	57.9%

Table 5.2 Percentage of improvement in listening

The intergroup comparison, when the variables Group/Proficiency level were taken into account, showed that there were no statistically significant differences between the two groups, CLIL High and Control High, at any of the times tested. Descriptive statistics, however, showed that at times T0, T2 and T3, High achievers in the CLIL group obtained slightly better scores than High achievers in the Control Group. However, at T1, High achievers in the Control Group seemed to be slightly better than those students in the CLIL group. Nevertheless, the mean differences at all times were always between 2% and 4% (see Table 5.3 below).

	Proficiency	Group	Mean	StdErr	Lower	Upper	F Value	P Value
T0	High	CLIL	58.2%	4.3%	49.5%	66.3%	0.46	0.4966
	High	Control	54.0%	4.3%	45.5%	62.2%		
T1	High	CLIL	60.2%	4.5%	51.1%	68.7%	0.16	0.6926
	High	Control	62.7%	4.1%	54.2%	70.4%		
T2	High	CLIL	62.2%	4.2%	53.7%	70.1%	0.09	0.7695
	High	Control	60.4%	4.4%	51.6%	68.6%		
T3	High	CLIL	77.8%	3.4%	70.5%	83.8%	0.30	0.5862
	High	Control	74.9%	4.1%	66.0%	82.1%		

Table 5.3 Listening achievement results: Group/High achievers interaction

Table 5.4 below shows that the difference in the achievement scores between CLIL Low achievers compared to Control Low achievers was only statistically significant in favour of the CLIL Low achievers group at T3 ($F= 0.10$, $p= 0.0037$). At

this time, the mean difference between the two groups was 17.9% (Control Group mean 51.9%; CLIL group mean 69.8%).

	Proficiency	Group	Mean	StdErr	Lower	Upper	F Value	P Value
T0	Low	CLIL	29.3%	4.1%	21.9%	38.0%	0.02	0.8989
	Low	Control	28.6%	3.3%	22.6%	35.5%		
T1	Low	CLIL	45.0%	4.9%	35.7%	54.6%	0.06	0.7998
	Low	Control	43.4%	3.8%	36.1%	51.0%		
T2	Low	CLIL	51.3%	4.7%	42.1%	60.4%	0.10	0.7545
	Low	Control	53.2%	3.8%	45.6%	60.6%		
T3	Low	CLIL	69.8%	4.2%	60.9%	77.4%	8.56	0.0037
	Low	Control	51.9%	4.1%	43.9%	59.8%		

Table 5.4 Listening achievement results: Group/Low achievers interaction

Figure 5.2 and Table 5.5 below show the progress of the groups, when the variables proficiency level and group interacted. Although both groups progressed from T0 to T1, there were no statistically significant differences for any of them. From T1 to T2, High achievers in the CLIL group continued to improve whereas High achievers in the Control Group slightly decreased (-3.6%). Nevertheless, both groups, Control and CLIL, progressed significantly from T2 to T3 (Control Group/High achievers: $p=0.0382$ CLIL Group/High achievers $p= 0.0050$). As for the Low achievers, both the CLIL and the Control Groups progressed significantly from T0 to T1 (Control Group/Low achievers $p= 0.0037$; CLIL Group/Low achievers $p= 0.0219$). They continued to progress from T1 to T2, but the differences were not significant. From T2 to T3 the Low achievers in the CLIL group progressed significantly ($p= 0.0042$). As has already been pointed out in the achievement report, the mean percentage of this group (69.8%)

was significantly different and much higher than the mean percentage of the Control Group (51.9%). The CLIL Low achievers' group obtained a result (69.8%) which was relatively close to the result of the Control High achievers (74.9%) and which was much higher than the result obtained by the Control Low achievers' group (51.9%). Nevertheless, the comparison T0-T3 was significant for all the groups: CLIL High ($p=0.0003$), CLIL Low ($p= <.0001$), Control High ($p= 0.0009$) and Control Low ($p= <.0001$).

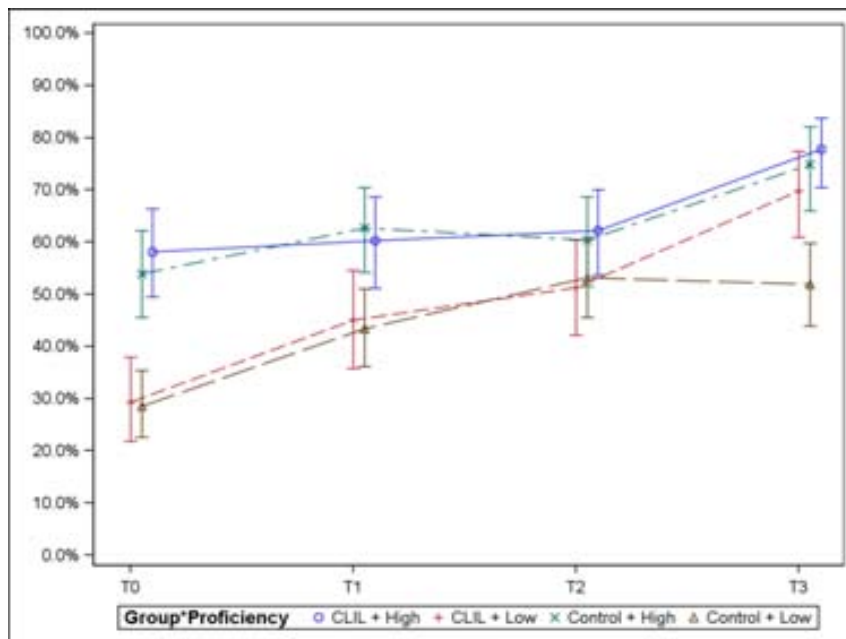


Figure 5.2 Listening improvement results: Group/Proficiency level interaction.

Proficiency	Group	I0_1	I1_2	I2_3	I0-3
Low	CLIL	53.7%	13.9%	36.1%	138.3%
	Control	51.8%	22.4%	-2.4%	81.3%
High	CLIL	3.5%	3.3%	25.1%	33.8%
	Control	16.1%	-3.6%	23.9%	38.8%

Table 5.5 Percentage of improvement in listening. Group/ Proficiency level interaction.

5.1.2 Reading test results

In this section the results for the reading tests will be presented. The general results for reading will be reported first of all, followed by the results for each of the four tests used to assess reading.

The intergroup comparisons carried out using the results of the reading tests administered to the CLIL and Control Groups did not show significant differences in achievement between the groups at any of the times tested. The comparison of the scores obtained at T0 showed that the differences between the CLIL and the Control Groups were not significant ($F= 0.22$, $p= 0.6369$); nor were the differences between the same groups at T1 ($F= 3.75$, $p= 0.0549$), T2 ($F= 0.00$, $p= 0.9446$) and T3 ($F= 0.99$, $p= 0.3216$) significant. Descriptive statistics showed that, although at T2 the scores were almost the same for both groups, there was a slight advantage for the Control Group at the end of the study (see Table 5.6).

	Group	Mean	Std	Lower	Upper	F value	P Value
T0	CLIL	42.5%	3.2%	36.3%	49.0%	0.22	0.6369
	Control	40.5%	2.6%	35.4%	45.8%		
T1	CLIL	51.7%	3.4%	45.1%	58.3%	3.75	0.0549
	Control	60.2%	2.6%	54.8%	65.3%		
T2	CLIL	51.1%	3.3%	44.6%	57.6%	0.00	0.9446
	Control	51.4%	2.8%	46.0%	56.8%		
T3	CLIL	64.0%	3.0%	57.8%	69.8%	0.99	0.3216
	Control	68.0%	2.5%	62.9%	72.8%		

Table 5.6 Reading achievement results

The intragroup comparison showed statistically significant differences in the improvement of both groups from T0 to T1 (CLIL, $p= 0.0002$ / Control $p=<.0001$). In

spite of the fact that both groups showed a decrease between T1 and T2, which was significant for the Control Group ($p < .0001$), the differences from T2 to T3 were statistically significant for both groups (CLIL, $p < .0001$ / Control $p < .0001$). The progress from T0 to T3 was significant for both groups ($p < .0001$); however, the percentage of improvement was higher for the Control group (68%) than it was for the CLIL group (50.7%) (see Table 5.7 and Figure 5.3).

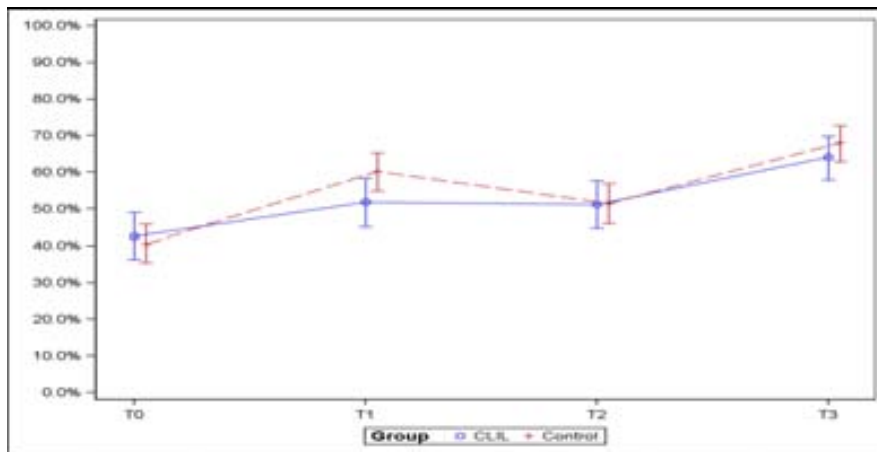


Figure 5.3: Reading improvement results

	I0_1	I1_2	I2_3	I0-3
CLIL	21.8%	-1.2%	25.2%	50.7%
Control	48.6%	-14.5%	32.3%	68.0%

Table 5.7 Percentage of improvement in reading

The interaction Group /High achievers showed that there were no statistically significant differences in reading in favour of High achievers at any of the times tested. Except for T0, when the High achievers in CLIL descriptively outperformed the High level students in the Control Group, the scores at the rest of the times seemed to be in favour of the High level students in the Control Group (see Table 5.8).

	Proficiency	Group	Mean	StdErr	Lower	Upper	F Value	P Value
T0	High	CLIL	56.0%	4.0%	47.9%	63.7%	0.08	0.7810
	High	Control	54.4%	3.9%	46.5%	62.0%		
T1	High	CLIL	62.7%	3.9%	54.7%	70.1%	2.81	0.0962
	High	Control	71.4%	3.3%	64.4%	77.5%		
T2	High	CLIL	53.0%	4.1%	44.9%	60.9%	0.30	0.5826
	High	Control	56.2%	4.0%	48.2%	63.8%		
T3	High	CLIL	72.5%	3.3%	65.4%	78.5%	0.85	0.3571
	High	Control	76.7%	3.1%	70.1%	82.3%		

Table 5.8 Reading achievement results. Group/High achievers interaction

The results of the Low achievers are similar to those of the High achievers in the sense that no statistically significant differences were found in the intergroup comparison when the variable Low proficiency level interacted with the variable Group. Descriptive statistics showed that at the end of the study, the Low achievers in the Control Group outperformed the Low achievers in the CLIL group in terms of reading (see Table 5.9).

	Proficiency	Group	Mean	StdErr	Lower	Upper	F Value	P Value
T0	Low	CLIL	30.1%	3.7%	23.2%	37.9%	0.19	0.6605
	Low	Control	28.0%	2.9%	22.6%	34.0%		
T1	Low	CLIL	40.6%	4.3%	32.5%	49.3%	1.58	0.2104
	Low	Control	47.7%	3.5%	40.8%	54.8%		
T2	Low	CLIL	49.3%	4.4%	40.7%	57.9%	0.21	0.6452
	Low	Control	46.7%	3.5%	39.8%	53.7%		
T3	Low	CLIL	54.6%	4.3%	46.0%	63.0%	0.32	0.5707
	Low	Control	57.8%	3.5%	50.8%	64.6%		

Table 5.9 Reading achievement results. Group/Low achievers interaction

Figure 5.4 and Table 5.10 below show the progress of the groups when the variables Proficiency level (High or Low) and Group interacted with each other. From

T0 to T1 there were no significant differences for the CLIL High achievers ($p = 0.0572$), but there were significant differences for the Control High achievers (P Value $<.0001$). From T1 to T2, both groups presented significant differences, although both groups decreased (CLIL High $p= 0.0018$; Control High $p= <.0001$). However, from T2 to T3 both groups improved significantly (CLIL High $p= <.0001$; Control High $p= <.0001$). As has been shown in the achievement report, the Control High mean score at T3 (76.7%) was higher than the mean score of the CLIL high group (72.5%). The Low achievers in CLIL were the ones that reached the lowest score in reading. All the groups progressed significantly from T0 to T3 ($p= <.0001$).

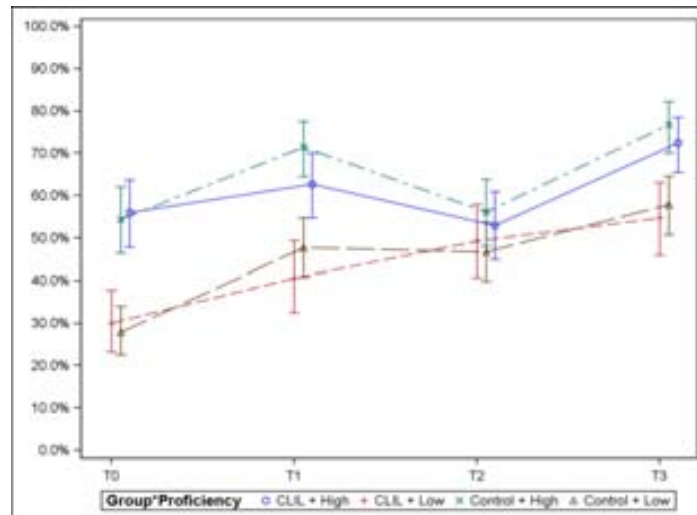


Figure 5.4: Reading improvement results: Group/Proficiency level interaction

Proficiency	Group	I0_1	I1_2	I2_3	I0-3
Low	CLIL	35.2%	21.3%	10.9%	81.8%
	Control	70.7%	-2.2%	23.9%	106.8%
High	CLIL	12.0%	-15.5%	36.8%	29.5%
	Control	31.3%	-21.4%	36.6%	41.1%

Table 5.10 Percentage of improvement in Reading. Group/Proficiency level interaction

5.1.2.1 Reading question 1

In Reading Question 1 the students were asked to read and match a definition with the corresponding word. The intergroup comparison scores obtained at T0 showed that, despite the slightly higher percentage obtained by the Control Group (53.2%), the difference between the CLIL and the Control Groups was not statistically significant ($F= 0.04$, $p= 0.8354$), which indicates that both groups started the study roughly with the same proficiency level as far as the Reading Question 1 was concerned; the differences between the same groups at T1 ($F= 0.62$, $p= 0.4327$) and T2 ($F= 0.45$, $p= 0.5010$) were not significant either, although descriptive statistics showed that the CLIL group slightly outperformed the Control Group at T2. The results at T3 did not show a statistically significant advantage for any of the groups in this question (see Table 5.11).

	Group	Mean	StdErr	Lower	Upper	F Value	P Value
T0	CLIL	51.8%	5.3%	41.4%	62.0%	0.04	0.8354
	Control	53.2%	4.4%	44.6%	61.7%		
T1	CLIL	70.4%	4.7%	60.4%	78.8%	0.62	0.4327
	Control	75.0%	3.5%	67.5%	81.3%		
T2	CLIL	60.3%	5.0%	50.1%	69.7%	0.45	0.5010
	Control	55.7%	4.4%	46.9%	64.2%		
T3	CLIL	76.9%	3.9%	68.4%	83.6%	0.28	0.5967
	Control	79.6%	3.4%	72.2%	85.5%		

Table 5.11 Reading achievement results: Reading Question1

The intragroup comparison showed that both groups significantly progressed from T0 to T1 (CLIL $p= 0.0004$, Control G $p=<.0001$). However, from T1 to T2, both groups decreased but this decrease was only significant for the Control Group ($p=<.0001$). From T2 to T3, the CLIL and the Control Groups significantly progressed

again (CLIL $p = 0.0002$, Control G $p = <.0001$). The Control Group progressed 42.9% whereas the progress of the CLIL Group was 27.4%. The improvement from T0 to T3 was significant for both groups ($p = <.0001$) (see Figure 5.5 and Table 5.12).

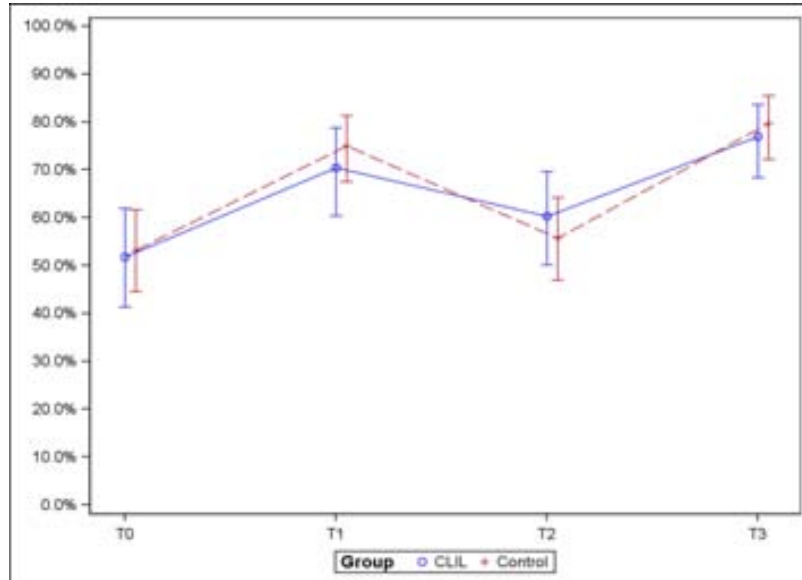


Figure 5.5 Reading improvement results: Question 1

	10_1	11_2	12_3	10_3
CLIL	36.0%	-14.3%	27.4%	45.8%
Control	41.0%	-25.8%	42.9%	49.6%

Table 5.12 Percentage of improvement: Reading Question1

5.1.2.2 Reading question 2

Reading Question 2 was a True/False reading comprehension task. The students were asked to read five statements and to write Yes or No considering what they could see in the picture in the task. The intergroup analysis at T0 showed no statistically significant differences between the CLIL and Control Groups. Although descriptively the CLIL group seemed to be slightly better at time T0 and T2, at the end of the study, the Control Group attained slightly better mean percentages (see Table 5.13).

	Group	Mean	StdErr	Lower	Upper	F Value	P Value
T0	CLIL	75.9%	3.1%	69.2%	81.5%	3.82	0.0516
	Control	67.2%	2.9%	61.2%	72.7%		
T1	CLIL	79.2%	3.2%	72.1%	84.8%	0.46	0.4962
	Control	81.9%	2.4%	76.8%	86.1%		
T2	CLIL	59.1%	3.7%	51.6%	66.1%	0.10	0.7578
	Control	57.5%	3.2%	51.2%	63.6%		
T3	CLIL	67.6%	3.5%	60.4%	74.1%	0.12	0.7317
	Control	69.3%	3.2%	62.6%	75.2%		

Table 5.13 Reading achievement results: Reading Question 2

In terms of progress, significant differences were established between T1 and T2 for both groups and from T2 to T3 only for the Control Group. This group showed a higher improvement, although not statistically significant from T0 to T1 than the CLIL group. Both groups decreased from T1 to T2 and then they both improved again from T2 to T3, although, as has been said, improvement was only significant in the case of the Control Group ($p= 0.0320$). The progress made by both groups from T0 to T3 was not significant for any of them (see Figure 5.6 and Table 5.14).

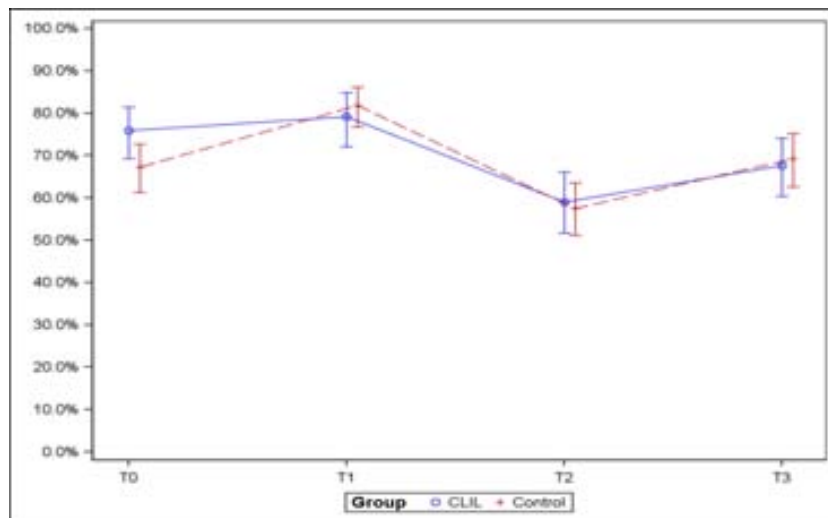


Figure 5.6 Reading improvement results: Question 2

	I0_1	I1_2	I2_3	I0_3
CLIL	4.3%	-25.4%	14.4%	-10.9%
Control	21.9%	-29.8%	20.4%	3.1%

Table 5.14 Percentage of improvement: Reading Question 2

5.1.2.3 Reading question 3

Reading Question 3 was a gap filling exercise consisting of a text and some gaps with some missing words. Next to it, there were some small pictures and words. Children had to decide which word to use and copy it into the gap. This exercise was a cloze type test designed to test the students' use of English.

As can be seen in Table 5.15 below, there were no statistically significant differences at any of the times tested for any of the groups. Descriptive statistics showed an advantage for the Control Group at T0 and at T1. The results at T2 were almost the same for both groups; at T3, the mean percentage result was again slightly higher for the Control Group.

	Group	Mean	StdErr	Lower	Upper	F Value	P Value
T0	CLIL	37.7%	5.3%	28.0%	48.4%	1.60	0.2062
	Control	46.5%	4.3%	38.2%	55.0%		
T1	CLIL	46.2%	5.7%	35.5%	57.3%	2.04	0.1541
	Control	56.6%	4.3%	48.0%	64.8%		
T2	CLIL	50.0%	5.2%	39.9%	60.0%	0.00	0.9476
	Control	49.5%	4.5%	40.8%	58.3%		
T3	CLIL	62.1%	4.9%	52.0%	71.3%	0.88	0.3483
	Control	68.4%	4.4%	59.3%	76.4%		

Table 5.15 Reading achievement results: Reading Question 3

The intragroup comparison showed a steady progression for the CLIL group, although the results for this group were not statistically significant at any of the time periods tested. The Control Group showed a very similar increase to that of the CLIL group from T0 to T1; the progression decreased from T1 to T2, but the results of the group improved significantly from T2 to T3 ($p= 0.0039$). the progress made by the groups from T0 to T3 was statistically significant: CLIL group ($p= 0.0007$), Control Group ($p= 0.0004$). Descriptively, the average achievement level of the Control Group was higher than that of the CLIL group in this particular question (see Figure 5.7 and Table 5.16).

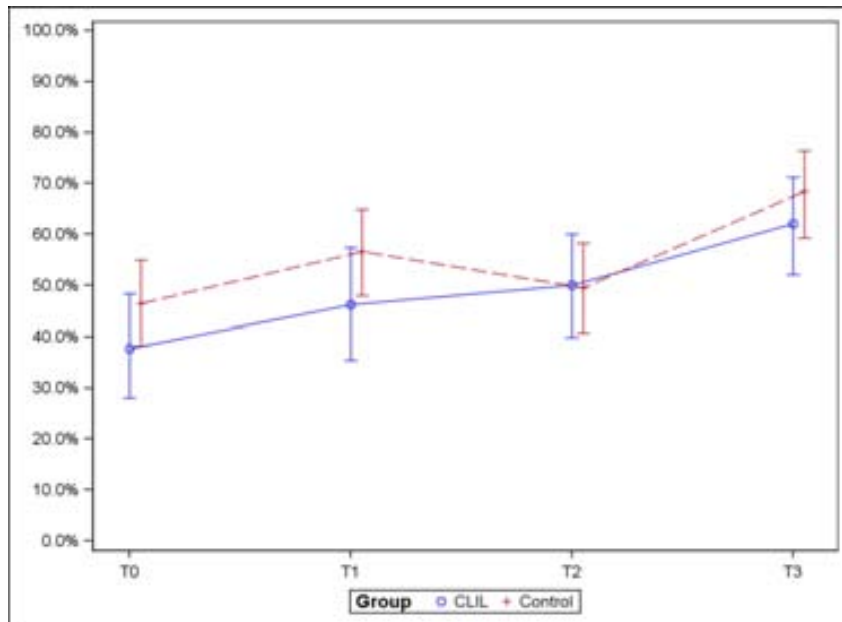


Figure 5.7 Reading improvement results: Question 3

	I0_1	I1_2	I2_3	I0_3
CLIL	22.7%	8.1%	24.3%	65.0%
Control	21.6%	-12.5%	38.3%	47.2%

Table 5.16 Percentage of improvement: Reading Question 3

5.1.2.4 Reading question 4 (Comprehension)

This test was a story in three parts. Each part of the story had a matching picture. After each part, the students had to complete 10 sentences about the story using one, two or three words. This task was scored under two different categories: comprehension and accuracy. This section will deal with the results in comprehension and the next section will address the results in accuracy.

As shown in Table 5.17, no significant differences were found in the intergroup comparisons between the CLIL and the Control Group in this particular task. Descriptive statistics showed, however, a slight advantage for the Control Group at all times tested, except for T0 at which the CLIL group performed slightly better. Although T1 showed the highest mean difference between the groups (11.1%), the CLIL group almost caught up with the Control Group at T3: the mean difference between the groups was 3%

	Group	Mean	StdErr	Lower	Upper	F Value	P Value
T0	CLIL	31.4%	4.7%	23.0%	41.3%	0.33	0.5681
	Control	28.0%	3.6%	21.4%	35.7%		
T1	CLIL	41.4%	5.3%	31.5%	52.1%	2.49	0.1162
	Control	52.5%	4.4%	44.0%	61.0%		
T2	CLIL	47.3%	5.1%	37.4%	57.4%	0.09	0.7638
	Control	49.4%	4.4%	40.9%	57.9%		
T3	CLIL	61.3%	4.9%	51.3%	70.3%	0.22	0.6390
	Control	64.4%	4.3%	55.5%	72.3%		

Table 5.17 Reading achievement results: Question 4 (Comprehension)

The intragroup comparisons revealed statistically significant differences in the improvement of the Control Group from T0 to T1 ($p = <.0001$), as well as from T2 to T3, for both groups (Control Group ($p = 0.0021$) and CLIL group ($p = 0.0094$)). Although the Control Group progressed more than the CLIL group from T0 to T1, its progression

from T1 to T2 decreased; the progression made from T2 to T3 was almost the same as the progression made by the CLIL group at the same time period. Both groups improved significantly throughout the study, T0 to T3: Control ($p = <.0001$), CLIL ($p = <.0001$) (see Figure 5.8 and Table 5.18).

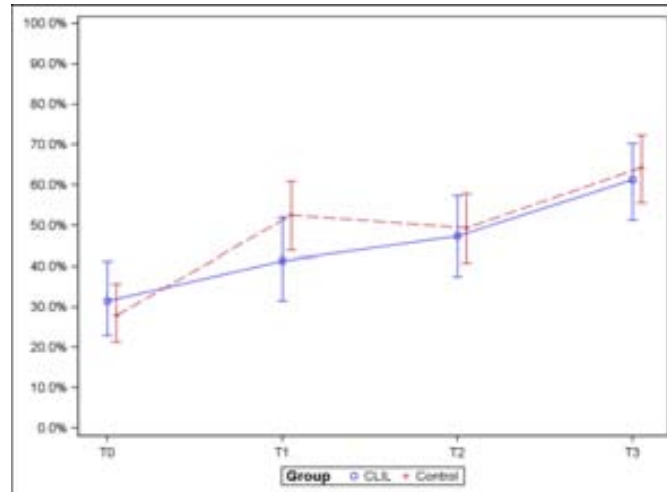


Figure 5.8: Reading improvement results: Question 4 (Comprehension)

	I0_1	I1_2	I2_3	I0_3
CLIL	31.8%	14.2%	29.5%	95.0%
Control	87.7%	-6.0%	30.4%	130.0%

Table 5.18 Percentage of improvement: Reading Question 4 (Comprehension)

5.1.2.5 Reading question 4 (Accuracy)

As has been mentioned in the previous section, the answers provided by the students in the Reading Question 4 were analysed for accuracy, which means that only accurate answers were taken into consideration. Table 5.19 showed that no significant differences were found at any of the times tested. Although the CLIL group performed better at T0 (7%), descriptive statistics showed that the mean results of the Control Group at T1 were also 7% higher than those of the CLIL group. Nevertheless, at T2 the difference between the two groups was only 2%. This difference increased again at T3, although not significantly, to almost 6% in favour of the Control Group.

	Group	Mean	StdErr	Lower	Upper	F Value	P Value
T0	CLIL	18.0%	3.3%	12.4%	25.4%	3.38	0.0667
	Control	11.0%	2.0%	7.6%	15.7%		
T1	CLIL	27.2%	4.3%	19.6%	36.3%	1.44	0.2315
	Control	34.2%	3.8%	27.2%	42.0%		
T2	CLIL	40.0%	4.7%	31.1%	49.6%	0.11	0.7395
	Control	42.1%	4.0%	34.4%	50.2%		
T3	CLIL	52.4%	4.9%	42.8%	61.8%	0.77	0.3800
	Control	58.3%	4.3%	49.7%	66.4%		

Table 5.19 Reading achievement results: Reading Question 4 (Accuracy)

The intragroup analyses showed that from T0 to T1 there were statistically significant differences for the Control Group ($p = <.0001$) but no differences for the CLIL Group ($p = 0.0897$). From T1 to T2 only the CLIL Group improved significantly ($p = 0.0228$). From T2 to T3 the progress was significantly different for both groups (CLIL group $p = 0.0297$, Control Group $p = 0.0007$). The progress T0 to T3 was also significant for both groups (CLIL group $p = <.0001$, Control Group $p = <.0001$) (see Figure 5.9 and Table 5.20).

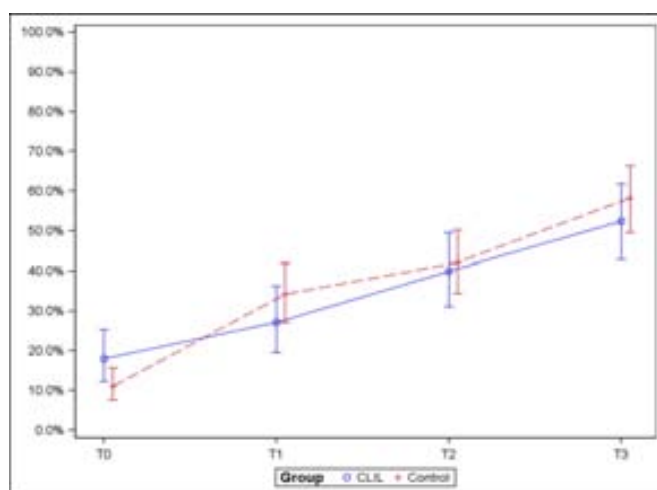


Figure 5.9 Reading improvement results. Question 4 (Accuracy)

	I0_1	I1_2	I2_3	I0_3
CLIL	51.1%	47.2%	31.1%	191.6%
Control	210.7%	23.0%	38.3%	428.0%

Table 5.20 Percentage of improvement: Reading Question 4 (Accuracy)

5.1.3 Summary of Science Listening and Reading Results

5.1.3.1 Summary of Listening and Reading Achievement Results.

Table 5.21 below shows the listening and reading results of the intergroup comparison between the Control and the CLIL Science group.

	Listening (total)		Reading (Total)	
	Control	CLIL	Control	CLIL
T0	40.7%	43.1%	40.5%	42.5%
	p= 0.5979		p= 0.6369	
T1	53.2%	52.7%	60.2%	51.7%
	p= 0.9197		p= 0.0549	
T2	56.8%	56.8%	51.4%	51.1%
	p= 0.9995		p= 0.9446	
T3	64.2%	74%	68.0%	64.0%
	p= 0.0289		p= 0.3216	

Table 5.21 Summary of Listening and Reading achievement results

Table 5.22 presents the achievement results for each of the questions used to measure reading.

Reading	Question1		Question2		Question 3		Question 4 (Comprehension)		Question 4 (Accuracy)	
	Control	CLIL	Control	CLIL	Control	CLIL	Control	CLIL	Control	CLIL
T0	53.2%	51.8%	67.2%	75.9%	46.5%	37.7%	28.0%	31.4%	11.0%	18.0%
	p= 0.8354		p= 0.0516		p= 0.2062		p= 0.5681		p= 0.0667	
T1	75.0%	70.4%	81.9%	79.2%	56.6%	46.2%	52.5%	41.4%	34.2%	27.2%
	p= 0.4327		p= 0.4962		p= 0.1541		p= 0.1162		p= 0.2315	
T2	55.7%	60.3%	57.5%	59.1%	49.5%	50.0%	49.4%	47.3%	42.1%	40.0%
	p= 0.5010		p= 0.7578		p= 0.9476		p= 0.7638		p= 0.7395	
T3	79.6%	76.9%	69.3%	67.6%	68.4%	62.1%	64.4%	61.3%	58.3%	52.4%
	p= 0.5967		p= 0.7317		p= 0.3583		p= 0.6390		p= 0.3800	

Table 5.22 Summary of Reading achievement results: individual reading questions

Table 5.23 shows a summary of the effects of the students' initial proficiency level on the results obtained in the listening test.

Listening (Total)	Group/Proficiency level			
	High		Low	
	Control	CLIL	Control	CLIL
T0	54%	58.2%	28.6%	29.3%
	p= 0.4966		p= 0.8989	
T1	62.7%	60.2%	43.4%	45.0%
	p= 0.6926		p= 0.7998	
T2	60.4%	62.2%	53.2%	60.4%
	p= 0.7695		p= 0.7545	
T3	74.9%	77.8%	51.9%	69.8%
	p= 0.5862		p= 0.0037	

Table 5.23 Summary of Listening achievement results. Group/Proficiency level interaction.

Table 5.24 below summarises the reading achievement results when the interaction Group/Proficiency level was taken into account.

Reading (Total)	Group/Proficiency level			
	High		Low	
	Control	CLIL	Control	CLIL
T0	54.4%	56.0%	28%	30.1%
	p= 0.7810		p= 0.6605	
T1	71.4%	62.7%	47.7%	40.06%
	p= 0.0962		p= 0.0214	
T2	56.2%	53.0%	46.7%	49.3%
	p= 0.5826		p= 0.6452	
T3	76.7%	72.5%	57.8%	54.6%
	p= 0.3571		p= 0.5707	

Table 5.24 Summary of Reading achievement results. Group/Proficiency level interaction.

5.1.3.2 Summary of Listening and Reading Improvement Results.

Table 5.25 below shows the improvement in Listening and Reading of the students exposed to EFL and CLIL Science.

	Listening		Reading (Total)	
	Control	CLIL	Control	CLIL
T0-T1	30.7% p=0.0028	22.1% p= 0.1370	48.6% p=<.0001	21.8% p= 0.0002
T1-T2	6.9% p=0.7346	7.9% p= 0.7720	-15.5% p=<.0001	-1.2% p= 0.9923
T2-T3	13.0% p=0.2194	30.2% p=<.0001	32.3% p=<.0001	25.2% p=<.0001
T0-T3	57.9% p=<.0001	71.5% p=<.0001	68.0% p=<.0001	50.7% p=<.0001

Table 5.25 Summary of Listening and Reading improvement results

The reading improvement results of each reading question is displayed in Tables 5.26 and 5.27.

Reading	Question1		Question2		Question 3	
	Control	CLIL	Control	CLIL	Control	CLIL
T0-T1	41% p=<.0001	36% p=<.0004	21.9% p=0.0003	4.3% p=0.8651	21.6% p=0.1989	22.7% p=0.5610
T1-T2	-25.8% p=<.0001	-14.3% p=0.1025	-29.8% p=<.0001	-25.8% p=0.0002	-12.5% p=0.5297	8.1% p=0.9380
T2-T3	42.9% p=<.0001	27.4% p=0.0002	20.4% p=0.0320	14.4% p=0.2684	38.3% p=0.0039	24.3% p=0.1757
T0-T3	49.6% p=<.0001	54.8% p=<.0001	3.1% p=0.9576	-10.9% p=0.2286	47.2% p=0.0004	65.0% p=0.0007

Table 5.26 Summary of Reading improvement results: Reading Questions 1, 2, 3

Reading	Question 4 (Comprehension)		Question 4 (Accuracy)	
	Control	CLIL	Control	CLIL
T0-T1	87.7% p= <.0001	31.8% p=0.1520	210.7% p=<.0001	51.1% p=0.0897
T1-T2	-6.0% p=0.8547	14.2 p=0.6074	23% p=0.1565	47.2% p=0.0228
T2-T3	30.4% p=0.0021	29.5% p=0.0094	38.3% p=0.0007	31.1% p=0.0297
T0-T3	130.0% p=<.0001	95.0% p=<.0001	428.8% p=<.0001	191.6% p=<.0001

Table 5.27 Summary of Reading improvement results: Reading Question 4

The intragroup comparison that was carried out in order to see the progress in the listening test of the Control and the CLIL groups when the variable Proficiency level was considered for the analysis is displayed in Table 5.28.

Listening	Group/Proficiency level			
	High		Low	
	Control	CLIL	Control	CLIL
T0-T1	16.1% p= 0.3089	3.5% p= 0.9800	51.8% p= 0.0037	53.7% p= 0.0219
T1-T2	-3.6% p= 0.9717	3.3% p= 0.9803	22.4% p= 0.1447	13.9% p= 0.6896
T2-T3	23.9% p= 0.0382	25.1% p= 0.0050	-2.4% p= 0.9929	36.1% p= 0.0042
T0-T3	38.8% p= 0.0009	33.8% p= 0.0003	81.3% p= <.0001	138.3% p= <.0001

Table 5.28 Summary of Listening improvement results. Group/Proficiency level interaction

The results of the interaction Group/Proficiency level, as far as reading is concerned, are shown in Table 5.29 below.

Reading	Group/Proficiency level			
	High		Low	
	Control	CLIL	Control	CLIL
T0-T1	31.3% p= <.0001	12.0% p= 0.0572	70.7% p= <.0001	35.2% p= 0.0006
T1-T2	-21.4% p= <.0001	-15.5% p= 0.0018	-2.2% p= 0.9674	21.3% p= 0.0137
T2-T3	36.6% p= <.0001	36.8% p= <.0001	23.9% p= <.0001	10.9% p= 0.2182
T0-T3	41.1% p= <.0001	29.5% p= <.0001	106.8% p= <.0001	81.8% p= <.0001

Table 5.29 Summary of Reading improvement results. Group/Proficiency level interaction.

5.1.4 Writing

In this section a description of the results obtained by CLIL Science students compared to those of their Control Group will be presented in order to answer the following question and subquestions: RQ2: Keeping the number of hours of exposure the same for both groups, CLIL and non-CLIL, do the CLIL students' writing skills measured in terms of complexity, accuracy and fluency benefit from their exposure to the CLIL experience? RQ 2.1 In terms of fluency, are there differences in achievement between CLIL and EFL learners significantly in favour of CLIL learners at different times (T0, T1, T2, T3)? RQ 2.2 In terms of fluency, are there differences in progress between CLIL and EFL learners significantly in favour of CLIL learners after one year (T0-T1) and two years (T2-T3) of CLIL implementation? Are there differences in favour of CLIL students when we consider their progress from T0-T3? RQ 2.3 In terms of accuracy, are there differences in achievement between CLIL and EFL learners significantly in favour of CLIL learners at different times (T0, T1, T2, T3)? RQ 2.4 In terms of accuracy, are there differences in their progress between CLIL and EFL learners significantly in favour of CLIL learners after one year (T0-T1) and two years (T2-T3) of CLIL implementation? Are there differences in favour of CLIL students

when we consider their progress from T0-T3? RQ 2.5 In terms of complexity (lexical and syntactic), are there differences in achievement between CLIL and EFL learners significantly in favour of CLIL learners at different times (T0, T1, T2, T3)? RQ 2.6 In terms of complexity (lexical and syntactic), are there differences in achievement between CLIL and EFL learners significantly in favour of CLIL learners after one year (T0-T1), two years (T2-T3) of CLIL implementation? Are there differences in favour of CLIL students when we consider their progress from T0-T3? RQ 2.7 How does the initial level of English proficiency affect the students' performance in writing in the CLIL and the EFL groups?

5.1.4.1 Fluency

This section presents the results obtained by the Control and the CLIL groups in the area of Fluency. Four different measures were used to analyse Fluency: Total Number of Words (TNW), Total Number of Words in English (TNWE), Total Number of Units (TNU), and the ratio between the Total Number of Words in English and the Total Number of Words (TNWE/TNW). For the TNWE and the TNU, the results of the interaction between Group/and Proficiency Level (High and Low achievers) will also be reported.

5.1.4.1.1 Total Number of Words

Table 5.30 shows the results in achievement for each group regarding the Total Number of Words written by the students at different times. The results show that, although both groups started at very similar points, students in the Control Group significantly outperformed their CLIL peers at T1 ($F= 13.33$, $p= 0.0003$). At T2, the level of achievement seemed to be higher for the Control Group, although the results were not significantly different. At T3 the average number of words written by CLIL

group was slightly higher than that of the Control Group. However, the difference were not significant either.

	Group	Mean	StdErr	Lower	Upper	F Value	P value
T0	CLIL	63.55	6.47	50.80	76.29	0.02	0.8748
	Control	64.87	5.12	54.78	74.96		
T1	CLIL	65.70	6.47	52.95	78.44	13.33	0.0003
	Control	96.35	5.12	86.26	106.4		
T2	CLIL	88.88	6.47	76.13	101.6	1.35	0.2468
	Control	98.63	5.12	88.53	108.7		
T3	CLIL	121.4	6.47	108.6	134.1	2.70	0.1018
	Control	107.6	5.13	97.46	117.7		

Table 5.30 Fluency achievement results: TNW

The intragroup comparison did not show a steady linear increase for either the CLIL or the Control Groups. The Control Group improved significantly from T0 to T1 ($p = <.0001$). It slightly progressed (2.4%) from T1 to T2 and it continued to grow (9.1%) from T2 to T3 but the progress was not significant. The CLIL group, however, improved significantly from T1-T2 ($p = 0.0048$) and from T2-T3 ($p = <.0001$) (see Table 5.31 and Figure 5.10).

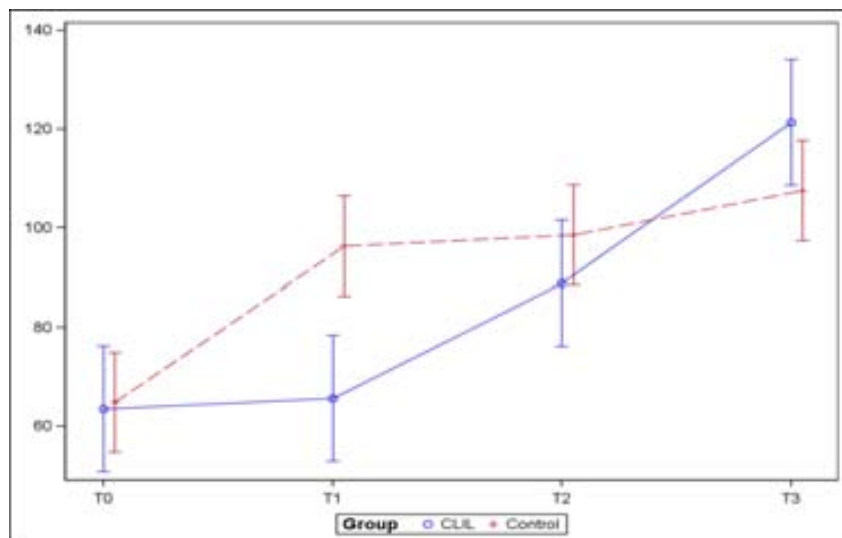


Figure 5.10: Fluency improvement results: TNW

	I0_1	I1_2	I2_3	I0_3
CLIL	3.4%	35.3%	36.6%	91.0%
Control	48.5%	2.4%	9.1%	65.8%

Table 5.31 Percentage of improvement in Fluency: TNW

5.1.4.1.2 Total Number of Words in English

As for the TNWE, no significant differences were found at T0 between the two groups. The results showed that the Control Group at T1 significantly outperformed the CLIL group ($F= 5.47$ $p=0.0203$). Even though descriptive statistics at T3 revealed that the CLIL group wrote, on average, more words in English (110) than the Control Group (99.38), these results were not statistically significantly different (see Table 5.32).

	Group	Mean	StdErr	Lower	Upper	F value	P Value
T0	CLIL	51.16	6.20	38.93	63.39	0.47	0.4934
	Control	45.64	4.91	35.95	55.32		
T1	CLIL	56.33	6.20	44.10	68.56	5.47	0.0203
	Control	75.16	4.91	65.47	84.84		
T2	CLIL	77.22	6.20	64.99	89.45	0.63	0.4286
	Control	83.61	4.91	73.92	93.30		
T3	CLIL	110.2	6.20	97.94	122.4	1.79	0.1821
	Control	99.38	4.92	89.68	109.1		

Table 5.32 Fluency achievement results: TNWE

The intragroup comparison showed that, from T0 to T1, the increase for the Control Group (64.7%) was higher than that of the CLIL group (10.1%); the results from T1-T2 were only significant for the CLIL group ($p= 0.0038$). However, both groups progressed significantly from T2 to T3 (CLIL $p= <.0001$, Control $p= 0.0066$). The CLIL group increased from 77.2 to 110.2, whereas the mean number of words in

English at T3 for the Control Group is 99.38. Nevertheless, the progress from T0 to T3 was significant for both groups ($p = <.0001$).

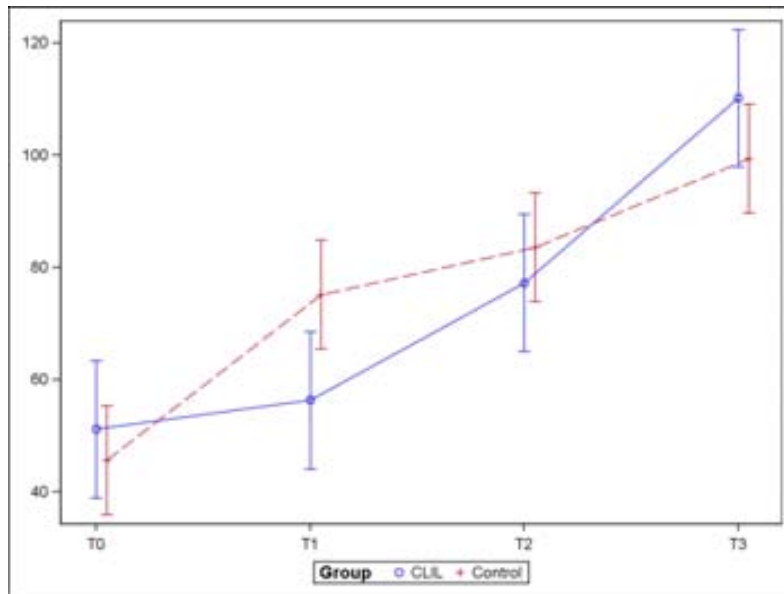


Figure 5.11 Fluency improvement results: TNWE

	I0_1	I1_2	I2_3	I0_3
CLIL	10.1%	37.1%	42.7%	115.3%
Control	64.7%	11.2%	18.9%	117.8%

Table 5.33 Percentage of improvement in Fluency: TNWE

The interaction Group/High achievers in terms of TNWE can be seen in Table 5.34 below. High achievers in the Control group significantly outperformed High achievers in the CLIL group at T1 ($F = 4.44$ $p = 0.0364$). Even though the results at the rest of the times were not significant, at T2 the High achievers in both groups obtained almost the same mean results. At T3, the High achievers in the CLIL group wrote, on average, more words (122.3) than the High achievers in the Control Group.

	Proficiency	Group	Mean	StdErr	Lower	Upper	F Value	P Value
T0	High	CLIL	56.79	7.79	41.41	72.16	0.56	0.4557
	High	Control	48.76	7.19	34.59	62.93		

	Proficiency	Group	Mean	StdErr	Lower	Upper	F Value	P Value
T1	High	CLIL	63.01	7.79	47.64	78.38	4.44	0.0364
	High	Control	85.63	7.19	71.46	99.80		
T2	High	CLIL	91.92	7.79	76.55	107.3	0.00	0.9825
	High	Control	91.68	7.19	77.51	105.9		
T3	High	CLIL	123.3	7.80	107.9	138.7	1.05	0.3067
	High	Control	112.3	7.19	98.13	126.5		

Table 5.34 Fluency achievement results: TNWE. Group/High achievers interaction

Table 5.35 shows the achievement results of the interaction Group/Low achievers in terms of TNWE. As can be seen, there were no significant results at any of the times tested for any of the groups. Descriptive statistics, however, showed that although at T0 of data collection the Low achievers in the CLIL group performed better than their low peers in the Control Group, at T1 and T2, the Control Group achieved better mean scores. At the end of the two years, at T3, the Low achievers in the CLIL group wrote, on average, more words (97.02) in English than the Low level students in the Control Group (86.45).

	Proficiency	Group	Mean	StdErr	Lower	Upper	F Value	P Value
T0	Low	CLIL	45.54	8.12	29.53	61.55	0.08	0.7716
	Low	Control	42.51	6.38	29.92	55.10		
T1	Low	CLIL	49.64	8.12	33.63	65.65	2.09	0.1499
	Low	Control	64.69	6.38	52.10	77.28		
T2	Low	CLIL	62.53	8.12	46.52	78.54	1.56	0.2129
	Low	Control	75.53	6.38	62.95	88.12		
T3	Low	CLIL	97.02	8.12	81.01	113.0	1.05	0.3067
	Low	Control	86.45	6.44	73.75	99.16		

Table 5.35 Fluency achievement results: TNWE. Group/Low achievers interaction

The improvement results when the interaction Group /Proficiency level was taken into account can be seen in Figure 5.12 and Table 5.36 below. From T0 to T1, the

improvement was significant for the Control Groups (Control High $p = <.0001$, Control Low $p = 0.0026$). From T1 to T2, it was only significant for the CLIL High ($p = 0.0011$). The progress during the final time period tested, T2 to T3, was significant for three of the groups: CLIL High $p = 0.0003$, Control High $p = 0.0194$, CLIL Low $p = 0.0001$. The results of the analysis T0 to T3 showed significant results for all the groups ($p = <.0001$).

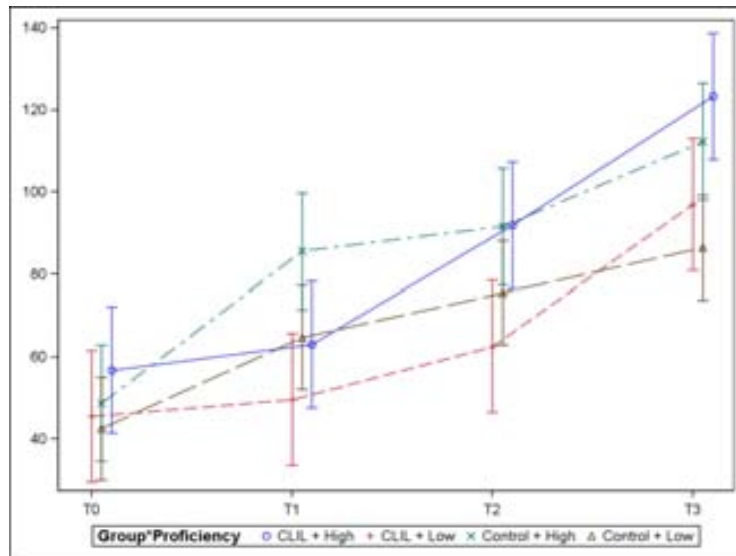


Figure 5.12 Fluency improvement results: TNWE. Group/Proficiency level interaction

Proficiency	Group	I0_1	I1_2	I2_3	I0_3
Low	CLIL	9.0%	26.0%	55.2%	113.0%
	Control	52.2%	16.8%	14.5%	103.4%
High	CLIL	11.0%	45.9%	34.2%	117.2%
	Control	75.6%	7.1%	22.5%	130.3%

Table 5.36 Percentage of improvement in Fluency: TNWE. Group/Proficiency level interaction

5.1.4.1.3 Ratio: Total Number of Words in English/Total Number of Words

As a measure of fluency, the ratio between the Total Number of Words in English and the Total Number of Words was also calculated. The results at T0 were significantly different in favour of the CLIL group ($F = 8.29$ $p = 0.0042$). At T1, after the students had already been exposed to CLIL, the ratio was also significant in favour of the CLIL group ($F = 6.82$ $p = 0.0095$). However, although descriptive statistics showed a

very similar achievement level for the two groups at T3, there were no statistically significant differences between the groups. In fact, the mean percentage difference was only 1%.

	Group	Mean	StdErr	Lower	Upper	F value	P value
T0	CLIL	78.1%	3.4%	71.8%	85.1%	8.29	0.0042
	Control	66.3%	2.3%	61.9%	71.1%		
T1	CLIL	84.7%	3.4%	78.3%	91.6%	6.82	0.0095
	Control	74.4%	2.1%	70.4%	78.7%		
T2	CLIL	84.4%	3.1%	78.6%	90.7%	0.42	0.5179
	Control	81.9%	2.2%	77.7%	86.4%		
T3	CLIL	89.4%	2.7%	84.2%	95.0%	0.07	0.7886
	Control	90.4%	2.2%	86.1%	94.9%		

Table 5.37 Fluency achievement results: % TNWE/TNW

The percentage of increase in the progression of the two groups showed a different pattern for each group, but the final results were very similar. From T0 to T1 there were no statistically significant differences for the CLIL group ($p= 0.4361$), but there were significant differences for the Control Group ($p= 0.0257$). The results between T1 and T2 were only significant for the Control Group ($p= 0.0262$). There was no significant increase for the CLIL group during this time period. When we considered T2 to T3, the results were significant for the Control Group ($p= 0.0094$). Both groups improved significantly from T0 to T3: Control Group $p= <.0001$, CLIL group $p= 0.0219$. Nevertheless, as has already been reported in the achievement results, the difference in the mean percentage at T3 between the CLIL and the Control Groups was only 1 %.

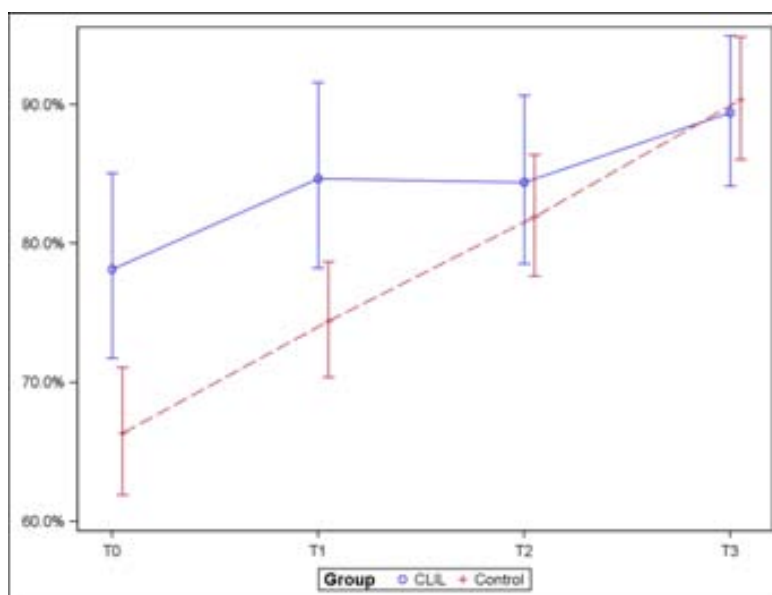


Figure 5.13 Fluency improvement results: %TNWE/TNW

	I0_1	I1_2	I2_3	I0_3
CLIL	8.3%	-0.3%	5.9%	14.4%
Control	12.2%	10.1%	10.3%	36.2%

Table 5.38 Percentage of improvement in Fluency: % TNWE/TNW

5.1.4.1.4 Total Number of Units

As was pointed out in Chapter 4, the concept of ‘unit’ in this dissertation refers to any ‘meaningful chunk of language that contains, at least, one finite or non-finite verb. Table 5.39 below describes the results of the intergroup analyses as for the TNU. The table shows no significant differences between the Control and the CLIL groups at the different times tested. Descriptive statistics, however, revealed that the achievement results for the Control Group at T1 and T2 were higher than those of the CLIL Group. Nevertheless, at T3, the CLIL group outperformed the Control Group. The CLIL group wrote an average of 17.68 units per essay whereas the Control Group wrote 14.95.

	Group	Mean	StdErr	Lower	Upper	F Value	P Value
T0	CLIL	9.81	1.08	7.69	11.94	2.07	0.1519
	Control	7.80	0.85	6.11	9.48		
T1	CLIL	8.86	1.08	6.73	10.98	3.01	0.0841
	Control	11.29	0.85	9.60	12.97		
T2	CLIL	12.99	1.08	10.86	15.12	1.63	0.2027
	Control	14.78	0.85	13.09	16.46		
T3	CLIL	17.68	1.08	15.55	19.81	3.81	0.0524
	Control	14.95	0.86	13.26	16.63		

Table 5.39 Fluency achievement results: TNU

In terms of progression, the students' scores as for the TNU increased for both groups during different time periods. The Control Group progressed significantly from T0 to T1 ($p= 0.0004$) and from T1 to T2 ($p= 0.0004$), but there were no statistically significant differences from T2 and T3. The progress of the CLIL group showed no statistically significant differences in the first year, from T0 to T1 ($p= 0.8162$). However, between T1 and T2, the CLIL group progressed significantly ($p= 0.0010$) and continued to improve from T2 to T3 ($p= 0.0001$). The progress T0 to T3 was significant for the Control Group ($p=<.0001$) as well as for the CLIL group ($p=<.0001$) (see Figure 5.14 and Table 5.40 below).

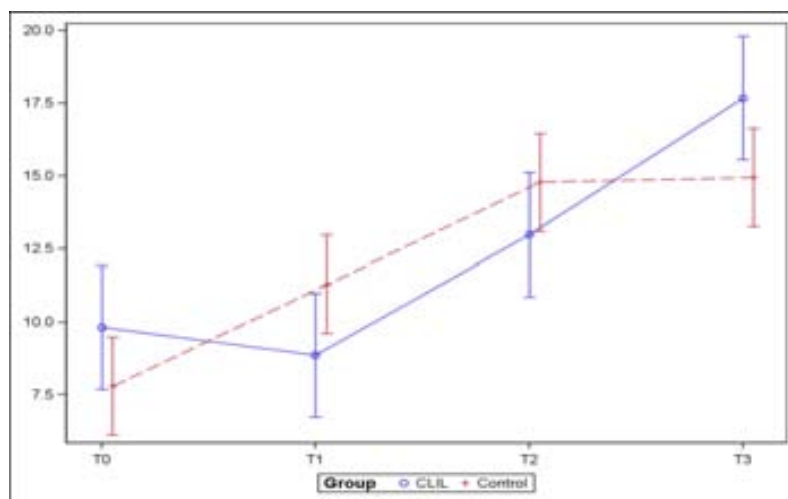


Figure 5.14 Fluency improvement results: TNU

	I0_1	I1_2	I2_3	I0_3
CLIL	-9.7%	46.7%	36.1%	80.2%
Control	44.7%	30.9%	1.1%	91.6%

Table 5.40 Percentage of improvement in Fluency: TNU

The achievement results of the interaction Group/High achievers as for the TNU can be seen in Table 5.41 below. There were no statistically significant differences for High achievers when the interaction between Group/Proficiency level was taken into account. Descriptive statistics show that, at T0, the mean score for the CLIL group was higher than that of the Control Group. Nevertheless, at T2 and T3, the mean score was higher for the Control Group. However, at the end of the study, the mean number of units was higher for the High achievers in the CLIL group (19.33) compared to the mean of the High achievers in the Control Group (16.02).

	Proficiency	Group	Mean	StdErr	Lower	Upper	F Value	P Value
T0	High	CLIL	10.55	1.36	7.88	13.22	1.33	0.2501
	High	Control	8.40	1.25	5.93	10.86		
T1	High	CLIL	9.68	1.36	7.00	12.35	2.45	0.1193
	High	Control	12.60	1.25	10.13	15.06		
T2	High	CLIL	15.41	1.36	12.74	18.08	0.07	0.7927
	High	Control	15.90	1.25	13.44	18.37		
T3	High	CLIL	19.33	1.36	16.66	22.01	3.07	0.0815
	High	Control	16.06	1.25	13.60	18.53		

Table 5.41 Fluency achievement results: TNU. Group/High achievers interaction

As with High achievers, no statistically significant differences were found when the interaction Group/Low achievers was taken into account. Although at T0, the Low achievers in the CLIL group performed better in terms of the TNU, at T1 and T2 the mean number of units was in favour of the Control Group. But, at the end of the study (T3), the average number of units written by the CLIL Group (16.02) was higher than

the number of units written by the Control Group (13.83). It is also interesting to notice that at T3 Low achievers in the CLIL group obtained almost the same results (16.02) as the High achievers in the Control Group (16.6) (see Table 5.42).

	Proficiency	Group	Mean	StdErr	Lower	Upper	F Value	P Value
T0	Low	CLIL	9.07	1.41	6.29	11.86	1.07	0.3018
	Low	Control	7.20	1.11	5.01	9.39		
T1	Low	CLIL	8.04	1.41	5.26	10.82	1.15	0.2851
	Low	Control	9.98	1.11	7.79	12.17		
T2	Low	CLIL	10.57	1.41	7.78	13.35	2.91	0.0896
	Low	Control	13.65	1.11	11.46	15.84		
T3	Low	CLIL	16.02	1.41	13.24	18.81	1.46	0.2278
	Low	Control	13.83	1.12	11.62	16.04		

Table 5.42 Fluency achievement results: TNWE. Group/Low achievers interaction

The progress of High and Low achievers in terms of the TNU can be seen in Figure 5.15 and Table 5.43 below. High and Low achievers in the CLIL group do not show any significant progress from T0 to T1. In fact, their results decrease but their decrease is not significant. High and Low achievers in the Control Groups progress but the improvement was only significant in the case of High achievers ($p= 0.0054$). From T1 to T2 all students improve and this improvement is significant for most of the groups: Control High (P Value 0.0452); CLIL High ($p= 0.0002$) and Control Low ($p= 0.0064$), except for the CLIL Low ($p= 0.2877$). The progress in the final time period is only significant for the CLIL groups: High ($p= 0.0230$) and Low ($p= 0.0009$). This particular group, the CLIL Low group is the one that decreased the most in the first time period but, as has been pointed out in the achievement report, their mean percentage of units increases substantially and their results are Higher in the end than those of the Control Low group. In fact, results obtained at T3 (16.02) are almost the same as those

obtained by the Control High achievers at the same time (16.06). The progress T0 to T3 was significant for all the groups ($p < .0001$).

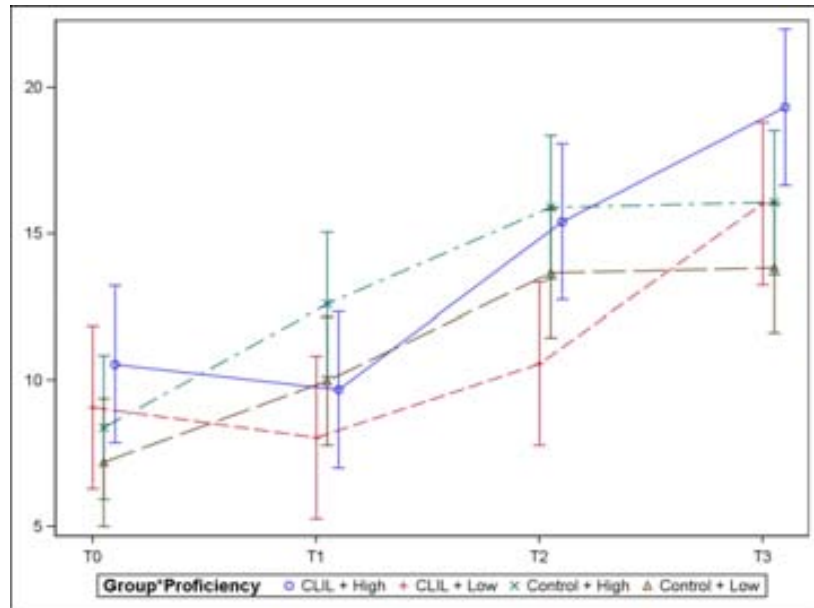


Figure 5.15 Fluency improvement results: TNU. Group/Proficiency level interaction

Proficiency	Group	I0_1	I1_2	I2_3	I0_3
Low	CLIL	-11.4%	31.4%	51.7%	76.6%
	Control	38.6%	36.8%	1.3%	92.1%
High	CLIL	-8.3%	59.3%	25.5%	83.3%
	Control	50.0%	26.3%	1.0%	91.3%

Table 5.43 Percentage of improvement in Fluency: TNU. Group/Proficiency level interaction

5.1.4.2. Accuracy

The results presented in this section show the differences between the Control and the CLIL groups in the area of Accuracy. Two different measures were used to analyse Accuracy: Total Number of Error Free Units (TNEFU) and the ratio between the Total Number of Error Free Units in relation to the Total Number of Units (TNEFU/TNU). As for the TNEFU, the results of the interaction Group/Proficiency level will also be presented.

5.1.4.2.1 Total Number of Error Free Units

The comparison of the scores obtained at T0 shows no statistically significant differences between the groups at T0, which indicates that both groups started the study roughly with the same proficiency level as for the number of error free units. At T1 both groups obtained roughly the same scores, and at T2 the Control Group slightly outperformed the CLIL one. Although there are no statistically significant differences, the descriptive results at T3 show a slightly higher advantage in the mean scores for the number of Error Free Units in favour of the CLIL group (see Table 5.44)

	Group	Mean	StdErr	Lower	Upper	F Value	P Value
T0	CLIL	1.88	0.40	1.24	2.85	1.15	0.2840
	Control	2.49	0.38	1.84	3.37		
T1	CLIL	3.07	0.58	2.12	4.44	0.15	0.6978
	Control	3.37	0.48	2.53	4.47		
T2	CLIL	4.67	0.80	3.32	6.56	0.35	0.5556
	Control	5.31	0.69	4.11	6.86		
T3	CLIL	7.35	1.13	5.43	9.96	0.99	0.3229
	Control	6.02	0.76	4.70	7.72		

Table 5.44 Accuracy achievement results: TNEFU

In terms of progression, none of the groups improved significantly from T0 to T1. The Control Group increased significantly from T1 to T2 ($p= 0.0022$). However, the progression from T2 to T3 was only significant for the CLIL group ($p= 0.0087$). The improvement from T0 to T3 was significant for both groups ($p= <.0001$) (see Figure 5.16 and Table 5.45 below).

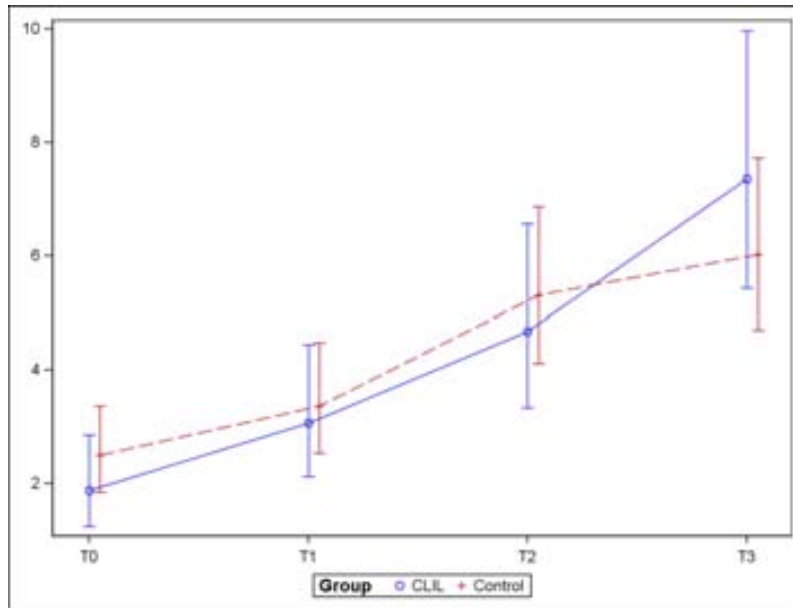


Figure 5.16 Accuracy improvement results: TNEFU

	I0_1	I1_2	I2_3	I0_3
CLIL	63.2%	52.3%	57.4%	291.4%
Control	35.0%	57.8%	13.4%	141.8%

Table 5.45 Percentage of improvement in accuracy: TNEFU

The results of the intergroup comparison when the variable High achievers was taken into account showed no significant differences. Descriptively, the High achievers in the Control Group obtained better mean scores at T0, T1 and T2; however, at T3, the mean score (9.09) was higher for the CLIL group (see Table 5.46)

	Proficiency	Group	Mean	StdErr	Lower	Upper	F Value	P Value
T0	High	CLIL	1.81	0.47	1.08	3.03	2.41	0.1219
	High	Control	3.04	0.61	2.04	4.52		
T1	High	CLIL	4.01	0.89	2.59	6.21	0.20	0.6540
	High	Control	4.58	0.86	3.15	6.64		
T2	High	CLIL	6.34	1.29	4.24	9.46	0.17	0.6842
	High	Control	7.08	1.24	5.00	10.01		
T3	High	CLIL	9.09	1.70	6.26	13.18	0.52	0.4736
	High	Control	7.55	1.30	5.36	10.63		

Table 5.46 Accuracy improvement results: TNEFU. Group/High achievers interaction

As in the case of Low achievers, no significant differences were found. Descriptively, even though the Control group performed better at T0, T1 and T2, the CLIL group attained better mean scores (5.95) at T3 than the scores obtained by the Control Group (4.81).

	Proficiency	Group	Mean	StdErr	Lower	Upper	F Value	P Value
T0	Low	CLIL	1.95	0.51	1.17	3.26	0.02	0.8831
	Low	Control	2.05	0.40	1.39	3.02		
T1	Low	CLIL	2.34	0.57	1.45	3.79	0.03	0.8604
	Low	Control	2.47	0.47	1.71	3.59		
T2	Low	CLIL	3.44	0.77	2.22	5.34	0.27	0.6024
	Low	Control	3.99	0.68	2.85	5.59		
T3	Low	CLIL	5.95	1.20	3.99	8.86	0.66	0.4177
	Low	Control	4.81	0.80	3.46	6.67		

Table 5.47 Accuracy improvement results: TNEFU. Group/Low achievers interaction

Table 5.48 and Figure 5.17 show the progress of High and Low achievers in terms of TNEFU. From T0 to T1, only the High achievers in CLIL progressed significantly ($p= 0.0117$). During the second time period, significant differences were found for the Control High ($p=0.0116$) and the Control Low ($p=0.0210$). The CLIL High and the CLIL Low did not show significant differences during this time period. From T2 to T3, only the CLIL Low improved significantly ($p= 0.0146$). The progress throughout the study, T0 to T3, was significant for all the groups ($p= <.0001$). The groups with the highest improvement percentage were the CLIL High and the CLIL Low groups. The lowest improvement percentage was for the Control Low group. As has already been shown, the Control High, which started with the highest mean as for TNEFU obtained a lower mean (7.55) than the CLIL high (9.09) at the end of the study.

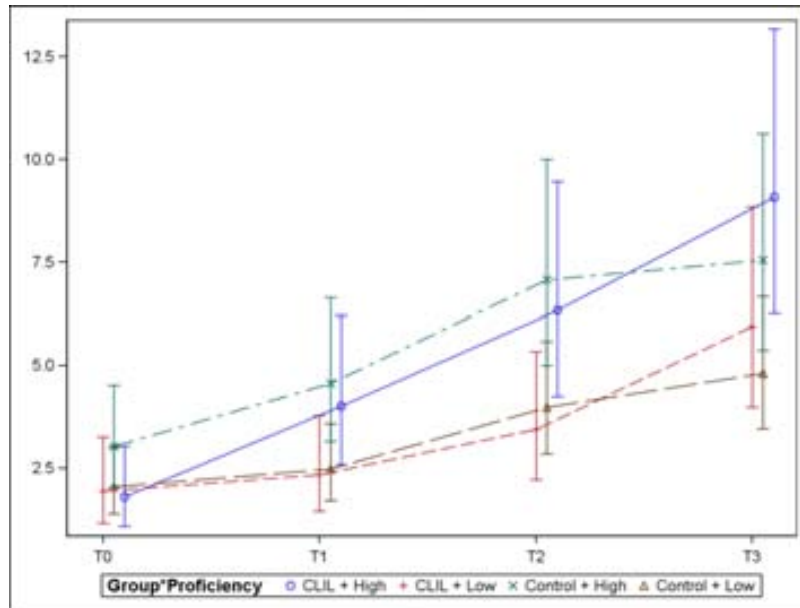


Figure 5.17 Accuracy improvement results: TNEFU. Group Proficiency level interaction

Proficiency	Group	I0_1	I1_2	I2_3	I0_3
Low	CLIL	20.2%	46.9%	72.8%	205.1%
	Control	21.0%	61.1%	20.6%	135.0%
High	CLIL	121.6%	58.0%	43.4%	402.0%
	Control	50.7%	54.6%	6.7%	148.8%

Table 5.48 Percentage of improvement results: TNEFU. Group /Proficiency level interaction

5.1.4.2.2 Ratio: Total Number of Error Free Units/Total Number of Units

As can be seen in Table 5.49, the ratio TNEFU/TNU showed statistically significant differences at T0 in favour of the Control Group ($p= 0.0098$). No other significant differences in achievement were found for the rest of the times tested. The achievement percentages of the CLIL and the Control Groups at T2, even though they were not significant, were almost the same (41.2 and 41.6%) but, at the end of the project, at T3, the Control Group obtained a slightly higher percentage (46.3%) in the TNEFU compared to that of the CLIL group (44.3%). Even though the mean difference in percentages at T0 was around 14%, the difference at the end of the study (T3), was only 2%. The CLIL group seemed to catch up with the Control Group.

	Group	Mean	StdErr	Lower	Upper	F Value	P Value
T0	CLIL	21.1%	3.4%	15.4%	28.9%	6.75	0.0098
	Control	35.2%	4.0%	28.2%	44.1%		
T1	CLIL	37.6%	5.1%	28.8%	49.2%	0.17	0.6807
	Control	35.0%	3.7%	28.5%	43.0%		
T2	CLIL	41.2%	5.0%	32.4%	52.3%	0.00	0.9447
	Control	41.6%	3.6%	35.0%	49.4%		
T3	CLIL	44.3%	4.5%	36.2%	54.3%	0.10	0.7508
	Control	46.3%	3.9%	39.2%	54.5%		

Table 5.49 Accuracy achievement results: % TNWE/TNW

However, when we analysed the progress made by these same students, it is interesting to note that from T0-T1 the CLIL group improved substantially (78%) and significantly ($p= 0.0098$). From T1 to T2, and from T2 to T3, the CLIL group continued to improve, but improvement was not significant. The progress made by the Control Group decreased from T0 to T1, but then it improved from T1 to T2 and T2 to T3. In spite of the different patterns of improvement, the level of achievement was almost the same for both groups at T3, with a slight advantage (2%) for the Control Group. The CLIL group improved significantly from T0 to T3 ($p= <.0001$). The Control Group improved, but the progress was not significant ($p= 0.0843$)

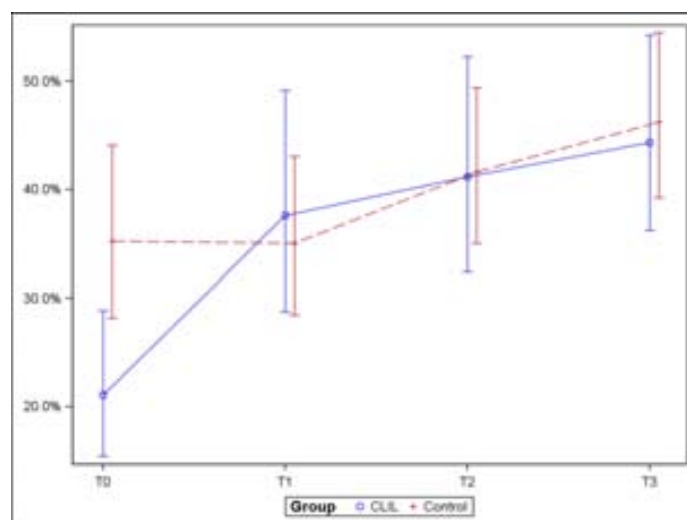


Figure 5.18 Accuracy improvement results: % TNEFU/TNU

	I0_1	I1_2	I2_3	I0_3
CLIL	78.2%	9.5%	7.7%	110.1%
Control	-0.6%	18.8%	11.2%	31.2%

Table 5.50 Percentage of improvement in accuracy: % TNEFU/TNU

As is clear from Table 5.51 below, the interaction Group/Proficiency level revealed that at T0 the score for the High achievers in the CLIL group was significantly Lower (18.5% $p=0.0019$) than that of the Control Group (39.8%). At T1, however, although the mean score was very similar, the CLIL group slightly outperformed the Control one, but the differences were not statistically significant. At T2 and T3, even though the scores were slightly higher for the Control Group (3%), they were statistically significant.

	Proficiency	Group	Mean	StdErr	Lower	Upper	F Value	P Value
T0	High	CLIL	18.5%	3.6%	12.6%	27.2%	9.81	0.0019
	High	Control	39.8%	5.7%	30.0%	52.7%		
T1	High	CLIL	43.9%	7.0%	32.1%	60.0%	0.02	0.8944
	High	Control	42.7%	5.4%	33.3%	54.9%		
T2	High	CLIL	45.5%	6.2%	34.8%	59.7%	0.29	0.5896
	High	Control	50.2%	5.7%	40.0%	63.0%		
T3	High	CLIL	49.2%	6.0%	38.7%	62.6%	0.20	0.6583
	High	Control	53.0%	5.9%	42.5%	66.0%		

Table 5.51 Accuracy achievement results: %TNEFU/TNU. Group/High achievers interaction

In terms of Low achievers, there were no significant differences between the CLIL and the Control Groups. Even though at T0 the mean percentage of the Control Group was higher (31.2%) than that of the CLIL group (24.1%), at T2 and T3, the CLIL group seemed to catch up with the score of the Control Group. At T3, the ratio was almost exactly the same for both groups (see Table 5.52).

	Proficiency	Group	Mean	StdErr	Lower	Upper	F Value	P Value
T0	Low	CLIL	24.1%	4.6%	16.5%	35.2%	1.14	0.2860
	Low	Control	31.2%	4.6%	23.3%	41.8%		
T1	Low	CLIL	32.2%	5.6%	22.9%	45.4%	0.27	0.6057
	Low	Control	28.7%	4.0%	21.9%	37.7%		
T2	Low	CLIL	37.2%	5.9%	27.3%	50.7%	0.15	0.6964
	Low	Control	34.5%	3.9%	27.5%	43.2%		
T3	Low	CLIL	40.0%	5.4%	30.6%	52.2%	0.00	0.9496
	Low	Control	40.4%	4.5%	32.5%	50.2%		

Table 5.52 Accuracy achievement results: %TNEFU/TNU. Interaction Group/Low achievers

The differences in improvement between High and Low achievers in the CLIL and the Control Groups can be seen in Figure 5.19 and Table 5.53 below. High achievers in CLIL experienced a marked significant increase from T0 to T1 ($p=0.0006$). Although from T1 to T2 and from T2 to T3 they continued to progress, their progress was not significantly different (T1-T2 $p=0.9964$, T2-T3 $p=0.9388$). Their counterparts in the Control Group improved during the three times tested, but their improvement was not significant at any of the times (T0-T1 $p=0.9581$; T1-T2 $p=0.5635$; T2-T3 $p=0.9594$). As for Low level students, they all progressed, even though their progress was not significant at any of the times. From T0 to T3, all groups progressed, but the CLIL High achievers and the CLIL Low achievers were the groups that progressed the most. As has been pointed out in the achievement report, the mean percentages of Low achievers in the CLIL and the Control Groups was almost the same at T3.

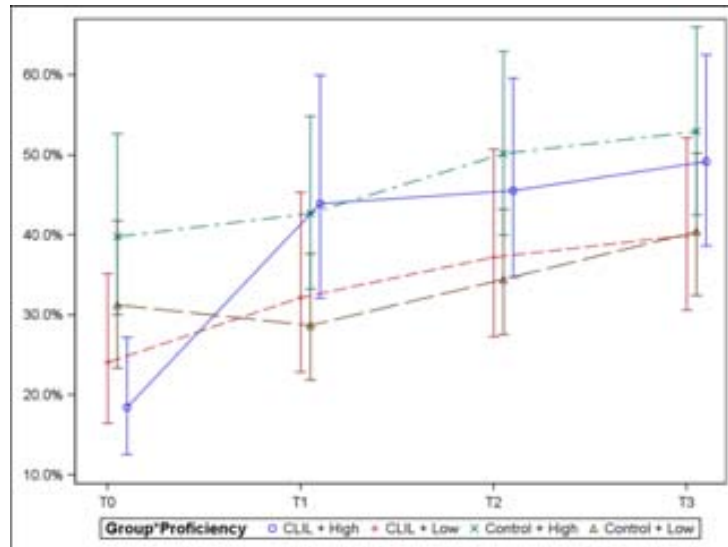


Figure 5.19 Accuracy improvement results: %TNEFU/TNU. Group/Proficiency level interaction

Proficiency	Group	I0_1	I1_2	I2_3	I0_3
Low	CLIL	33.7%	15.6%	7.4%	66.0%
	Control	-8.1%	20.2%	17.1%	29.4%
High	CLIL	137.5%	3.7%	8.0%	166.0%
	Control	7.4%	17.4%	5.5%	33.1%

Table 5.53 Percentage of improvement in accuracy: %TNEFU/TNU. Group/Proficiency level

5.1.4.3 Complexity

5.1.4.3.1 Lexical Complexity

Two measures were taken to determine the lexical complexity of the CLIL and EFL writings: The Total Number of Adjectives (TNAdj) and the Total Number of Lexical Verbs (TNLV). The ratios between the TNLV in relation to the TNWE and the ratio between the TNAdj in relation to the TNWE will also be presented. As for the TNAdj and the TNLV the differences between High and Low achievers will be reported as well.

5.1.4.3.1.1 Total Number of Lexical Verbs

The comparison of the scores obtained shows that there are no statistically significant differences at any of the times tested in terms of achievement. Descriptive statistics show almost no differences at T0 between the groups. However, in terms of the mean scores obtained by the CLIL group at T1 (2.52), that is, at the end of the first year, and at T3 (4.04), the CLIL group shows a slight advantage over the results of the Control Group.

	Group	Mean	StdErr	Lower	Upper	F Value	P Value
T0	CLIL	1.48	0.27	1.03	2.13	0.04	0.8472
	Control	1.55	0.22	1.18	2.04		
T1	CLIL	2.52	0.40	1.84	3.44	2.00	0.1598
	Control	1.88	0.25	1.45	2.44		
T2	CLIL	2.59	0.41	1.89	3.55	0.44	0.5079
	Control	2.97	0.36	2.34	3.77		
T3	CLIL	4.04	0.58	3.03	5.38	0.89	0.3468
	Control	3.37	0.40	2.67	4.27		

Table 5.54 Lexical Complexity achievement results: TNLV

The intergroup comparison carried out to determine the progression of each of the groups showed a higher statistically significant progression for the CLIL Group from T0 to T1 ($p= 0.0079$). The progression from T1 to T2 was higher and statistically significant for the Control Group ($p= 0.0002$). The Control Group attained a slightly higher mean score (2.97) over the CLIL group (2.59) at the end of this time period. From T2 to T3, however, the improvement was statistically significant for the CLIL group ($p= 0.002$) which improved substantially to reach a mean score of 4.04, higher than that obtained by the Control Group (3.37) (see Figure 5.20 and Table 5.55).

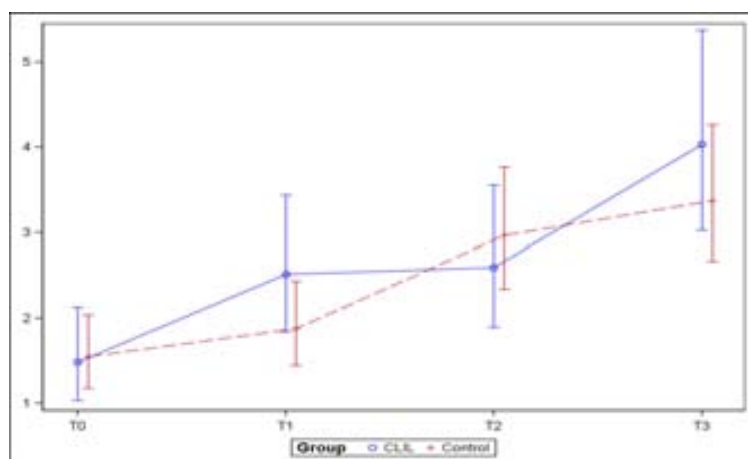


Figure 5.20 Lexical Complexity improvement results: TNLV

	I0_1	I1_2	I2_3
CLIL	69.6%	3.1%	55.5%
Control	20.9%	58.2%	13.7%

 Table 5.55 Percentage of improvement in Lexical Complexity: %TNEFU/TNU.
 Group/Proficiency level interaction

The intergroup comparison when the interaction Group/Proficiency level was taken into account revealed that there were no significant differences for any of the groups. Table 5.56 shows that CLIL High achievers at the end of the study outperformed the High achievers in the Control Group in the mean number of lexical verbs used.

	Proficiency	Group	Mean	StdErr	Lower	Upper	F Value	P Value
T0	High	CLIL	2.19	0.46	1.45	3.32	0.00	0.9608
	High	Control	2.16	0.40	1.51	3.10		
T1	High	CLIL	3.41	0.64	2.34	4.96	1.90	0.1710
	High	Control	2.37	0.42	1.67	3.38		
T2	High	CLIL	3.43	0.65	2.36	4.99	0.28	0.5997
	High	Control	3.92	0.65	2.82	5.45		
T3	High	CLIL	5.07	0.90	3.57	7.21	0.67	0.4142

Proficiency	Group	Mean	StdErr	Lower	Upper	F Value	P Value
High	Control	4.15	0.68	2.99	5.76		

Table 5.56 Lexical Complexity achievement results: TNLV. Group/High achievers

As in the case of High achievers, no significant differences can be found in the case of Low achievers. However, the mean number of lexical verbs used by CLIL Low achievers (3.21) was higher than the mean number of lexical verbs used by their counterparts in the Control Group (2.74) (see Table 5.57).

	Proficiency	Group	Mean	StdErr	Lower	Upper	F Value	P Value
T0	Low	CLIL	1.00	0.24	0.62	1.62	0.11	0.7366
	Low	Control	1.11	0.21	0.77	1.61		
T1	Low	CLIL	1.86	0.39	1.23	2.82	0.68	0.4117
	Low	Control	1.48	0.26	1.05	2.09		
T2	Low	CLIL	1.96	0.41	1.30	2.97	0.26	0.6133
	Low	Control	2.25	0.36	1.63	3.09		
T3	Low	CLIL	3.21	0.62	2.19	4.71	0.40	0.5291
	Low	Control	2.74	0.43	2.01	3.74		

Table 5.57 Lexical Complexity achievement results: TNLV. Group/Low achievers

Figure 5.21 below describes the progress of High and Low achievers as for the TNLV. CLIL High achievers improved significantly from T2 to T3 ($p= 0.0253$), whereas the High achievers in the Control Group improved significantly from T1 to T2 ($p= 0.0012$). As has already been pointed out, the High achievers in CLIL obtained the highest mean score in the Total Number of Lexical Verbs (5.07). As for the CLIL Low achievers, their progress was significant during the first time period, T0 to T1 ($p= 0.0281$), as well as during the final time period, T2 to T3 ($p= 0.0148$). The Low achievers in the Control Group only progressed significantly from T1 to T2 ($p= 0.0273$). All groups improved significantly from T0 to T3 ($p= <.0001$). The groups that progress

the most were the CLIL groups. In spite of the progress, the group with the lowest achievement score at T3 was the Control Low achievers group.

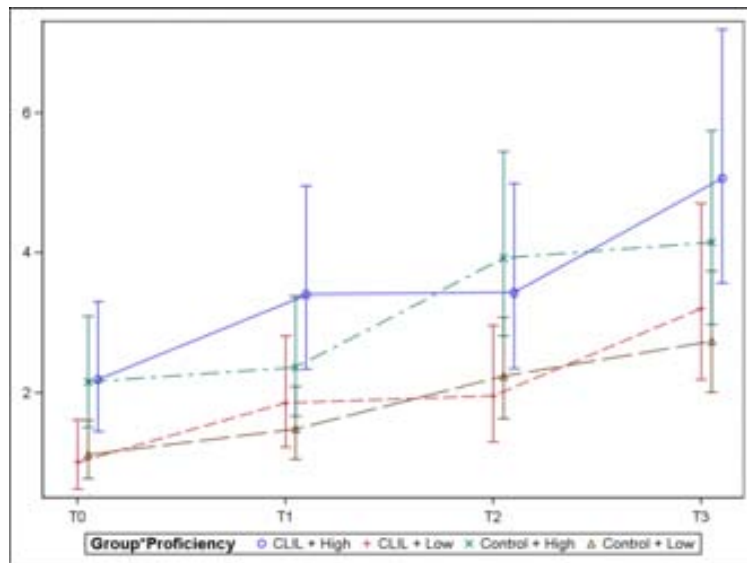


Figure 5.21 Lexical Complexity improvement results: TNLV. Group/Proficiency level interaction

Proficiency	Group	I0_1	I1_2	I2_3	I0_3
Low	CLIL	85.0%	5.6%	63.6%	219.5%
	Control	33.1%	51.4%	22.1%	146.1%
High	CLIL	55.5%	0.6%	47.8%	131.3%
	Control	9.8%	65.2%	5.8%	92.0%

Table 5.58 Percentage of improvement in Lexical Complexity: TNLV. Group/Proficiency level interaction

5.1.4.3.1.2 Ratio: Total Number of Lexical Verbs/ Total Number of words in English

As can be seen in Table 5.59, the achievement in the ratio TNLV/TNWE was only significant at T1 in favour of the CLIL group ($F= 10.83$ $p= 0.0011$). The Control Group started at an average mean of 4% (T0) and finished exactly at 4% (T3). Descriptive statistics show that the Control Group mean level of achievement did not

show many changes and stayed at around 4%. However, in spite of the significant level of achievement at T1, the mean value for the ratio TNLV/TNWE for both groups at T3 was around 4%.

	Group	Mean	StdErr	Lower	Upper	F Value	P Value
T0	CLIL	3.2%	0.5%	2.3%	4.5%	1.04	0.3092
	Control	4.0%	0.5%	3.2%	5.0%		
T1	CLIL	5.1%	0.6%	4.0%	6.5%	10.83	0.0011
	Control	3.0%	0.3%	2.4%	3.7%		
T2	CLIL	3.9%	0.5%	3.0%	5.0%	0.39	0.5312
	Control	4.3%	0.4%	3.6%	5.1%		
T3	CLIL	4.1%	0.4%	3.3%	5.0%	0.02	0.9006
	Control	4.0%	0.4%	3.4%	4.8%		

Table 5.59 Lexical Complexity achievement results: % TNLV/TNWE

The progress of the two groups showed different patterns which were not statistically significant at any of the time periods tested. From T0 to T1 the CLIL group progressed, but from T1 to T2 its progression decreased; however, from T2 to T3 it slightly increased again, although the improvement was not statistically significant. The Control Group decreased from T0 to T1, but then it improved again from T1 to T2 to a level which was very similar to the initial one (4.3%), and then it slightly decreased again. The level of achievement at T3 was the same as at T1 (4%). The progress throughout the study, T0 to T3, was not significant for any of the groups.

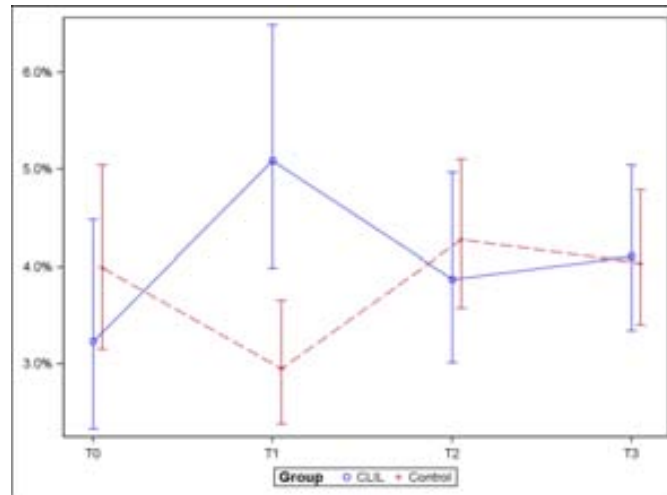


Figure 5.22 Lexical Complexity improvement results: %TNLV/TNWE

	I0_1	I1_2	I2_3	I0_3
CLIL	57.5%	-23.9%	6.1%	27.1%
Control	-26.1%	44.8%	-5.5%	1.1%

Table 5.60 Percentage of improvement in Lexical Complexity: %TNLV/TNWE.

5.1.4.3.1. 3 Total Number of Adjectives

As can be seen in Table 5.61, there were statistically significant differences at T1 and T3 of data collection. The differences at T1 were favourable to the Control Group ($p= 0.0029$). Even though the mean score for the CLIL group was slightly higher at T0, the Control Group more than doubled the mean score for adjectives at T1 (1.65 vs. 0.69). At T2, the differences were still descriptively favourable for the Control Group. However, at T3, the CLIL group significantly outperformed the Control Group ($p= 0.0212$).

	Group	Mean	StdErr	Lower	Upper	F Value	P value
T0	CLIL	1.11	0.23	0.74	1.66	0.79	0.3756
	Control	0.99	0.17	0.70	1.40		
T1	CLIL	0.69	0.17	0.42	1.13	9.01	0.0029
	Control	1.65	0.23	1.26	2.16		
T2	CLIL	1.80	0.31	1.29	2.52	0.79	0.3756

	Group	Mean	StdErr	Lower	Upper	F Value	P value
T3	Control	2.18	0.27	1.71	2.78	5.40	0.0212
	CLIL	3.36	0.45	2.57	4.38		
	Control	2.18	0.27	1.70	2.78		

Table 5.61 Lexical Complexity achievement results: TNAdj

Figure 5.23 shows the different patterns in the progress made by the two groups in terms of the TNAdj. From T0 to T1 the Control Group progressed significantly ($p=0.0456$). From T1 to T2 it was the CLIL group that progressed the most and its improvement was statistically significant ($p=0.0024$). From T2 to T3 only the CLIL group continued to progress significantly ($p=0.0016$). As has already been said, at T3, the mean score (3.36) was higher for the CLIL group than the mean score for the Control Group (2.18). Both groups progressed significantly from t0 to T3: Control Group ($p=0.0002$), CLIL group ($p<.0001$).

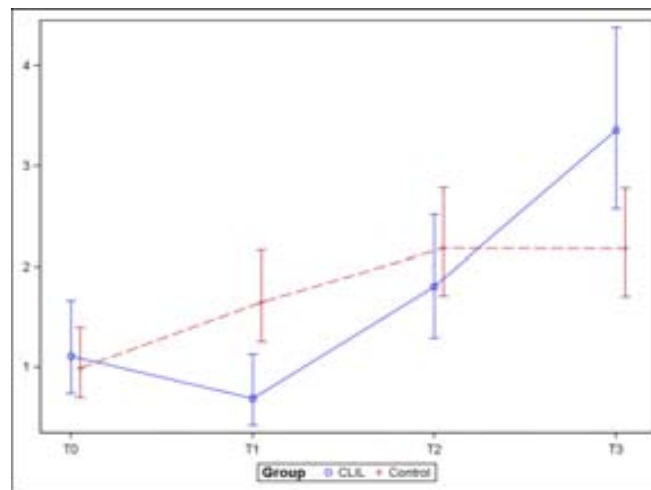


Figure 5.23 Lexical Complexity improvement results: TNAdj

	I0_1	I1_2	I2_3	I0_3
CLIL	-37.8%	161.4%	86.1%	202.3%
Control	66.5%	32.4%	-0.2%	199.9%

Table 5.62 Percentage of improvement in Lexical Complexity: TNAdj

Table 5.63 below shows the results of the intergroup comparison when the interaction Group/High achievers was taken into account. Except for T1, at which the mean number of adjectives was higher for the Control Group, the rest of the times, High achievers in the CLIL group used, on average, more adjectives than High achievers in the Control Group. Although at T0 and T2 there were no significant differences between the CLIL and the Control Groups, at T3, the differences were statistically significant in favour of the High achievers in the CLIL group ($p= 0.0255$).

	Proficiency	Group	Mean	StdErr	Lower	Upper	F Value	P Value
T0	High	CLIL	1.35	0.32	0.86	2.14	2.41	0.1217
	High	Control	0.80	0.19	0.51	1.28		
T1	High	CLIL	0.90	0.25	0.53	1.54	3.34	0.0685
	High	Control	1.67	0.32	1.15	2.43		
T2	High	CLIL	2.72	0.50	1.88	3.92	0.53	0.4654
	High	Control	2.25	0.39	1.60	3.16		
T3	High	CLIL	3.76	0.62	2.72	5.21	5.07	0.0255
	High	Control	2.18	0.38	1.55	3.07		

Table 5.63 Lexical Complexity achievement results: TNAdj. Group/High achievers interaction

As can be seen in Table 5.64, the interaction between Low achievers in the CLIL and the Control Groups showed statistically significant differences at T1 ($p= 0.0030$) and T2 ($p= 0.0492$) in favour of the Low achievers in the CLIL group. Even though the differences at T3 were not statistically significant, the mean number of adjectives was higher for the Low achievers in the CLIL group (2.99) than that of the Control group (2.17).

	Proficiency	Group	Mean	StdErr	Lower	Upper	F Value	P Value
T0	Low	CLIL	0.91	0.25	0.53	1.55	0.74	0.3905
	Low	Control	1.22	0.24	0.83	1.79		
T1	Low	CLIL	0.53	0.18	0.27	1.02	8.91	0.0030

	Proficiency	Group	Mean	StdErr	Lower	Upper	F Value	P Value
T2	Low	Control	1.63	0.28	1.16	2.27	3.90	0.0492
	Low	CLIL	1.20	0.29	0.75	1.92		
T3	Low	Control	2.11	0.33	1.55	2.87	1.76	0.1857
	Low	CLIL	2.99	0.55	2.09	4.29		
	Low	Control	2.17	0.34	1.60	2.95		

Table 5.64 Lexical Complexity achievement results: TNAdj. Group/Low achievers interaction

The intragroup comparison showed the progress of High and Low achievers in the different groups as for the TNAdj. From T0 to T1, the High achievers in the Control Group increased significantly ($p= 0.0245$), whereas there were no differences for the high achievers in the CLIL group ($p= 0.5750$). However, from T1 to T2, only the CLIL high achievers improved significantly ($p= 0.0007$). In the final time period, T2 to T3, the progress was not significant for the High achievers in any of the two groups (Control Group $p= 0.9981$; CLIL group $p= 0.2734$). As for the Low achievers, the progress made by the Low achievers in the Control Group was not significant at any of the time periods tested. As for the Low achievers in CLIL, the progress was not significant either during the first two periods. However, this group improved significantly from T2 to T3 ($p= 0.0011$). As can be seen in Table 5.64 above, the mean number of adjectives used by low achievers in the CLIL group (2.99) was higher than the mean number used by the Low achievers in the Control Group (2.17). High and Low achievers mean number of adjectives at T3 was the same (2.1). From to T3, all groups progressed significantly: Control High ($p= 0.0004$), CLIL High ($p= <.0001$), Control Low ($p= 0.0264$), CLIL Low ($p= 0.0001$). The CLIL Low was the group that improved the most.

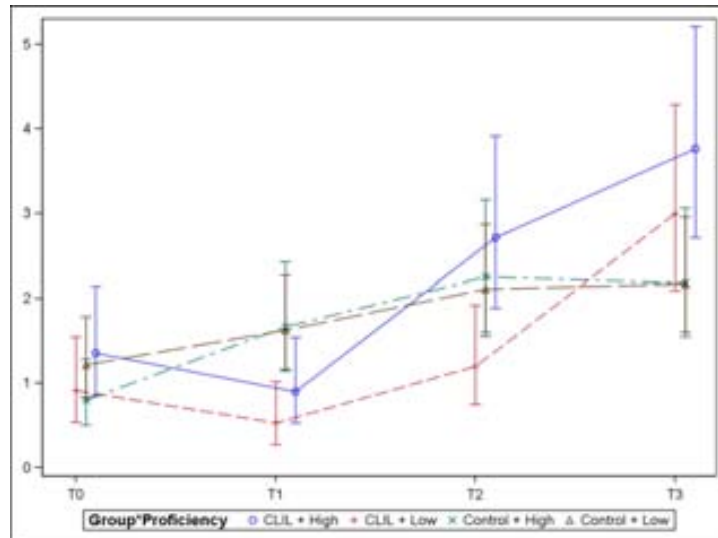


Figure 5.24 Lexical Complexity improvement results: TNAdj. Group/Proficiency level interaction.

Proficiency	Group	I0_1	I1_2	I2_3	I0_3
Low	CLIL	-42.1%	127.1%	149.9%	228.9%
	Control	33.6%	30.0%	2.9%	78.7%
High	CLIL	-33.3%	200.8%	38.6%	177.9%
	Control	107.4%	34.8%	-3.2%	170.6%

Table 5.65 Percentage of improvement in Lexical Complexity:TNAdj. Group/Proficiency level interaction

5.1.4.3.1.4 Ratio: Total Number of Adjectives/Total Number of Words in English

Table 5.66 shows the results of the ratio TNAdj/TNWE. At T0 there were no differences between the CLIL and the Control Groups, however, at T1, the Control Group significantly outperformed the results obtained by the CLIL group ($p= 0.0309$). At T2 and T3, no significant differences were found between the groups.

	Group	Mean	StdErr	Lower	Upper	F Value	P Value
T0	CLIL	2.3%	0.5%	1.5%	3.4%	0.25	0.6140

	Group	Mean	StdErr	Lower	Upper	F Value	P Value
T1	Control	2.6%	0.4%	1.9%	3.6%	4.70	0.0309
	CLIL	1.4%	0.3%	0.9%	2.2%		
T2	Control	2.5%	0.3%	2.0%	3.2%	0.53	0.4660
	CLIL	2.6%	0.4%	1.9%	3.5%		
T3	Control	3.0%	0.3%	2.4%	3.7%	3.27	0.0718
	CLIL	3.2%	0.4%	2.6%	4.0%		
	Control	2.4%	0.3%	2.0%	3.0%		

Table 5.66 Lexical Complexity Achievement results. %TNAdj/TNWE

The results of the intragroup analyses carried out to determine the improvement made by the groups showed similar developmental patterns. Both groups decreased from T0 to T1, the CLIL group being the one that decreased the most. From T1 to T2 both groups increased. However, from T2 to T3, the CLIL group improved whereas the Control Group decreased again. None of the improvement percentages was statistically significant. The improvement from T0 to T3 was not significant for any of the groups.

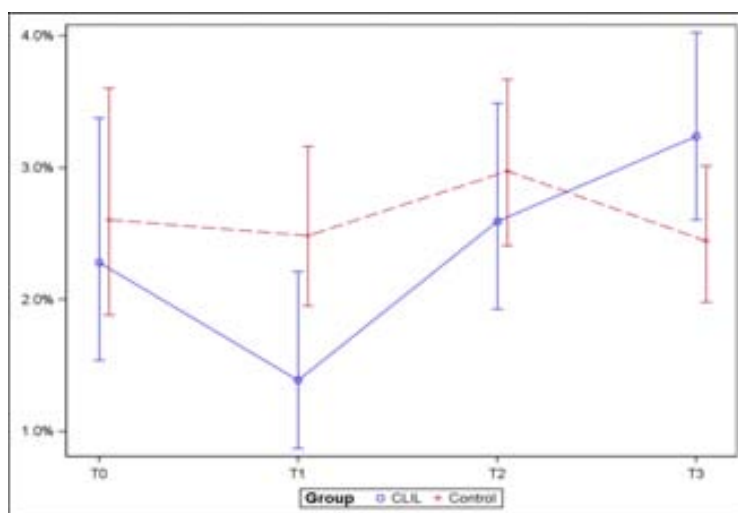


Figure 5.25 Lexical Complexity improvement results: %TNAdj/TNWE

	I0_1	I1_2	I2_3	I3_0
CLIL	-39.1%	86.9%	24.8%	42.1%
Control	-4.6%	19.7%	-17.8%	-6.2%

Table 5.67 Percentage of improvement in Lexical Complexity: %TNAdj/TNWE

5.1.4.3.2 Syntactic Complexity

Two measures were taken into consideration for the analysis of the syntactic complexity: Instances of Coordinated Units (ICU), that is, the number of examples of coordination and Instances of Subordinate Units (ISU), the number of examples of subordination.

5.1.4.3.2.1 Instances of Coordinated Units

None of the achievement results for ICU was statistically significantly different at any of the times tested. The mean results obtained by the Control Group were slightly better at T1 than those obtained by the Control Group. However, at Times 2 and 3, it was the CLIL group that showed more positive mean results (see Table 5.68).

	Group	Mean	StdErr	Lower	Upper	F Value	P Value
T0	CLIL	1.22	0.24	0.83	1.79	2.54	0.1122
	Control	0.78	0.15	0.53	1.15		
T1	CLIL	1.00	0.22	0.64	1.55	1.47	0.2270
	Control	1.40	0.22	1.02	1.91		
T2	CLIL	2.40	0.38	1.75	3.28	0.57	0.4495
	Control	2.04	0.28	1.55	2.68		
T3	CLIL	2.70	0.42	1.99	3.65	0.64	0.4246
	Control	2.29	0.30	1.76	2.97		

Table 5.68 Syntactic Complexity achievement results: ICU

Figure 5.26 and Table 5.69 show a steady improvement in the progression of the Control Group. Nevertheless, only the improvement of the Control Group from T0 to T1 was statistically significantly different ($p= 0.0452$). The CLIL group did not improve significantly from T0 to T1; however, the improvement of the CLIL group from T1 to T2 was statistically significant (P Value 0.0008). The progress from T2 to T3 was not

significant for any of the groups but the progress throughout the study, T0 to T3, was significant for the CLIL ($p= 0.0004$) and the Control Groups ($p= <.0001$).

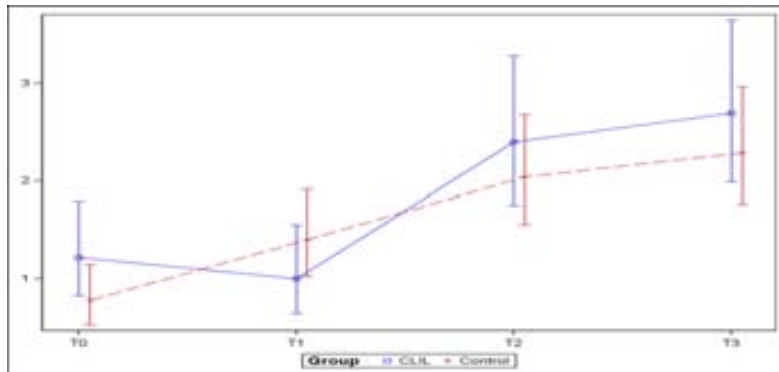


Figure 5.26 Syntactic Complexity improvement results: ICU

	I0_1	I1_2	I2_3	I0_3
CLIL	-17.9%	139.9%	12.4%	121.6%
Control	79.9%	45.9%	12.1%	194.2%

Table 5.69 Percentage of improvement in Syntactic Complexity: ICU

The results of the interaction Group/ High achievers show that there were no statistically significant differences for High achievers in the CLIL and Control groups at any of the times tested in terms of ICU except for T0 in favour of CLIL High achievers ($p= 0.0030$). Descriptive statistics revealed that at T1, High achievers in CLIL obtained lower mean results (1.53) than High achievers in the Control Group (1.87). At T2 and at T3, however, the High achievers in CLIL performed better than their counterparts in the Control Group. (See Table 5.70)

	Proficiency	Group	Mean	StdErr	Lower	Upper	F Value	P Value
T0	High	CLIL	2.06	0.46	1.32	3.19	8.95	0.0030
	High	Control	0.72	0.19	0.42	1.21		
T1	High	CLIL	1.53	0.38	0.94	2.48	0.39	0.5321
	High	Control	1.87	0.38	1.25	2.80		
T2	High	CLIL	2.94	0.57	2.01	4.30	0.71	0.4000

	Proficiency	Group	Mean	StdErr	Lower	Upper	F Value	P Value
T3	High	Control	2.33	0.44	1.61	3.38	0.13	0.7157
	High	CLIL	2.71	0.52	1.85	3.97		
	High	Control	2.45	0.45	1.71	3.52		

Table 5.70 Syntactic Complexity achievement results. Group/High achievers interaction

As for Low achievers, no differences were found between the CLIL and the Control Groups. Even though at T1 and at T2, the Low achievers in the Control Group attained better mean differences, at T2 and T3, the Low achievers in CLIL performed better than the Low achievers in the Control Group. The mean results of the CLIL Low at T3 (2.69) were almost the same as the results of the CLIL High (2.71). The group with the lowest results at T3 was the Control Low (see Table 5.71).

	Proficiency	Group	Mean	StdErr	Lower	Upper	F Value	P Value
T0	Low	CLIL	0.72	0.21	0.41	1.28	0.18	0.6704
	Low	Control	0.84	0.20	0.53	1.33		
T1	Low	CLIL	0.65	0.20	0.36	1.19	1.59	0.2085
	Low	Control	1.05	0.22	0.69	1.58		
T2	Low	CLIL	1.96	0.42	1.29	2.99	0.11	0.7357
	Low	Control	1.78	0.32	1.26	2.53		
T3	Low	CLIL	2.69	0.53	1.81	3.98	0.77	0.3801
	Low	Control	2.13	0.36	1.53	2.98		

Table 5.71 Syntactic Complexity achievement results. Group/Low achievers interaction

The results of the intragroup comparison when the variable High achievers was taken into account showed that, from T0 to T1, the progress of the High achievers in the Control Group ($p= 0.0047$) was significant whereas the progress of the High achievers in CLIL was not. From T1 to T2, only the CLIL group improved significantly ($p= 0.0437$). During the final time period, T2 to T3, no significant differences were found

for any of the groups. As in the case of High achievers, the Low achievers in CLIL progressed significantly from T1 to T2 ($p= 0.0029$). The progress throughout the study was only significant for three of the groups: Control High ($p= <.0001$), Control Low ($p= 0.0008$) and CLIL Low ($p= <.0001$) (see Table 5.72 and Figure 5.27).

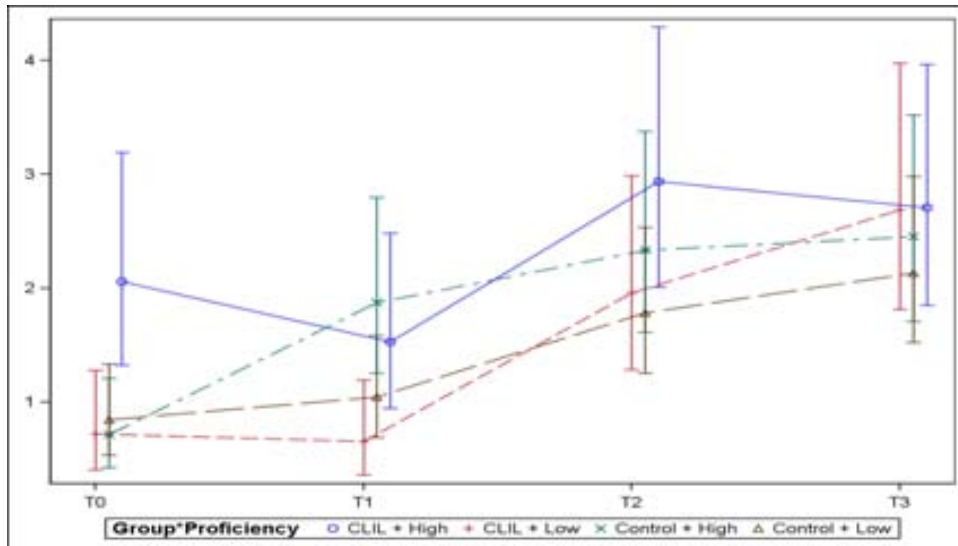


Figure 5.27 Syntactic Complexity improvement results. Group/Proficiency level interaction

Proficiency	Group	I0_1	I1_2	I2_3	I0_3
Low	CLIL	-9.2%	199.7%	37.1%	272.9%
	Control	23.8%	70.6%	19.6%	152.6%
High	CLIL	-25.6%	92.1%	-7.8%	31.6%
	Control	161.4%	24.7%	5.1%	242.6%

Table 5.72 Percentage of improvement in Syntactic Complexity. Group/Proficiency level interaction

5.1.4.3.2.2 Instances of Subordinate Units

There were no instances of subordination at T0. No significant differences were found at any of the other times tested. The Control Group seemed to slightly outperform the CLIL group at T1 and T2. However, although the differences were no statistically

significantly different, the mean score for the CLIL group (1.16) was higher at T3 than the score obtained by the Control Group (0.89).

	Group	Mean	StdErr	Lower	Upper	F Value	P Value
T1	CLIL	0.15	0.08	0.06	0.42	3.38	0.0672
	Control	0.44	0.11	0.26	0.74		
T2	CLIL	0.58	0.18	0.31	1.09	0.40	0.5286
	Control	0.75	0.17	0.47	1.19		
T3	CLIL	1.16	0.33	0.66	2.03	0.53	0.4695
	Control	0.89	0.20	0.56	1.39		

Table 5.73 Syntactic Complexity achievement results: ISU

In terms of progress, the CLIL group improved significantly from T1 to T2 ($p=0.0156$) and from T2 to T3 ($p=0.0056$). The improvement for the Control Group was only statistically significant from T1 to T2 ($p=0.0350$).

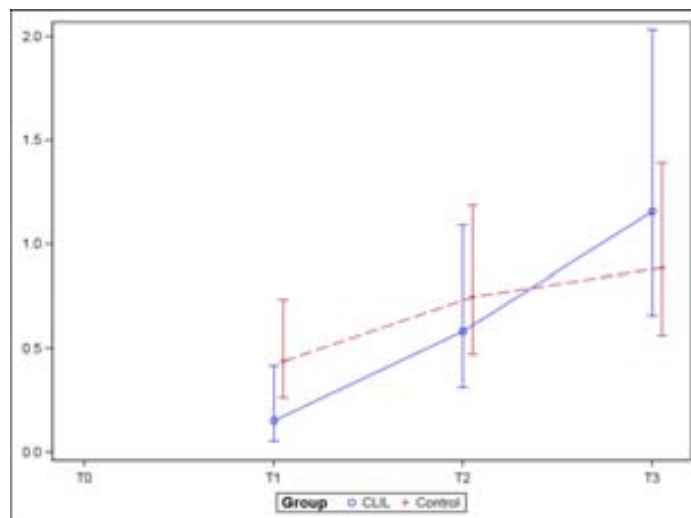


Figure 5.28 Syntactic Complexity Improvement results: ISU

	I1_2	I2_3	I1_3
CLIL	283.3%	98.4%	660.3%
Control	70.4%	18.2%	101.3%

Table 5.74 Percentage of improvement in Syntactic Complexity: ISU

As can be seen in Table 5.75 below, there were no statistically significant differences for High achievers between the Control and CLIL groups. Descriptive statistics show that, even though the results of the Control group were higher at T1 and T2, at the end of the study, at T3, the CLIL group slightly outperformed the Control Group.

	Proficiency	Group	Mean	StdErr	Lower	Upper	F Value	P Value
T1	High	CLIL	0.19	0.10	0.06	0.56	3.29	0.0709
	High	Control	0.62	0.21	0.31	1.22		
T2	High	CLIL	1.07	0.38	0.52	2.17	0.00	0.9609
	High	Control	1.09	0.34	0.58	2.04		
T3	High	CLIL	1.55	0.52	0.80	3.03	0.28	0.5994
	High	Control	1.22	0.38	0.66	2.26		

Table 5.75 Syntactic Complexity achievement results: ISU. Group/High achievers

The picture for Low achievers was the same as for High achievers. No significant differences were found in the number of Instances of Subordinate Units. The mean number of subordinate units used by the Low achievers was lower than the mean number used by the High achiever students. The Low achievers in the Control Group outperformed the Low achievers in the CLIL group at T1 and at T2. At T3, even though the mean score was still very low, the score of the CLIL group (0.86) was slightly higher than that of the Control group (0.64).

	Proficiency	Group	Mean	StdErr	Lower	Upper
T1	Low	CLIL	0.12	0.07	0.04	0.40
	Low	Control	0.31	0.11	0.16	0.62
T2	Low	CLIL	0.32	0.14	0.14	0.74
	Low	Control	0.52	0.16	0.28	0.95
T3	Low	CLIL	0.86	0.33	0.41	1.82

Proficiency	Group	Mean	StdErr	Lower	Upper
Low	Control	0.64	0.20	0.35	1.18

Table 5.76 Syntactic Complexity achievement results: ISU. Group/Low achievers

The progress of the groups when the interaction Group/Proficiency level was taken into account showed that the CLIL High achievers progressed significantly from T1 to T2 ($p= 0.0023$) and that the CLIL Low achievers significantly improved from T2 to T3 ($p= 0.0033$). The Control groups did not show any significant differences during any of these time periods. However, the improvement throughout the study was significant for all the groups: CLIL High ($p=<.0001$), CLIL Low ($p= 0.0007$), Control High ($p= 0.0151$) and Control Low ($p= 0.0280$) (see Table 5.77 and Figure 5.29).

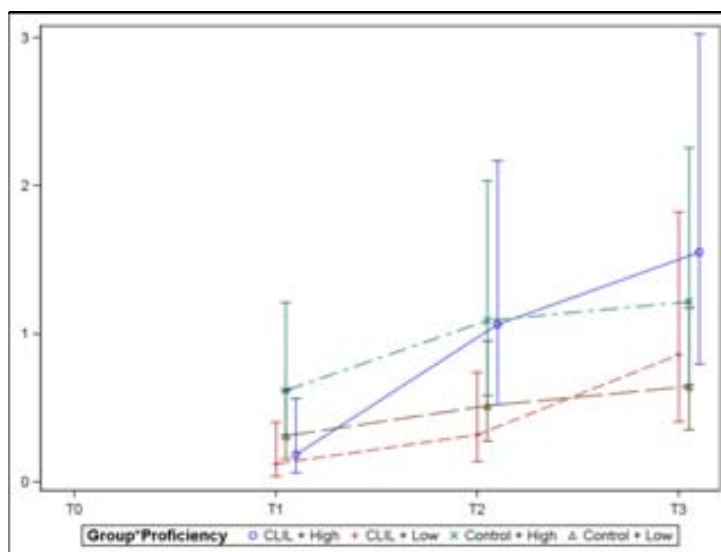


Figure 5.29 Syntactic Complexity improvement results: ISU. Group/Low achievers interaction

Proficiency	Group	I1_2	I2_3	I1_3
Low	CLIL	156.3%	169.8%	591.6%
	Control	64.4%	25.1%	105.7%
High	CLIL	473.2%	45.8%	735.8%
	Control	76.5%	11.6%	97.0%

Table 5.77 Percentage of Syntactic Complexity improvement: ISU. Group/Proficiency level interaction

5.1.5 Summary of CLIL Science Writing Results

5.1.5.1 Summary of Writing Achievement Results

This section presents a summary of the fluency achievement results obtained by CLIL Science students compared to their Control Group at different times. Table 5.78 displays a summary of the different measures used to assess the results in Fluency of CLIL Science students.

FLUENCY	TNW		TNWE		%TNWE/TNW		TNU	
	Control	CLIL	Control	CLIL	Control	CLIL	Control	CLIL
T0	64.87	63.55	45.64	51.16	66.35%	78.15%	7.80	9.81
	P Value 0.8748		P Value 0.4934		P Value 0.0042		P Value 0.1519	
T1	96.35	65.70	75.16	56.33	74.4%	84.7%	11.29	8.86
	P Value 0.0003		P Value 0.0203		P Value 0.0095		P Value 0.0841	
T2	96.35	88.88	83.61	77.22	81.9%	84.4%	14.78	12.99
	P Value 0.2468		P Value 0.4286		P Value 0.5179		P Value 0.2027	
T3	121.4	107.6	99.38	110.2	90.4%	89.4%	14.95	17.68
	P Value 0.1018		P Value 0.1821		P Value 0.7886		P Value 0.0524	

Table 5.78 Summary of Fluency achievement results

Table 5.79 provides a summary of Fluency results for the TNWE and TNU when the interaction Group/Proficiency level was taken into consideration.

FLUENCY	TNWE				TNU			
	HIGH		LOW		HIGH		LOW	
	Control	CLIL	Control	CLIL	Control	CLIL	Control	CLIL
T0	48.76	56.79	42.51	45.54	8.40	10.55	7.20	9.07
	P Value 0.0457		P Value 0.7716		P Value 0.2501		P Value 0.3018	
T1	85.63	63.01	64.69	49.64	12.60	9.68	9.98	8.04
	P Value 0.0364		P Value 0.1499		P Value 0.1193		P Value 0.2851	

T2	91.68	91.92	75.53	62.53	15.90	15.41	13.65	10.57
	P Value 0.9825		P Value 0.2129		P Value 0.7927		P Value 0.0896	
T3	112.3	123.3	86.45	97.02	16.06	19.33	13.83	16.02
	P Value 0.3067		P Value 0.3128		P Value 0.0815		P Value 0.2278	

Table 5.79 Summary of Fluency achievement results. Group/Proficiency level interaction

Table 5.80 sums up the Accuracy achievement results and Table 5.81 sums up the achievement results as for the TNEFU when the interaction Group/ Proficiency level was taken into consideration.

ACCURACY	TNEFU		%TNEFU/TNU	
	Control	CLIL	Control	CLIL
T0	2.49	1.88	35.2%	21.1%
	P Value 0.2840		P Value 0.0098	
T1	3.37	3.07	35.0%	37.6%
	P Value 0.6978		P Value 0.6807	
T2	5.31	4.67	41.6%	41.2%
	P Value 0.5556		P Value 0.9447	
T3	6.02	7.35	46.3%	44.3%
	P Value 0.3229		P Value 0.7508	

Table 5.80 Summary of Accuracy achievement results

ACCURACY	TNEFU			
	HIGH		LOW	
	Control	CLIL	Control	CLIL
T0	3.04	1.81	2.05	1.95
	P Value 0.1219		P Value 0.8831	
T1	4.58	4.01	2.47	2.34
	P Value 0.6540		P Value 0.8604	
T2	7.08	6.34	3.99	3.44
	P Value 0.6842		P Value 0.6024	

Language Competence of Young Learners Exposed to EFL and CLIL

T3	7.55	9.09	4.81	5.95
	P Value 0.4736		P Value 0.4177	

Table 5.81 Summary of Accuracy achievement results: TNEFU. Interaction Group/ Proficiency Level

Table 5.82 displays a summary of the different measures in Lexical Complexity achievement results.

LEXICAL COMPLEXITY	TNLV		%TNLV/TNWE		TNAdj		TNAdj/TNWE	
	Control	CLIL	Control	CLIL	Control	CLIL	Control	CLIL
T0	1.55	1.48	4.0%	3.2%	0.99	1.11	2.6%	2.3%
	P Value 0.8472		P Value 0.3092		P Value 0.6737		P Value 0.6140	
T1	1.88	2.52	3.0%	5.1%	1.65	0.69	2.5%	1.4%
	P Value 0.1598		P Value 0.0011		P Value 0.0029		P Value 0.0309	
T2	2.97	2.59	3.9%	4.3%	2.18	1.80	3.0%	2.6%
	P Value 0.5079		P Value 0.5312		P Value 0.3756		P Value 0.4660	
T3	3.37	4.04	4.0%	4.1%	2.18	3.36	2.4%	3.2%
	P Value 0.3468		P Value 0.9006		P Value 0.0212		P Value 0.0718	

Table 5.82 Summary of Lexical Complexity achievement results

Table 5.83 below shows the results in Lexical Complexity (TNLV and TNAdj) when the interaction Group/Proficiency level was considered.

LEXICAL COMPLEXITY Group/proficiency	TNLV				TNAdj			
	HIGH		LOW		HIGH		LOW	
	Control	CLIL	Control	CLIL	Control	CLIL	Control	CLIL
T0	2.16	2.19	1.11	1.00	0.80	1.35	1.22	0.91
	P Value 0.9608		P Value 0.7366		P Value 0.1217		P Value 0.3905	
T1	2.37	3.41	1.48	1.86	1.67	0.90	1.63	0.53

Language Competence of Young Learners Exposed to EFL and CLIL

	P Value 0.1710		P Value 0.4117		P Value 0.0685		P Value 0.0030	
T2	3.92	3.43	2.25	1.96	2.25	2.72	2.11	1.20
	P Value 0.5997		P Value 0.6133		P Value 0.4654		P Value 0.0492	
T3	4.15	5.07	2.74	3.21	2.18	3.76	2.17	2.99
	P Value 0.4142		P Value 0.5291		P Value 0.0255		P Value 0.1857	

Table 5.83 Summary of Lexical Complexity achievement results. Group/Proficiency level interaction

Table 5.84 sums up the Syntactic Complexity achievement results: ICU and ISU.

SYNTACTIC COMPLEXITY	ICU		ISU	
	Control	CLIL	Control	CLIL
T0	0.78	1.22		
	P Value 0.1122		P Value	
T1	1.40	1.00	0.44	0.15
	P Value 0.2270		P Value 0.0672	
T2	2.04	2.40	0.75	0.58
	P Value 0.4495		P Value 0.5286	
T3	2.29	2.70	0.89	1.16
	P Value 0.4246		P Value 0.4695	

Table 5.84 Summary of Syntactic Complexity achievement results

Table 5.85 shows the achievement results in Syntactic Complexity for High and Low achievers.

SYNTACTIC COMPLEXITY Group/proficiency	ICU				ISU			
	HIGH		LOW		HIGH		LOW	
	Control	CLIL	Control	CLIL	Control	CLIL	Control	CLIL
T0	0.72	2.06	0.84	0.72				
	P Value 0.0030		P Value 0.6704		P Value		P Value	

T1	1.87	1.53	1.05	0.65	0.62	0.19	0.31	0.12
	P Value 0.5321		P Value 0.2085		P Value 0.0709		P Value 0.1820	
T2	2.33	2.94	1.78	1.96	1.09	1.07	0.52	0.32
	P Value 0.4000		P Value 0.7357		P Value 0.9609		P Value 0.3693	
T3	2.45	2.71	2.13	2.69	1.22	1.55	0.64	0.86
	P Value 0.7157		P Value 0.3801		P Value 0.5994		P Value 0.5508	

Table 5.85 Summary of Syntactic Complexity achievement results. Group/Proficiency level interaction

5.1.5.2 Summary of Writing Improvement Results

Table 5.86 sums up the improvement in the different measures taken to assess Fluency.

FLUENCY	TNW		TNWE		%TNWE/TNW		TNU	
	Control	CLIL	Control	CLIL	Control	CLIL	Control	CLIL
T0-T1	48.5% p=<.0001	3.4% p=0.9894	64.7% p=<.0001	10.1% p=0.8304	12.2% p=0.0257	8.3% p=0.4361	44.7% p=0.0004	-9.7% p=0.8162
T1-T2	2.4% p=0.9755	35.3% p=0.0048	11.2% p=0.2071	37.1% p=0.0038	10.1% p=0.0262	-0.3% p=0.9999	30.9% p=0.0004	46.7% p=0.0010
T2-T3	9.1% p=0.3593	36.6% p=<.0001	18.9% p=0.0066	42.7% p=<.0001	10.3% p=0.0094	5.9% p=0.4838	1.1% p=0.9974	36.1% p=0.0001
T0-T3	65.8% p= <.0001	91.0% p=<.0001	117.8% p=<.0001	115.3% p=<.0001	36.2% p=<.0001	14.4% p=0.0219	91.6% p=<.0001	80.2% p=<.0001

Table 5.86 Summary of Fluency improvement results

The description of the improvement results for the TNWE and the TNU, when the variable Group interacted with Proficiency level is displayed in Table 5.87 below.

FLUENCY Group/Proficiency level interaction	TNWE				TNU			
	HIGH		LOW		HIGH		LOW	
	Control	CLIL	Control	CLIL	Control	CLIL	Control	CLIL
T0-T1	75.6% p=<.0001	11.0% p=0.8474	52.2% p=0.0026	9.0% p=0.9552	50.0% p=0.0054	-8.3% p=0.9190	38.6% p=0.0652	-11.4% p=0.8862

Language Competence of Young Learners Exposed to EFL and CLIL

T1-T2	7.1% p=0.8253	45.9% p=0.0011	16.8% p=0.3082	26.0% p=0.3694	26.3% p=0.0452	59.3% p=0.0002	36.8% p=0.0064	31.4% p=0.2877
T2-T3	22.5% p=0.0194	34.2% p=0.0003	14.5% p=0.3106	55.2% p=0.0001	1.0% p=0.9993	25.5% p=0.0230	1.3% p=0.9987	51.7% p=0.0009
T0-T3	130.3% p<<.0001	117.2% p<<.0001	103.4% p<<.0001	113.0% p<<.0001	91.3% p<<.0001	83.3% p<<.0001	92.1% p<<.0001	76.6% p<<.0001

Table 5.87 Summary of Fluency improvement results. Group/Proficiency Level interaction

The findings from the intragroup comparison as for their Accuracy improvement are summarized in Table 5.88.

ACCURACY	TNEFU		%TNEFU/TNU	
	Control	CLIL	Control	CLIL
T0-T1	35.0% p= 0.195	63.2% p=0.1130	-0.6% p=1.0000	78.2% p=0.0100
T1-T2	57.8% p= 0.0022	52.3% p=0.0897	18.8% p=0.3922	9.5% p=0.9339
T2-T3	13.4% p= 0.6315	57.4% p=0.0087	11.2% p=0.6364	7.7% p=0.9310
T0-T3	13.4% p= <.0001	57.4% p=0.0087	11.2% p=0.6364	7.7% p=0.9310

Table 5.88 Summary of Accuracy improvement results

Table 5.89 below displays the improvement results of High and Low achievers in terms of the TNEFU.

ACCURACY Group/Proficiency level interaction	TNEFU			
	HIGH		LOW	
	Control	CLIL	Control	CLIL
T0-T1	50.7% p=0.0987	121.6% p=0.0117	21.0% p=0.7577	20.2% p=0.8983
T1-T2	54.6% p=0.0164	58.0% p=0.0956	61.1% p=0.0210	46.9% p=0.3262
	6.7%	43.4%	20.6% %	72.8% %

Language Competence of Young Learners Exposed to EFL and CLIL

T2-T3	p=0.9540	p=0.1049	p=0.5351	p=0.0146
T0-T3	148.8% p=<.0001	402% p=<.0001	135.0% p=<.0001	205.1% p=<.0001

Table 5.89 Summary of Accuracy improvement results. Group/Proficiency level interaction

The progress of the different groups as for their Lexical Complexity is summarized in Table 5.90 below.

LEXICAL COMPLEXITY	TNLV		%TNLV/TNWE		TNAdj		%TNAdj/TNWE	
	Control	CLIL	Control	CLIL	Control	CLIL	Control	CLIL
T0-T1	20.9% p=0.4571	69.6% p=0.0079	-26.1% p=0.1458	57.5% p=0.0669	66.5% p=0.0456	-37.8% p=0.3720	-4.6% p=0.9952	-31.9% p=0.3499
T1-T2	58.2% p=0.0002	3.1% p=0.9962	44.8% p=0.0115	-23.9% p=0.2585	32.4% p=0.2419	161.4% p=0.0024	19.7% p=0.6331	86.9% p=0.0973
T2-T3	13.7% p=0.5221	55.5% p=0.0021	-5.5% p=0.9714	6.1% p=<.0001	-0.2% p=1.0000	86.1% p=0.0016	-17.8% p=0.4893	24.8% p=0.5661
T0-T3	117.3% p=<.0001	171.9% p=<.0001	1.1% p=0.9997	27.1% p=0.5000	119.9% p=0.0002	202.3% p=<.0001	-6.2% p=0.3588	42.1% p=0.9863

Table 5.90 Summary of Lexical Complexity improvement results

The findings from the intragroup comparison when the variable Proficiency level was taken into account can be seen in Table 5.91 below.

LEXICAL COMPLEXITY Group/Proficiency level interaction	TNLV				TNAdj			
	HIGH		LOW		HIGH		LOW	
	Control	CLIL	Control	CLIL	Control	CLIL	Control	CLIL
T0-T1	9.8% p=0.9297	55.5% p=0.0646	33.1% p=0.3601	85% p=0.0281	107.4% p=0.0245	-33.3% p=0.5750	33.6% p=0.5418	-41.2% p=0.5022
T1-T2	65.2% p=0.0012	0.6% p=1.0000	51.4% p=0.0273	5.6% p=0.9909	34.8% p=0.4459	200.8% p=0.0007	30.0% p=0.4721	127.1% p=0.1187
T2-T3	5.8% p=0.9599	47.8% p=0.0253	22.1% p=0.3975	63.6% p=0.0148	-3.2% p=0.9981	38.6% p=0.2734	2.9% p=0.9983	149.9% p=<.0001
T0-T3	92.0% p=<.0001	131.3% p=<.0001	146.1% p=<.0001	219.5% p=<.0001	170.6% p=0.0004	177.9% p=<.0001	78.7% p=0.0264	228.9% p=0.0001

Table 5.91 Summary of Lexical Complexity improvement results. Group/ Proficiency level interaction

Table 5.92 sums up the improvement results in Syntactic Complexity.

SYNTACTIC COMPLEXITY	ICU		ISU	
	Control	CLIL	Control	CLIL
T0-T1	79.9% p=0.0452	-17.9% p=0.8663		
T1-T2	45.9% p=0.1318	139.9% p=0.0008	70.4% p=0.0350	283.3% p=0.0156
T2-T3	12.1% p=0.8694	12.4% p=0.8807	18.2% p=0.5902	98.4% p=0.0056
T0-T3	194.2% p<.0001	121.6% p=0.0004	101.3% p<.0001	660.3% p<.0001

Table 5.92 Summary of Syntactic Complexity improvement results

The results of the intragroup comparison when the interaction Group/Proficiency level was taken into account are displayed in Table 5.93 below.

SYNTACTIC COMPLEXITY Group/Proficiency level interaction	ICU				ISU			
	HIGH		LOW		HIGH		LOW	
	Control	CLIL	Control	CLIL	Control	CLIL	Control	CLIL
T0-T1	161.4% p=0.0047	-25.6% p=0.5967	23.8% p=0.8601	-9.2% p=0.9937				
T1-T2	24.7% p=0.7276	92.1% p=0.0437	70.6% p=0.0877	199.7% p=0.0029	76.5% p=0.0563	473.2% p=0.0023	64.4% p=0.1958	156.3% p=0.201
T2-T3	5.1% p=0.9940	-7.8% p=0.9753	19.6% p=0.7773	37.1% p=0.4318	11.6% p=0.8472	45.8% p=0.2436	25.1% p=0.6114	169.8% p=0.0033
T0-T3	242.6% p<.0001	31.6% p=0.6153	152.6% p=0.0008	272.9% p<.0001	97.5% p=0.0151	735.8% p<.0001	105.7% p=0.0280	591.6% p=0.0007

Table 5.93 Summary of Syntactic Complexity improvement results. Group/ Proficiency level.

5.2 CLIL Arts & Crafts results

In sections 5.2.1 and 5.2.2 a description of the results obtained by CLIL Arts & crafts students compared to those in their Control Groups will be presented in order to answer the following research questions and sub questions:

RQ1: Keeping the number of hours of exposure to English the same for both groups, CLIL and EFL, do the CLIL students' listening and reading skills benefit from their

exposure to the CLIL experience? RQ1.1 Are there any differences in achievement between CLIL and EFL learners statistically significantly in favour of CLIL learners at different times (T0, T1, T2, T3)? RQ1.2 Are there any differences in progress between CLIL and EFL learners significantly in favour of CLIL learners after one year (T0-T1) and two years (T2-T3) of CLIL implementation? Are there differences in favour of CLIL students when we consider their progress from T0-T3? RQ 1.3 How does the initial level of English proficiency affect the students' performance in the CLIL and the EFL groups?

5.2.1 Listening test results

Intergroup comparisons between the CLIL and the Control Group at different times do not yield the same results for all the times. The comparison of the mean scores obtained at T0 showed that, despite the slightly higher percentage obtained by the students in the Control Group, the differences between the CLIL and the Control Groups were not statistically significant ($F=0.17$, $p=0.6845$) and the differences at T1 were not significant either. However, the comparisons at T2 and T3 showed statistically significant differences in favour of the Control Group (T2: $F=7.65$, $p=0.0060$; T3: $F=9.02$, $p=0.0029$).

	Group	Mean	StdErr	Lower	Upper	F value	P Value
T0	CLIL	27.2%	3.2%	21.4%	34.0%	0.17	0.6845
	Control	29.0%	2.9%	23.6%	35.1%		
T1	CLIL	42.9%	3.7%	35.9%	50.2%	0.76	0.3825
	Control	47.3%	3.5%	40.6%	54.2%		
T2	CLIL	46.2%	3.7%	39.1%	53.6%	7.65	0.0060
	Control	60.4%	3.4%	53.5%	67.0%		
T3	CLIL	50.6%	3.6%	43.6%	57.6%	9.02	0.0029
	Control	65.5%	3.3%	58.7%	71.7%		

Table 5.94: Listening achievement results

Intragroup comparisons carried out to determine the students' improvement in percentages at different time periods showed a similar increase for both the CLIL and the Control Groups in the first year. Both groups increased significantly from T0 to T1 (Control Group $p < .0001$; CLIL Group $p = 0.0009$). Between T1 and T2, there were only significant differences for the Control Group ($p = 0.0062$). Between T2 and T3, both groups progressed although the improvement was not statistically significant and, as has been pointed out, it was the Control Group the one that obtained significant achievement results at T3. Nevertheless, the improvement T0-T3 was statistically significant for both the CLIL ($p < .0001$) and the Control Groups ($p < .0001$) (see Figure 5.30 and Table 5.95)

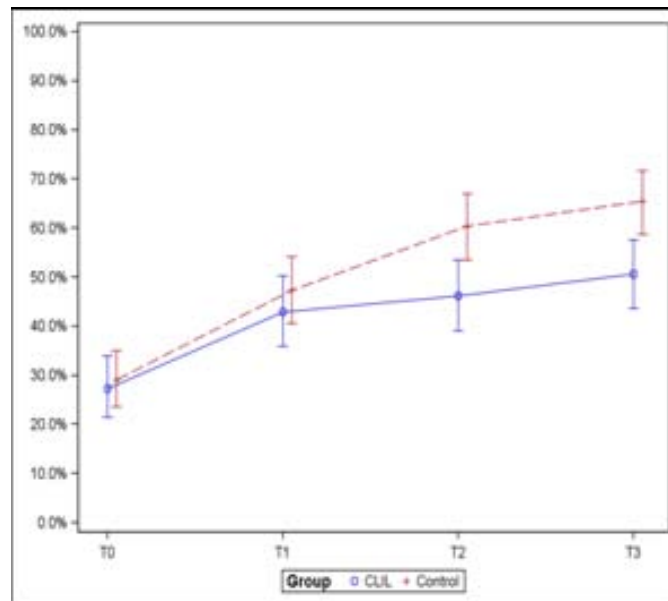


Figure 5.30: Listening improvement results

	I0_1	I1_2	I2_3	I0_3
CLIL	57.4%	7.8%	9.5%	85.9%
Control	63.1%	27.7%	8.4%	125.8%

Table 5.95 Improvement percentages in listening

Table 5.96 presents the listening achievement results when the interaction Group/High achiever was taken into account. As can be seen, even though CLIL High

achievers started with a slightly higher mean percentage at T0, there were statistically significant differences at T1 ($F= 7.86$ $p=0.0054$), T2 ($F= 8.93$ $p=0.0030$) and T3 ($F= 10.29$ $p=0.0015$) in favour of the High achievers in the Control Group.

	Proficiency	Group	Mean	StdErr	Lower	Upper	F Value	P Value
T0	High	CLIL	41.3%	3.8%	34.1%	49.0%	0.06	0.8092
	High	Control	39.7%	5.5%	29.6%	50.8%		
T1	High	CLIL	46.3%	4.0%	38.5%	54.2%	7.86	0.0054
	High	Control	65.4%	5.1%	54.8%	74.8%		
T2	High	CLIL	55.0%	3.9%	47.2%	62.5%	8.93	0.0030
	High	Control	74.1%	4.5%	64.3%	81.9%		
T3	High	CLIL	57.4%	3.9%	49.7%	64.8%	10.29	0.0015
	High	Control	77.8%	4.3%	68.2%	85.2%		

Table 5.96 Listening achievement results: Group/High achievers interaction

As for Low achievers (see Table 5.97), there were no statistically significant differences for any of the groups at any of the times tested. Descriptive statistics, however, showed that, except at T1 when the CLIL group achievement percentage was higher (39.6%) than that of the Control Group (29.9%), the rest of the times, the percentages were higher for the Control group.

	Proficiency	Group	Mean	StdErr	Lower	Upper	F Value	P Value
T0	Low	CLIL	16.6%	4.0%	10.1%	26.1%	0.53	0.4661
	Low	Control	20.3%	2.7%	15.5%	26.0%		
T1	Low	CLIL	39.6%	6.3%	28.1%	52.4%	1.99	0.1589
	Low	Control	29.9%	3.3%	23.9%	36.6%		
T2	Low	CLIL	37.7%	6.2%	26.5%	50.5%	0.97	0.3251
	Low	Control	45.0%	3.8%	37.7%	52.5%		
T3	Low	CLIL	43.8%	6.1%	32.4%	55.9%	0.94	0.3328
	Low	Control	50.7%	3.7%	43.5%	57.9%		

Table 5.97 Listening achievement results: Group/Low achievers interaction

Figure 5.31 and Table 5.98 below show the progress in listening when the variables proficiency level and group interacted. Even though High and Low achievers in the CLIL and the Control Groups seemed to improve, the differences were not always significant. The results of the High achievers in CLIL were not significant for any of the time periods whereas the results of the High achievers in the Control Group were statistically significant from T0 to T1 ($p= 0.0003$) but not significant during the other time periods. As for Low achievers, the ones in the CLIL group also improved significantly from T0 to T1 ($p= 0.0018$) but the differences were not significant from T1 to T2 and from T2 to T3. Low achievers in the Control Group improved significantly during the first two time periods: from T0 to T1 ($p= 0.0310$) and from T1 to T2 ($p= 0.0015$). From T2 to T3 their improvement was not significant. As has been reported, even though High and Low achievers in the CLIL and the Control Groups started at very similar points at T0, High and Low students in the Control Groups achieved better mean scores than their counterparts in the CLIL groups. The lowest scores were obtained by the Low achievers in the CLIL group. The comparison T0-T3 showed statistically significant results for all the groups: CLIL High achievers ($p= 0.0016$), Control High achievers ($p=<.0001$), CLIL Low achievers ($p= 0.0001$) and Control Low achievers ($p=<.0001$). As Table 5.98 displays, the improvement percentages of Low achievers were much higher than those of High achievers. Low achievers in the CLIL group were the ones with the highest improvement percentage even though their achievement percentage (43.8%) was lower than that obtained by their counterparts in the control group (50.7%).

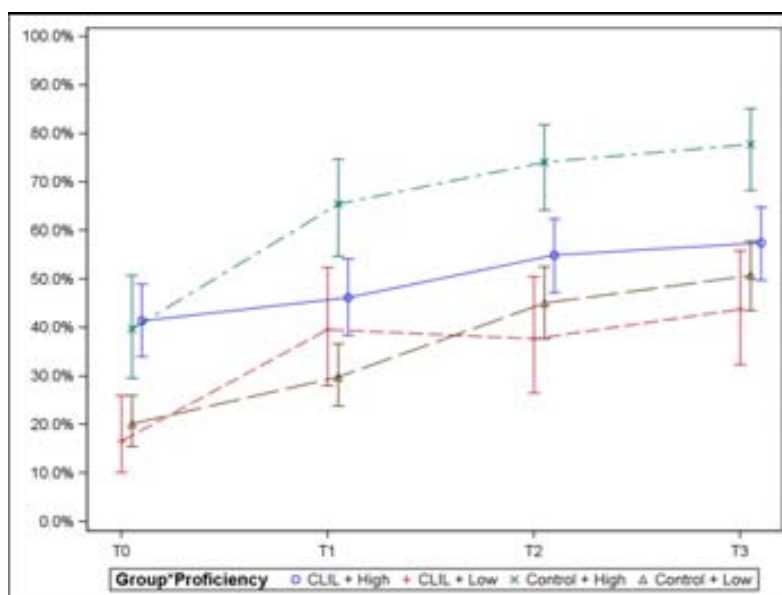


Figure 5.31 Listening improvement results. Group/Proficiency level interaction

Proficiency	Group	I0_1	I1_2	I2_3	I0_3
Low	CLIL	138.2%	-4.6%	16.0%	163.7%
	Control	47.5%	50.6%	12.7%	150.5%
High	CLIL	11.9%	18.8%	4.5%	39.0%
	Control	64.9%	13.2%	5.1%	96.1%

Table 5.98 Percentage of improvement in listening. Group/Proficiency level interaction

5.2.2 Reading test results

The total results obtained in the reading tests when the CLIL and the Control Groups were compared showed no significant differences at T0. However, at T1 and T2 the results in achievement were both statistically significant in favour of the Control Group (T1 $F = 0.00$ $p = 0.0007$; T2 $F = 0.9754$ $p < .0001$). At T3 there were no significant results. Nevertheless, at T3, the students in the CLIL group seem to catch up with the ones in the Control Group; the difference in the mean scores between the CLIL and the Control Group was much lower at T3 (4%) than the differences at T1 (15%) and T2 (18%) (see Table 5.99).

	Group	Mean	StdErr	Lower	Upper	F Value	P Value
T0	CLIL	31.0%	2.7%	25.8%	36.7%	0.00	0.9754
	Control	30.9%	2.6%	26.0%	36.2%		
T1	CLIL	42.5%	3.1%	36.5%	48.8%	12.21	0.0007
	Control	57.7%	2.9%	51.8%	63.4%		
T2	CLIL	46.9%	3.2%	40.7%	53.2%	17.34	<.0001
	Control	64.9%	2.8%	59.1%	70.2%		
T3	CLIL	59.6%	3.0%	53.4%	65.4%	0.95	0.3327
	Control	63.6%	2.8%	57.9%	68.9%		

Table 5.99 Reading achievement results

Table 5.100 displays the reading improvement results. Despite the fact that the CLIL and the Control Groups started at almost the same point and that both groups progressed, the intragroup comparison showed that the Control Group improved significantly from T0 to T1 ($p < .0001$) and from T1 to T2 ($p = 0.0017$). The progression from T2 to T3 was not significant for any of the groups, although, as has been pointed out in the description of achievement, the differences at T3 between the two groups were shorter. However, the progress made by both groups during the study, that is, from T0 to T3, was statistically significant for both, the CLIL ($p < .0001$) and the Control Groups ($p < .000$). (See Table 5.100 and Figure 5.32).

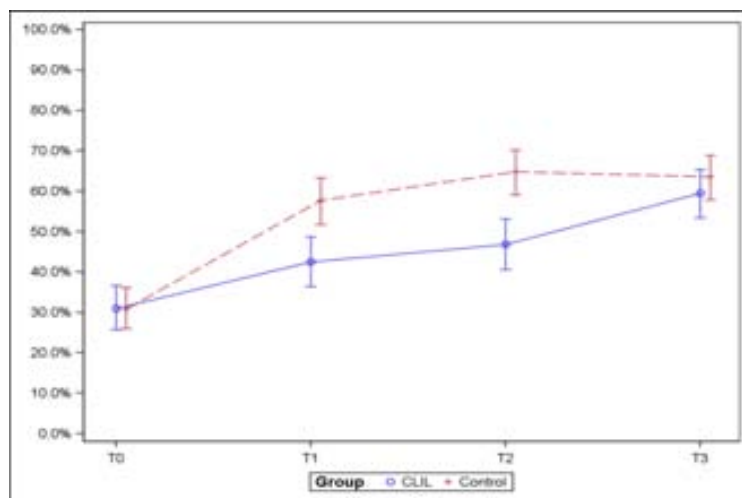


Figure 5.32 Reading improvement results

	I0_1	I1_2	I2_3	I0-3
CLIL	37.2%	10.3%	27.1%	92.3%
Control	87.1%	12.4%	-2.0%	106.1%

Table 5.100 Percentage of improvement in reading

The intergroup comparison taking into account the proficiency level of students showed that there were statistically significant differences at T1 ($F= 9.99$ $p= 0.0020$), T2 ($F= 14.51$ $p= 0.0002$) and T3 ($F= 4.77$ $p= 0.0309$) in favour of the High achievers in the Control Group (see Table 5.101).

	Proficiency	Group	Mean	StdErr	Lower	Upper	F Value	P Value
T0	High	CLIL	42.6%	3.4%	36.1%	49.4%	1.17	0.2816
	High	Control	49.1%	4.9%	39.5%	58.7%		
T1	High	CLIL	51.6%	3.5%	44.8%	58.4%	9.99	0.0020
	High	Control	69.7%	4.2%	60.9%	77.3%		
T2	High	CLIL	58.1%	3.4%	51.3%	64.6%	14.51	0.0002
	High	Control	78.0%	3.5%	70.4%	84.1%		
T3	High	CLIL	65.3%	3.1%	58.8%	71.2%	4.77	0.0309
	High	Control	76.3%	3.6%	68.3%	82.7%		

Table 5.101 Reading achievement results: Group/High achievers interaction

As can be seen in Table 5.102 below, the results of the comparison when the variables group/Low achievers were taken into account showed statistically significant differences between the Low achievers in the CLIL and Control Groups at T2 ($p= 0.0393$) in favour of the Control Group. Even though the CLIL group obtained a better mean percentage at T0, the score at T1 was higher for the Control Group. However, at the end of the study, at T3, the mean percentage in reading was higher for the Low achievers in the CLIL group (53.6%) than it was for the Low achievers in the Control Group (48.7%).

	Proficiency	Group	Mean	StdErr	Lower	Upper	F Value	P Value
T0	Low	CLIL	21.3%	3.8%	14.8%	29.7%	1.07	0.3023
	Low	Control	17.1%	1.9%	13.7%	21.3%		
T1	Low	CLIL	33.8%	5.0%	24.7%	44.3%	3.18	0.0774
	Low	Control	44.7%	3.2%	38.5%	51.1%		
T2	Low	CLIL	36.0%	5.1%	26.5%	46.6%	4.35	0.0393
	Low	Control	49.0%	3.3%	42.6%	55.4%		
T3	Low	CLIL	53.6%	5.4%	42.9%	64.0%	0.60	0.4387
	Low	Control	48.7%	3.2%	42.4%	55.1%		

Table 5.102 Reading achievement results: Group /Low achievers interaction

Figure 5.33 and Table 5.103 below display the progress of High and Low achievers when the interaction Group/Proficiency level was taken into account. High achievers in the CLIL group progressed significantly during the three time periods tested: from T0 to T1 ($p= 0.0004$), from T1 to T2 ($p= 0.0192$) and from T2 to T3 ($p= 0.0049$). Their progression during the study, from T0 to T3, was also significant ($p= <.0001$). However, High achievers in the Control Group progressed significantly only during the first two time periods tested, from T0 to T1 ($p= <.0001$) and from T1 to T2 ($p= 0.0104$), but their progress was not significant during the final period, from T2 to T3 ($p= 0.8982$). In spite of this, as in the case of High achievers in the CLIL group, their progression was also significant from T0 to T3 ($p= <.0001$).

Regarding Low achievers in the CLIL group, they progressed significantly from T0 to T1 ($p= 0.0004$) and from T2 to T3 ($p= <.0001$). Their counterparts, Low achievers in the Control Group only improved significantly from T0 to T1 ($p= <.0001$). Nevertheless, there were significant differences from T0 to T3 for both the CLIL ($p= <.0001$) and the Control groups ($p= <.0001$). As has already been explained in the achievement report, Low achievers in the CLIL group, obtained better mean scores at

T3 (53.6%) than Low achievers in the Control Group (48.7%) even though the difference was not significant.

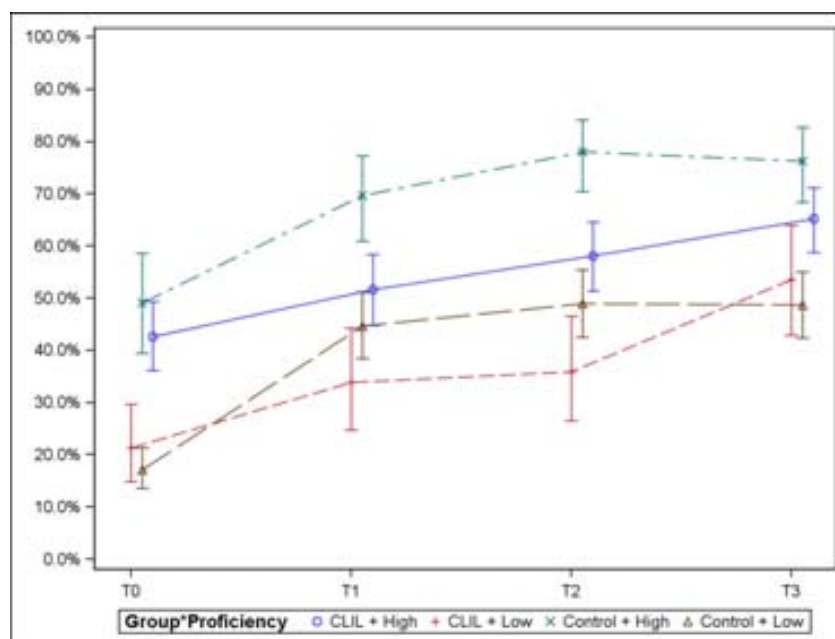


Figure 5.33 Reading improvement results: Group /Proficiency level interaction

Proficiency	Group	I0_1	I1_2	I2_3	I0_3
Low	CLIL	58.7%	6.2%	49.1%	151.3%
	Control	161.1%	9.5%	-0.6%	184.4%
High	CLIL	21.2%	12.6%	12.3%	53.2%
	Control	42.1%	11.9%	-2.2%	55.4%

Table 5.103 Percentage of improvement in reading. Group/Proficiency level interaction

So far, the total results in listening and reading achievement and improvement have been presented. In the next sections, the achievement and improvement results for each of the four questions involved in testing reading will be reported.

5.2.2.1 Reading question 1

In Reading question 1, students were asked to read and match a definition with the corresponding word. Even though the differences were not statistically significant at T0, as can be seen in Table 5.104, the mean results for the CLIL group were slightly

higher than those for the Control Group. However, achievement results at the rest of the times tested were all significantly different in favour of the Control Group: T1 ($p=0.0078$), T2 ($p=0.0004$) and T3 ($p=0.0095$). T2 displayed the highest mean difference between the CLIL and the Control Group (24.7%), whereas the mean differences at T1 and T3 were between 15% and 18%.

	Group	Mean	StdErr	Lower	Upper	F Value	P Value
T0	CLIL	31.1%	4.3%	23.2%	40.2%	2.03	0.1559
	Control	23.0%	3.6%	16.7%	30.9%		
T1	CLIL	42.5%	4.9%	33.1%	52.3%	7.26	0.0078
	Control	61.1%	4.6%	51.7%	69.8%		
T2	CLIL	38.5%	4.8%	29.6%	48.3%	12.92	0.0004
	Control	63.2%	4.5%	53.9%	71.6%		
T3	CLIL	59.1%	4.7%	49.6%	67.9%	6.87	0.0095
	Control	75.3%	3.9%	66.9%	82.2%		

Table 5.104: Reading achievement results: Reading question 1

The results of the intragroup comparison displayed significant improvements during the same time periods for both groups. Although the Control Group started from a lower mean percentage, it improved significantly from T0 to T1 ($p<.0001$) and from T2 to T3 ($p=0.0199$). The CLIL group progressed significantly at exactly the same time periods: T0-T1 ($p=0.0435$) and T2-T3 ($p<.0001$). However, as has already been pointed out, the Control Group reached higher achievement percentages at the end of the study (T3) in this question. As can be seen in the graph below, both groups progressed from T0 to T3 and their progress was statistically significant (CLIL group $p<.0001$ / Control Group $p<.0001$) (see Figure 5.34 and Table 5.105).

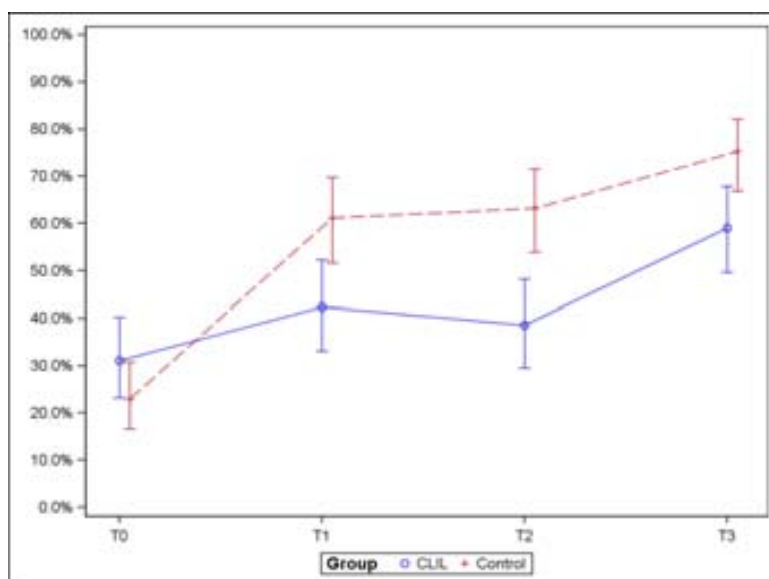


Figure 5.34 Reading improvement results. Reading Question 1

	I0_1	I1_2	I2_3	I0_3
CLIL	36.6%	-9.2%	53.3%	90.1%
Control	165.6%	3.4%	19.1%	227.1%

Table 5.105 Percentage of improvement: Reading question1

5.2.2.2 Reading question 2

Reading question 2 was a True/False reading comprehension task. Students were asked to read five statements and to write *Yes* or *No* considering what they could see in the picture accompanying the task. Table 5.106 presents statistics for question 2 in the reading test. The initial results at T0 showed no significant differences between the groups. Although at T1 the Control Group outperformed the CLIL Group, results were not significantly different. However, there were significant differences at T2 in favour of the Control Group ($F= 6.35$ $p= 0.0122$), which, descriptively, also seemed to outperform the CLIL Group at T3. The mean difference between the groups at the end of the study, at T3, was very low, around 3% (see Table 5.106).

	Group	Mean	StdErr	Lower	Upper	F Value	P Value
T0	CLIL	59.0%	3.8%	51.4%	66.2%	0.06	0.8041
	Control	60.3%	3.6%	53.0%	67.2%		
T1	CLIL	69.0%	3.7%	61.4%	75.8%	0.76	0.3828
	Control	73.3%	3.2%	66.6%	79.1%		
T2	CLIL	57.3%	4.0%	49.3%	64.9%	6.35	0.0122
	Control	70.8%	3.5%	63.5%	77.2%		
T3	CLIL	64.6%	3.7%	57.1%	71.4%	0.40	0.5273
	Control	67.8%	3.5%	60.6%	74.3%		

Table 5.106 Reading achievement results: Reading question 2

Figure 5.35 and Table 5.107 display the improvement made by both groups during the different time periods. The Control Group improved significantly from T0 to T1 ($p= 0.0146$). Then it slightly decreased from T1 to T2 and from T2 to T3. The highest mean result for this group occurred at T1 (73.3%) and not at T3. There were no significant differences in the progression of the CLIL group during the study. As has already been pointed out in the achievement report, the mean difference in achievement at T3 between the two groups was around 3% and even though it was not significantly different, the difference was in favour of the Control Group.

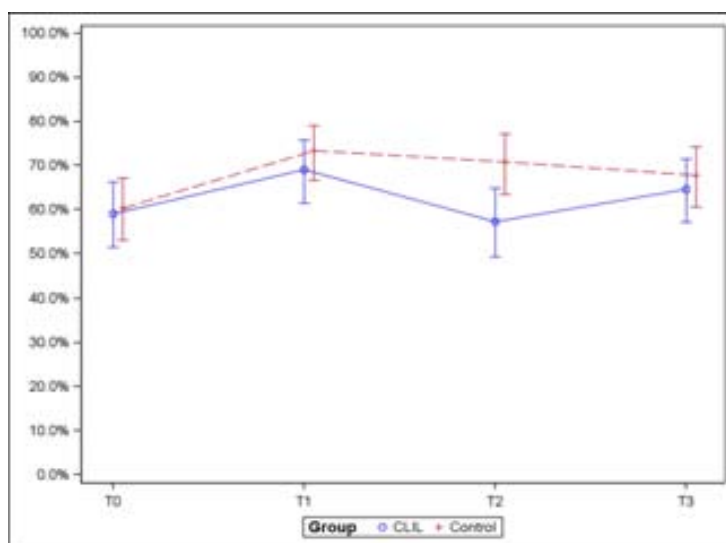


Figure 5.35 Reading improvement results. Reading Question 2

	I0_1	I1_2	I2_3	I0_3
CLIL	17.0%	-17.0%	12.7%	9.4%
Control	21.5%	-3.4%	-4.3%	12.4%

Table 5.107 Percentage of improvement: Reading Question 2

5.2.2.3 Reading question 3

Reading question 3 was a gap filling exercise which consisted of a text with some missing words. For each gap, students had a choice of three words. They had to decide which word to use and copy it in the gap. This exercise was a cloze type test designed to test the students' use of English.

As Table 5.108 displays, there were no statistically significant differences at any of the times tested for any of the groups. Descriptive statistics, however, showed an advantage for the Control Group at T0, T1 and T2. At T3, the CLIL group seemed to catch up with the Control Group. The mean results at T3, although not statistically different, were almost the same for both groups. The difference was around 1.1 %.

	Group	Mean	StdErr	Lower	Upper	F Value	P Value
T0	CLIL	32.8%	4.0%	25.4%	41.2%	1.20	0.2735
	Control	39.1%	3.9%	31.7%	47.1%		
T1	CLIL	39.0%	4.5%	30.6%	48.0%	2.48	0.1162
	Control	48.7%	4.1%	40.6%	56.8%		
T2	CLIL	51.3%	4.5%	42.5%	60.0%	0.65	0.4219
	Control	56.3%	4.3%	47.7%	64.6%		
T3	CLIL	58.1%	4.3%	49.6%	66.2%	0.03	0.8570
	Control	59.2%	4.2%	50.8%	67.2%		

Table 5.108 Reading achievement results. Reading Question 3

The intragroup comparison showed a similar linear progression for the groups. None of the percentages in improvement was statistically significant. The CLIL group

progressed slightly more from T1 to T2 whereas the Control Group seemed to progress more from T0 to T1. As has been pointed out in the description of the achievement results, the groups reached almost the same level at T3. Nevertheless, both groups improved significantly throughout the study (T0-T3): CLIL Group $p = <.0001$, Control group $p = 0.0014$) (see Figure 5.36 and Table 5.109).

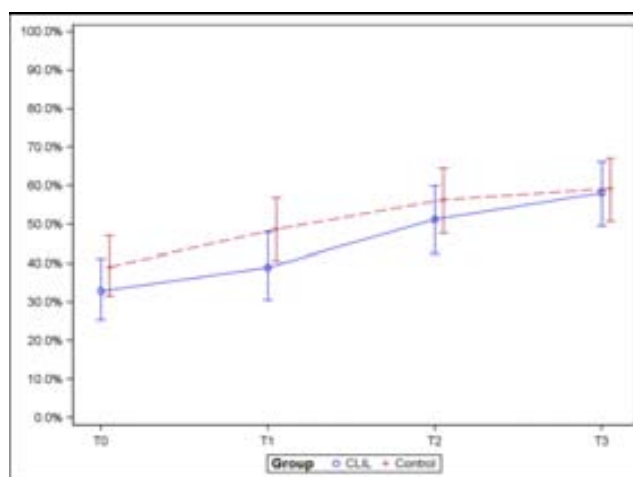


Figure 5.36 Reading improvement results. Reading Question 3

	I0_1	I1_2	I2_3	I0_3
CLIL	18.7%	31.6%	13.3%	77.0%
Control	24.5%	15.7%	5.2%	51.5%

Table 5.109 Percentage of improvement: Reading Question 3

5.2.2.4 Reading question 4 (Comprehension)

This question was a story in three parts. Each part of the story had a matching picture. After each part, the students had to complete 10 sentences about the story using one, two or three words. This task was scored under two different categories: comprehension and accuracy. This section will deal with the results on comprehension and the next section will address the results on accuracy.

As shown in Table 5.110, significant differences were found in the intergroup comparisons between CLIL and Non-CLIL learners in this particular task at T1 and T2 in favour of the Control Group (T1 $F= 11.21$ $p= 0.0010$ / T2 $F= 9.59$ $p= 0.0023$). Descriptive statistics showed, however, that at T3 both groups reached almost the same level of achievement. There was only 1.1% difference between the groups.

	Group	Mean	StdErr	Lower	Upper	F Value	P=
T0	CLIL	23.1%	3.6%	16.8%	30.8%	0.00	0.9852
	Control	23.2%	3.4%	17.1%	30.6%		
T1	CLIL	41.2%	4.6%	32.6%	50.4%	11.21	0.0010
	Control	62.6%	4.2%	54.1%	70.4%		
T2	CLIL	49.2%	4.7%	40.1%	58.4%	9.59	0.0023
	Control	68.6%	3.9%	60.4%	75.8%		
T3	CLIL	61.8%	4.3%	53.0%	69.9%	0.03	0.8542
	Control	62.9%	4.1%	54.6%	70.5%		

Table 5.110 Reading achievement results. Reading Question 4 (Comprehension)

Figure 5.37 shows the improvement made by both groups. The Control Group improved significantly from T0 to T1 ($p= <.0001$). The mean scores for this group continued to increase from T1 to T2 and decreased slightly from T2 to T3. The progress made by the CLIL group was significant from T0 to T1 ($p= 0.0001$) and from T2 to T3 ($p= 0.0224$) but not from T1 to T2. The progress throughout the study, T0 to T3, was statistically significant for the CLIL ($p= <.0001$) and the Control groups ($p= <.0001$). This last group improved slightly more (171.6%) than the CLIL one (167.9%).

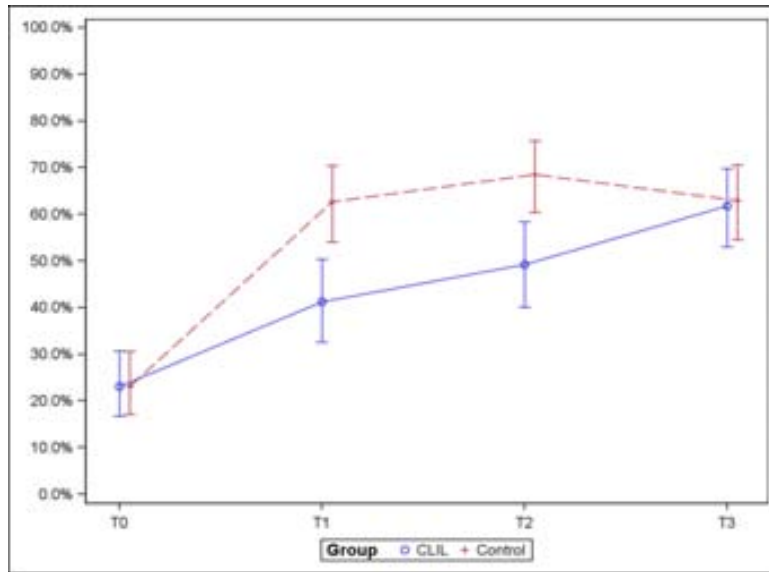


Figure 5.37 Reading improvement results. Reading Question 4 (Comprehension)

	I0_1	I1_2	I2_3	I0_3
CLIL	78.7%	19.4%	25.6%	167.9%
Control	170.5%	9.5%	-8.3%	171.6%

Table 5.111: Percentage of improvement: Reading Question 4 (Comprehension)

5.2.2.5 Reading question 4 (Accuracy)

The mean results obtained in reading question 4 (Accuracy) are, in general terms, lower than those obtained in Reading Question 4 when this question was assessed for Comprehension. At T1 and T2, achievement results are significantly different in favour of the Control Group (T1 $F=12.72$ $p= 0.0003$ / T2 $F= 13.85$ $p= 0.0003$), which at T0 had obtained a lower score. Descriptive statistics show, however, that at T3 both groups reached the same mean results in achievement with a slightly minor advantage for the Control Group (see Table 5.112).

	Group	Mean	StdErr	Lower	Upper	F Value	P Value
T0	CLIL	16.8%	3.0%	11.7%	23.6%	1.37	0.2426
	Control	12.3%	2.4%	8.3%	17.9%		
T1	CLIL	25.3%	3.6%	18.9%	33.0%	13.72	0.0003
	Control	46.1%	4.0%	38.3%	54.1%		

	Group	Mean	StdErr	Lower	Upper	F Value	P Value
T2	CLIL	42.7%	4.4%	34.4%	51.5%	13.85	0.0003
	Control	65.5%	3.9%	57.3%	72.8%		
T3	CLIL	55.2%	4.2%	46.8%	63.4%	0.00	0.9507
	Control	55.6%	4.1%	47.5%	63.4%		

Table 5.112 Reading achievement results. Reading Question 4 (Accuracy)

As for improvement, the Control Group improved significantly from T0 to T1 ($p < .0001$) and from T1 to T2 ($p < .0001$), but from T2 to T3 it experienced a decrease. The CLIL group, however, improved at all times tested, but only the improvements from T1 to T2 ($p = 0.0002$) and T2 to T3 ($p = 0.0214$) were statistically significant. At T3, as explained before, both groups reached the same level in terms of achievement. The progression T0 to T3 was significant for both groups: Control group $p < .0001$, CLIL group $p < .0001$.

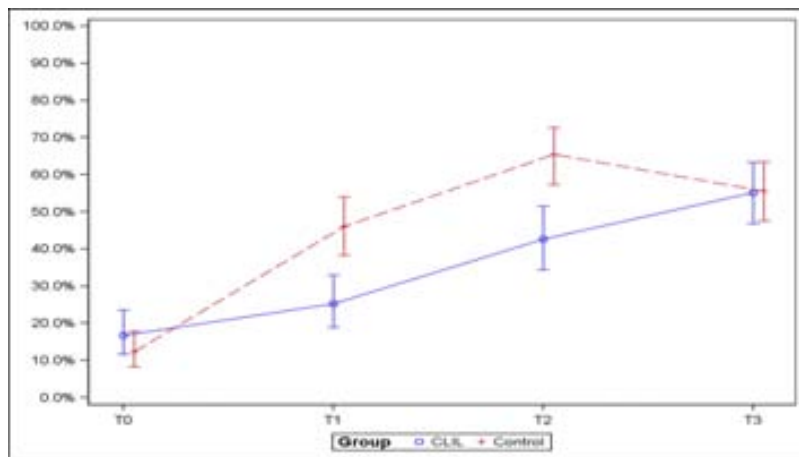


Figure 5.38 Reading improvement results. Reading Question 4 (Accuracy)

	I0_1	I1_2	I2_3	I0_3
CLIL	50.6%	68.7%	29.3%	228.6%
Control	274.8%	42.0%	-15.1%	352.2%

Table 5.113 Percentage of improvement: Reading Question 4 (Accuracy)

5.2.3 Summary of Arts & Crafts Listening and Reading Results

5.2.3.1 Summary of Listening and Reading Achievement Results

Table 5.114 shows a summary of the results of the intergroup comparison between the CLIL and the Control Groups in the listening and reading tests.

	Listening (total)		Reading (Total)	
	Control	CLIL	Control	CLIL
T0	29.0%	27.2%	30.09%	31.0%
	p= 0.6845		p= 0.9754	
T1	47.3%	42.9%	57.7%	42.5%
	p=0.3825		p= 0.0007	
T2	60.4%	46.2%	64.9%	46.9%
	p= 0.0060		p= <.0001	
T3	65.5%	50.6%	63.6%	59.6%
	p= 0.0029		p= 0.3327	

Table 5.114 Summary of Listening and Reading achievement results

Table 5.115 displays a summary of the intergroup comparison results for each of the reading questions.

Reading	Question1		Question2		Question 3		Question 4 (Comprehension)		Question 4 (Accuracy)	
	Control	CLIL	Control	CLIL	Control	CLIL	Control	CLIL	Control	CLIL
T0	23.0%	31.1%	60.3%	59.0%	39.1%	32.8%	23.2%	23.1%	12.3%	16.8%
	p= 0.1559		p= 0.8041		p 0.2735		p= 0.9852		p= 0.2426	
T1	61.1%	42.5%	73.3%	69.0%	48.7%	39.0%	62.6%	41.2%	46.1%	25.3%
	p= 0.0078		p= 0.3828		p= 0.1162		p 0.0010		p= 0.0003	
T2	63.2%	38.5%	70.8%	57.3%	56.3%	51.3%	68.6%	49.2%	65.5%	42.7%
	p= 0.0004		p= 0.0122		p 0.4219		p= 0.0023		p= 0.0003	
T3	75.3%	59.1%	67.8%	64.6%	59.2%	58.1%	62.9%	61.8%	55.6%	55.2%
	p= 0.0095		p= 0.5273		p= 0.8570		p= 0.8542		p= 0.9507	

Table 5.115 Summary of Reading achievement results: individual reading questions

Achievement results in Listening when the interaction Group/Proficiency level were taken into account are displayed in Table 5.116 below.

Listening	Group/Proficiency level			
	High		Low	
	Control	CLIL	Control	CLIL
T0	39.7%	41.3%	20.3%	16.6%
	p= 0.8092		p= 0.4661	
T1	65.4%	46.3%	29.9%	39.6%
	p= 0.0054		p= 0.1589	
T2	74.1%	55.0%	45.0	37.7%
	p= 0.0030		p= 0.3251	
T3	77.8%	57.4%	50.7%	43.8%
	p= 0.0015		p= 0.3328	

Table 5.116 Summary of Listening achievement results. Group/Proficiency level interaction

In terms of reading, Table 5.117 summarises the intergroup comparison when the variables Group/Proficiency level were taken into account.

Reading	Group/Proficiency level			
	High		Low	
	Control	CLIL	Control	CLIL
T0	49.1%	42.6%	17.1%	21.3%
	p= 0.2816		p= 0.3023	
T1	69.7%	51.6%	44.7%	33.8%
	p= 0.0020		p= 0.0774	
T2	78.0%	58.1%	49.0%	36.0%
	p= 0.0002		p= 0.0393	
T3	76.3%	65.3%	48.7%	53.6%
	p= 0.0309		p= 0.4387	

Table 5.117 Summary of Reading achievement results. Group/Proficiency level interaction

5.2.3.2 Summary of Listening and Reading Improvement Results

The progression of the Control Group and the CLIL group in terms of listening and reading during four different time periods can be seen in Table 5.118 below.

	Listening		Reading (Total)	
	Control	CLIL	Control	CLIL
T0-T1	63.1%	57.4%	87.1%	37.2%
	p= <.0001	p= 0.0009	p= <.0001	p= <.0001

Language Competence of Young Learners Exposed to EFL and CLIL

T1-T2	27.7% p= 0.0062	7.8% P= 0.8532	12.4% p= 0.0017	10.3% p= 0.1600
T2-T3	8.4% p= 0.5626	9.5% p= 0.7121	-2.0% p= 0.9137	27.1% p=<.0001
T0-T3	125.8% p=<.0001.	85.9% p=<.0001	101.6% p=<.0001	92.3% p=<.0001

Table 5.118 Summary of Listening and Reading Improvement Results

Table 5.119 and Table 5.120 below show the improvement of the Control and CLIL groups the different reading questions at different time periods.

Reading	Question1		Question2		Question 3	
	Control	CLIL	Control	CLIL	Control	CLIL
T0-T1	165.6% p=<.0001	36.6% p=0.0435	21.5% p=0.0146	17.0% p=0.1456	24.5% p=0.2640	18.7% p=0.6869
T1-T2	3.4% p=0.9631	-9.2% p=0.8111	-3.4% p=0.9343	-17.0% p=0.0725	15.7% p=0.5091	31.6% p=0.1551
T2-T3	19.1% p=0.0199	53.3% p=<.0001	-4.3% p=0.9044	12.7% p=0.4216	5.2% p=0.9534	13.3% p=0.6250
T0-T3	227.1% p=<.0001	90.1% p=<.0001	12.4% p=0.6253	9.4% p=0.3412	51.5% p=0.014	77.0% p=<.0001

Table 5.119 Summary of reading improvement results: reading questions 1, 2, and 3

Reading	Question 4 (Comprehension)		Question 4 (Accuracy)	
	Control	CLIL	Control	CLIL
T0-T1	170.5% p=<.0001	78.7% p=0.1520	274.8% p=<.0001	50.6% p=0.1009
T1-T2	9.5% p=0.4476	19.4% p=0.6074	42.0% p=<.0001	68.7% p=0.0002
T2-T3	-8.3% p=0.04671	25.6% p=0.0094	-15.1% p=0.0751	29.3% p=0.0214
T0-T3	171.6% p=<.0001	167.9% p=<.0001	352.2% p=<.0001	228.6% p=<.0001

Table 5.120 Summary of reading improvement results: reading questions 4 and 5

Table 5.121 below displays, the results in the listening test when the variables Group/Proficiency level interacted.

Listening	Group/Proficiency level			
	High		Low	
	Control	CLIL	Control	CLIL
T0-T1	64.9% p= 0.0003	11.9% p= 06811	47.5% p= 0.0310	138.2% p= 0.0018
T1-T2	13.2% p= 0.0015	18.8% p= 0.2121	50.6% p= 0.0015	-4.6% p= 0.9940
T2-T3	5.1% p= 0.5312	4.5% p= 0.9403	12.7% p= 0.5312	16.0% p= 0.8228
T0-T3	96.1% p= <.0001	39.0% p= 0.0016	150.5% p= <.0001	163.7% p= 0.0001

Table 5.121 Summary of the Listening improvement results. Group/Proficiency level interaction

The interaction Group/Proficiency level, as far as reading improvement is concerned, is summarised in Table 5.122 below.

Reading	Group/Proficiency level			
	High		Low	
	Control	CLIL	Control	CLIL
T0-T1	42.1% p=<.0001	21.2% p=0.0004	161.1% p=<.0001	58.7% p=0.0004
T1-T2	11.9% p=0.0104	12.6% p=0.0192	9.5% p=0.1832	6.2% p=0.9266
T2-T3	-2.2% p=0.8982	12.3% p=0.0049	-0.6% p=0.9992	49.1% p=<.0001
T0- T3	55.4% p=<.0001	53.2% p=<.0001	184.4% p=<.0001	151.3% p=<.0001

Table 5.122 Summary of the Interaction Group/Proficiency. Reading improvement results

5.2.4 Writing

5.2.4.1 Fluency

This section shows the results obtained by the Control and the CLIL Arts & Crafts groups in Fluency. Four different measures were used to analyse it: Total

Number of Words (TNW), Total Number of Words in English (TNWE), Total Number of Units (TNU) as well as the ratio between the Total Number of Words in English in relation to the Total Number of Words written by the students (TNWE/TNW). For two of the measures, TNWE and TNU, the results of the interaction between High and Low achievers will also be reported.

5.2.4.1.1 Total Number of Words

The intergroup comparison did not show statistically significant differences between the CLIL and the Control Groups at any of the times tested. Descriptive statistics, however, showed an advantage in the number of words written by students at T1 and T2 in favour of the Control Group. Nevertheless, at the end of the study (T3), the CLIL group wrote an average of 110 words as opposed to an average of 99.3 words written by the Control group.

	Group	Mean	StdErr	Lower	Upper	F Value	P Value
T0	CLIL	65.26	5.34	54.73	75.80	0.06	0.8047
	Control	63.47	4.90	53.80	73.14		
T1	CLIL	76.23	5.34	65.70	86.76	0.72	0.3981
	Control	82.37	4.90	72.70	92.04		
T2	CLIL	98.85	5.34	88.32	109.4	2.94	0.0883
	Control	86.42	4.90	76.75	96.09		
T3	CLIL	110.8	5.49	99.99	121.6	2.44	0.1199
	Control	99.31	4.90	89.64	109.0		

Table 5.123 Fluency achievement Results: TNW

The analysis of the improvement made by the two groups showed significant differences in improvement for the Control Group ($p= 0.0009$) during the first time period, T0 to T1, but no differences for the CLIL group. From T1 to T2, the CLIL Group improved significantly ($p= 0.0002$), however, the Control Group showed very

little progress which was not significant. Although both groups progressed from T2 to T3, improvement was only statistically significant for the Control Group ($p= 0.0459$). Nevertheless, as has been explained in the previous section, at T2 and T3, the mean scores were higher for the CLIL group. The progress made by the groups from T0 to T3 was significant for both of them: Control Group $p= <.0001$, CLIL group $p=<.0001$ (see Figure 5.39 and Table 5.124).

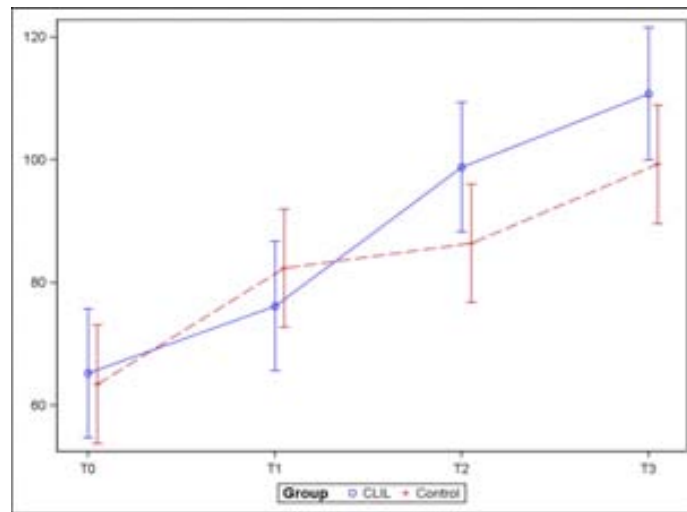


Figure 5.39 Fluency improvement results: TNW

	I0_1	I1_2	I2_3	I0_3
CLIL	16.8%	29.7%	12.1%	69.8%
Control	29.8%	4.9%	14.9%	56.5%

Table 5.124 Percentage of improvement in Fluency: TNW

The intergroup comparison, when the variables Group /High achievers were taken into account, displayed no significant differences for the High achievers at any of the times. At T0, the groups started at the same level and at the end of the first year, the High achievers in the Control Group outperformed those in CLIL. Nevertheless, at T2 and T3, the mean average number of words was higher for the CLIL group, although the difference was not significant (see Table 5.125).

	Proficiency	Group	Mean	StdErr	Lower	Upper	F Value	P Value
T0	High	CLIL	69.96	5.97	58.18	81.74	0.00	0.9461
	High	Control	69.28	8.06	53.38	85.18		
T1	High	CLIL	77.79	5.97	66.01	89.57	0.19	0.6642
	High	Control	82.18	8.06	66.27	98.08		
T2	High	CLIL	107.0	5.97	95.26	118.8	3.85	0.0511
	High	Control	87.26	8.06	71.36	103.2		
T3	High	CLIL	118.9	6.06	107.0	130.9	1.51	0.2201
	High	Control	106.4	8.06	90.54	122.3		

Table 5.125 Fluency achievement results: TNW. Group/High achievers interaction

As can be seen in Table 5.126, no significant differences were found for Low achievers at any of the times tested. CLIL Low achievers started at a slightly higher point at T0. As in the case of High achievers, at T1, descriptively, they were behind Control Low achievers. However, at T2 and T3, the mean score in the TNW was higher for the students in CLIL than for those in the Control group.

	Proficiency	Group	Mean	StdErr	Lower	Upper	F Value	P Value
T0	Low	CLIL	60.57	9.32	42.18	78.95	0.07	0.7865
	Low	Control	57.65	5.34	47.12	68.19		
T1	Low	CLIL	74.66	9.32	56.27	93.05	0.54	0.4622
	Low	Control	82.57	5.34	72.03	93.10		
T2	Low	CLIL	90.65	9.32	72.27	109.0	0.22	0.6364
	Low	Control	85.57	5.34	75.04	96.11		
T3	Low	CLIL	102.7	9.55	83.88	121.6	0.93	0.3365
	Low	Control	92.18	5.34	81.64	102.7		

Table 5.126 Fluency achievement results: TNW. Group/Low achievers interaction

5.2.4.1.2 Total Number of Words in English

As for the TNWE, no significant differences were found at T0, T1 and T2 between the Control and the CLIL groups. Descriptive statistics showed that at T1 the

mean number of words in English was the same for both groups. However, at T2, the mean of the CLIL Group was higher than that of the Control Group. At T3, the CLIL group significantly outperformed the Control Group ($F= 4.24$ $p= 0.0409$) (see Table 5.127).

	Group	Mean	StdErr	Lower	Upper	F Value	P Value
T0	CLIL	55.31	4.94	45.56	65.06	0.68	0.4101
	Control	49.77	4.53	40.82	58.72		
T1	CLIL	67.16	4.94	57.42	76.91	0.01	0.9186
	Control	67.85	4.53	58.90	76.80		
T2	CLIL	84.73	4.94	74.98	94.48	2.69	0.1027
	Control	73.72	4.53	64.77	82.67		
T3	CLIL	100.0	5.06	90.04	110.0	4.24	0.0409
	Control	86.03	4.53	77.08	94.98		

Table 5.127: Fluency achievement results: TNWE

Both groups progressed during different time periods although the progress was not always statistically significant. From T0 to T1 there were statistically significant differences for the Control Group ($p= 0.0002$), but no differences for the CLIL one. From T1 to T2 it was the CLIL group the one that improved significantly ($p= 0.0013$). From T2 to T3 both groups showed a significant improvement: Control Group $p= 0.0244$; CLIL Group $p= 0.0094$. However, as has already been said, the group with the highest significant achievement result in the TNWE was the CLIL Group at T3. Considering the progress made by the groups throughout the study, both groups progressed significantly from T0 to T3: Control Group $p=<.0001$, CLIL group $p=<.0001$ (see Table 5.128 and Figure 5.40).

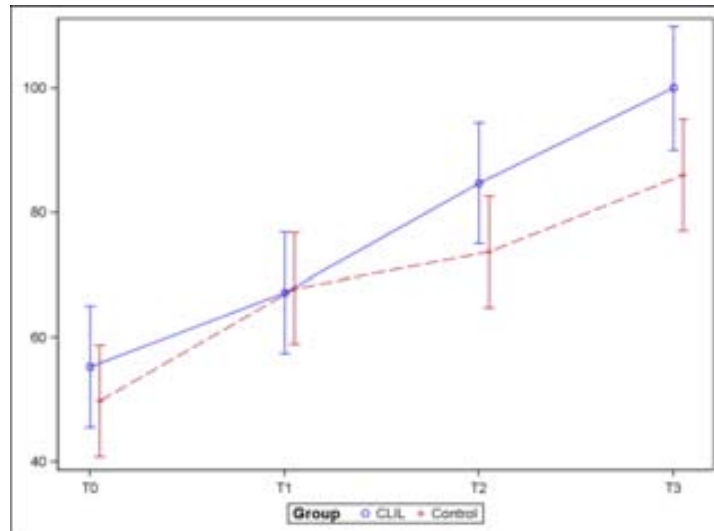


Figure 5.40 Fluency improvement results: TNWE

	I0_1	I1_2	I2_3	I0_3
CLIL	21.4%	26.2%	18.1%	80.9%
Control	36.3%	8.7%	16.7%	72.9%

Table 5.128 Percentage of improvement in Fluency: TNWE

Achievement results as for the TNWE when the interaction Group/High achievers was taken into account can be seen in Table 5.129 below. No significant differences were found for High achievers at any of the times. Descriptive statistics, however, showed that, even though at T0 and T1 the mean scores for both groups were very similar, at T2 and T3, the CLIL High achievers attained better mean scores.

	Proficiency	Group	Mean	StdErr	Lower	Upper	F Value	P Value
T0	High	CLIL	58.74	5.52	47.84	69.65	0.05	0.8250
	High	Control	56.68	7.45	41.96	71.40		
T1	High	CLIL	69.07	5.52	58.17	79.98	0.01	0.9209
	High	Control	70.00	7.45	55.28	84.72		
T2	High	CLIL	93.45	5.52	82.54	104.3	2.45	0.1196
	High	Control	78.87	7.45	64.15	93.59		
T3	High	CLIL	107.4	5.60	96.36	118.4	1.64	0.2021

Proficiency	Group	Mean	StdErr	Lower	Upper	F Value	P Value
High	Control	95.42	7.46	80.70	110.1		

Table 5.129 Fluency achievement results: TNWE. Group/High achievers interaction

As in the case of High achievers, no significant differences were found for Low achievers in any of the groups. Descriptively, the scores at T1 were the same for the Control and the CLIL groups. At T2 and T3, Low achievers in the CLIL group wrote, on average, more words in English (92.66) than their counterparts in the Control Groups (76.64). It is interesting to notice that at T2, the score of the Control High (78.87) is very similar to that of the CLIL Low (76.02) and that at T3, this difference is even smaller: Control High (95.42) and CLIL Low (92.66) (see Table 5.130).

	Proficiency	Group	Mean	StdErr	Lower	Upper	F Value	PValue
T0	Low	CLIL	51.87	8.62	34.86	68.89	0.83	0.3649
	Low	Control	42.86	4.94	33.10	52.61		
T1	Low	CLIL	65.25	8.62	48.23	82.27	0.00	0.9642
	Low	Control	65.70	4.94	55.95	75.45		
T2	Low	CLIL	76.02	8.62	59.00	93.03	0.56	0.4546
	Low	Control	68.58	4.94	58.82	78.33		
T3	Low	CLIL	92.66	8.82	75.27	110.1	2.52	0.1145
	Low	Control	76.64	4.94	66.89	86.39		

Table 5.130 Fluency achievement results: TNWE. Group/Low achievers interaction

When the progress for High and Low achievers was analysed, the results were statistically significant for High achievers in the CLIL group from T1 to T2 ($p < .0001$) and from T2 to T3 ($p = 0.0465$), whereas High achievers in the Control Group did not improve significantly during any of the time periods tested. The improvement of Low achievers in the CLIL group was not significant whereas there were significant differences in the progress of Low achievers in the Control Group during the first time

period, T0 to T1 ($p = <.0001$). Improvement from T0 to T3 was significant for all the groups ($p = <.0001$) (see Figure 5.41 and Table 131).

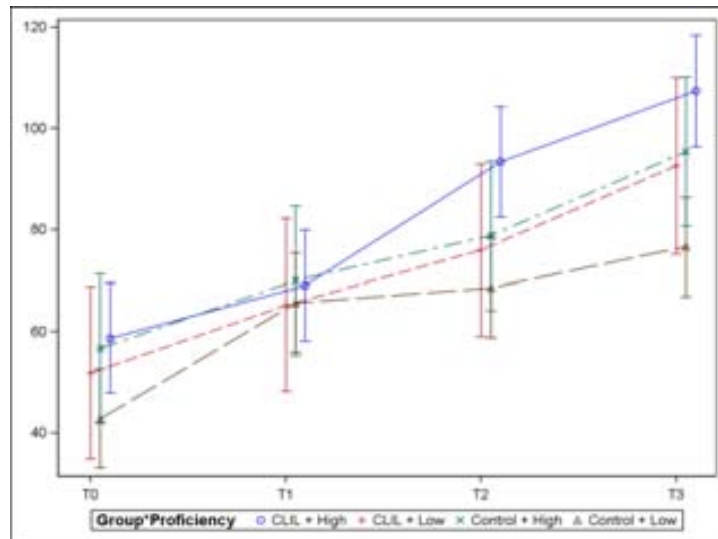


Figure 5.41 Fluency improvement results: Group/Proficiency level interaction

Proficiency	Group	I0_1	I1_2	I2_3	I0-3
Low	CLIL	25.8%	16.5%	21.9%	78.6%
	Control	53.3%	4.4%	11.8%	78.8%
High	CLIL	17.6%	35.3%	14.9%	82.8%
	Control	23.5%	12.7%	21.0%	68.3%

Table 5.131 Percentages of improvement in Fluency: TNWE. Group/Proficiency level interaction

5.2.4.1.3 Ratio: Total Number of Words in English/Total Number of Words

As a measure of Fluency, the ratio between the Total Number of Words in English and the Total Number of Words was also calculated. As seen in Table 5.132, the results for the CLIL Group were significantly different from those of the Control group at T0 ($P = 0.0425$). Yet, achievement results at T1, T2 and T3 did not yield any significant differences between the groups. However, descriptive statistics show an advantage for the CLIL group in the mean percentage of words written in English at all the times.

	Group	Mean	StdErr	Lower	Upper	F Value	P Value
T0	CLIL	83.8%	2.4%	79.2%	88.7%	4.16	0.0425
	Control	77.4%	2.1%	73.5%	81.6%		
T1	CLIL	86.9%	2.3%	82.4%	91.6%	3.58	0.0598
	Control	81.2%	2.0%	77.4%	85.2%		
T2	CLIL	85.1%	2.1%	81.0%	89.4%	0.18	0.6755
	Control	83.9%	2.0%	80.0%	87.9%		
T3	CLIL	89.8%	2.2%	85.5%	94.3%	2.43	0.1208
	Control	85.2%	1.9%	81.5%	89.1%		

Table 5.132 Fluency achievement results: %TNWE/TNW

The intragroup comparison carried out to determine the individual progress of each group showed no statistically significant differences at any of the time periods tested as for the TNWE. The improvement of the Control Group seemed to be much more linear than that of the CLIL Group, whose progression decreased slightly from T1 to T2. As pointed out in the achievement results, the difference in achievement at T3 was not statistically significant (see Figure 5.42 and Table 5.133).

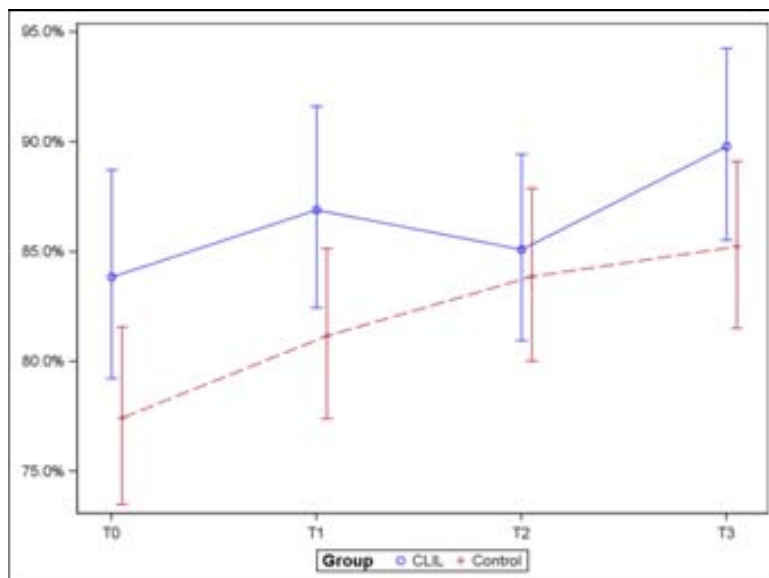


Figure 5.42 Fluency improvement results: %TNWE/TNW

	I0_1	I1_2	I2_3	I0_3
CLIL	3.6%	-2.1%	5.5%	7.1%
Control	4.8%	3.3%	1.6%	10.1%

Table 5.133 Percentage of improvement in Fluency: %TNWE/TNW

5.2.4.1.4 Total Number of Units

The number of Units, that is, the number of meaningful chunks of language that contain a finite or non-finite verb, was taken as a measure of Fluency. As can be seen in Table 5.134, statistically significant differences were found in favour of the CLIL group at T2 ($F= 14.00$ $p= 0.0002$) and T3 ($F= 15.60$ $p= 0.0001$). At T3, CLIL students wrote an average of 18.07 units per essay as opposed to an average of 13.22 in the Control Group.

	Group	Mu	StdErr	Lower	Upper	F Value	P Value
T0	CLIL	8.75	0.89	6.99	10.51	1.51	0.2211
	Control	7.26	0.82	5.65	8.88		
T1	CLIL	9.53	0.89	7.77	11.29	0.58	0.4464
	Control	10.46	0.82	8.84	12.08		
T2	CLIL	16.39	0.89	14.63	18.15	14.00	0.0002
	Control	11.85	0.82	10.24	13.47		
T3	CLIL	18.07	0.91	16.27	19.88	15.60	0.0001
	Control	13.22	0.82	11.61	14.84		

Table 5.134 Fluency achievement results: TNU

In terms of progress, both groups improved during the different time periods tested. From T0 to T1, the Control Group showed statistically significant differences in its progress ($p= 0.0004$). During the next time period, from T1 to T2, only the CLIL group increased significantly ($p= <.0001$). During the second year of the study, from T2 to T3, no significant differences were found for any of the groups. However, as has already been pointed out in the achievement results, the CLIL Group significantly

outperformed the Control Group at times T2 and T3. The progress throughout the study, T0 to T3, was significant for all the groups: Control Group $p = <.0001$, CLIL group $p = <.0001$. The improvement percentage for the CLIL group was higher than that of the Control group.

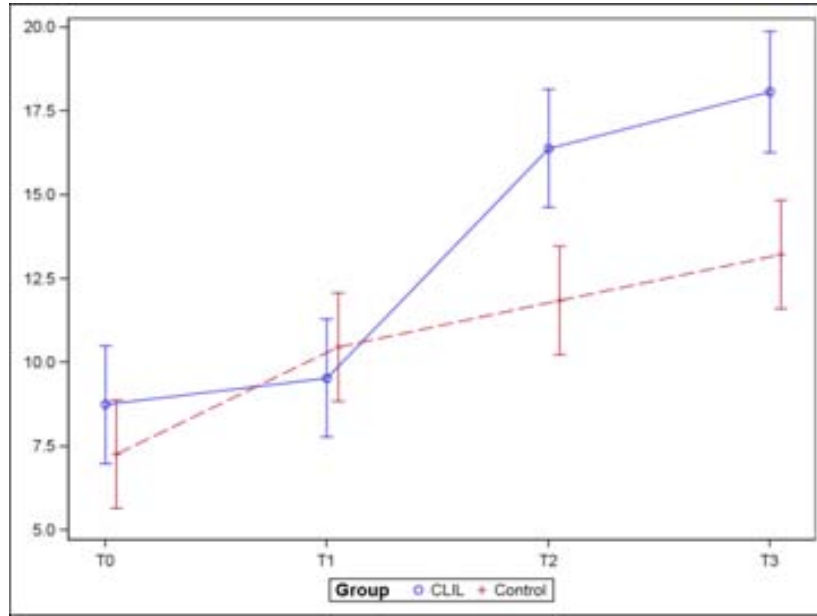


Figure 5.43 Fluency improvement results: TNU

	I0_1	I1_2	I2_3	I0_3
CLIL	8.9%	71.9%	10.3%	106.5%
Control	44.0%	13.3%	11.5%	82.0%

Table 5.135 Percentage of improvement in Fluency: TNU

The results of the intergroup comparison when the interaction Group/High achiever was taken into account can be seen in Table 5.136 below. There were statistically significant differences in favour of the High achievers in the CLIL group at T2 ($F = 5.27$ $p = 0.0229$) and at T3 ($F = 8.41$ $p = 0.0042$).

	Proficiency	Group	Mean	StdErr	Lower	Upper	F Value	P Value
T0	High	CLIL	9.86	1.00	7.89	11.83	0.77	0.3826
	High	Control	8.39	1.35	5.73	11.05		

	Proficiency	Group	Mean	StdErr	Lower	Upper	F Value	P Value
T1	High	CLIL	11.02	1.00	9.05	12.99	0.03	0.8718
	High	Control	10.75	1.35	8.09	13.41		
T2	High	CLIL	16.77	1.00	14.80	18.74	5.27	0.0229
	High	Control	12.91	1.35	10.25	15.57		
T3	High	CLIL	19.72	1.01	17.72	21.71	8.41	0.0042
	High	Control	14.81	1.35	12.15	17.47		

Table 5.136 Fluency achievement results: TNU. Group/High achievers interaction

As for Low achievers, the intergroup comparison showed significant differences in favour of the CLIL group at T2 ($F= 8.44$ $p= 0.0042$) and at T3 ($F= 6.92$ $p= 0.0093$). It is interesting to notice that CLIL students in Arts & Crafts wrote more units than their counterparts in the Control Group: the highest mean number of units was written by High achievers in the CLIL group (19.72) and the second highest number of mean units was written by Low achievers also in the CLIL group (16.43).

	Proficiency	Group	Mean	StdErr	Lower	Upper	F Value	P Value
T0	Low	CLIL	7.64	1.56	4.57	10.71	0.70	0.4028
	Low	Control	6.14	0.89	4.38	7.90		
T1	Low	CLIL	8.05	1.56	4.97	11.12	1.40	0.2383
	Low	Control	10.17	0.89	8.41	11.93		
T2	Low	CLIL	16.00	1.56	12.93	19.08	8.44	0.0042
	Low	Control	10.80	0.89	9.03	12.56		
T3	Low	CLIL	16.43	1.59	13.29	19.58	6.92	0.0093
	Low	Control	11.63	0.89	9.87	13.39		

Table 5.137 Fluency achievement results: TNU. Group/Low achievers interaction

In terms of progress, CLIL High achievers improved significantly during the second ($p= <.0001$) and the third time periods ($p= 0.0149$). No significant differences were found in the progress of High achievers in the Control Group. CLIL Low

achievers improved significantly during the second time period T1- T2 ($p = <.0001$) whereas Low achievers in the Control Group only improved significantly during the first time period analysed, T0-T1 ($p = <.0001$). Improvement from T0 to T3 was significant for all the groups ($p = <.0001$). As has already been reported in the achievement section, CLIL High and Low achievers attained their highest level of achievement in terms of number of units at T3.

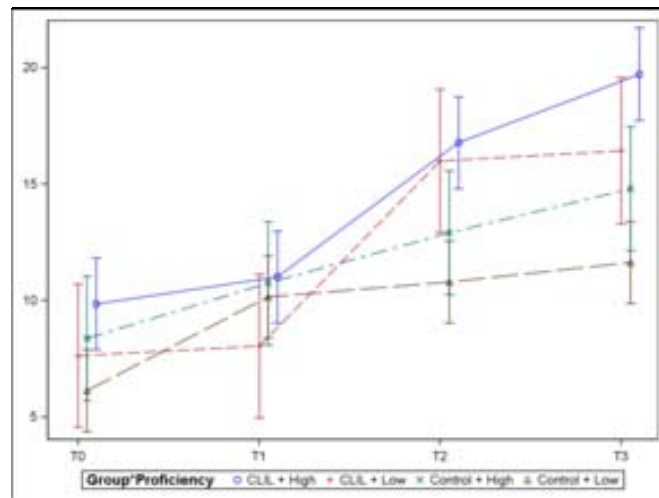


Figure 5.44 Fluency improvement results: TNU. Group/Proficiency level interaction

Proficiency	Group	I0_1	I1_2	I2_3	I0-3
Low	CLIL	5.3%	98.8%	2.7%	115.1%
	Control	65.7%	6.1%	7.8%	89.6%
High	CLIL	11.7%	52.2%	17.5%	99.9%
	Control	28.1%	20.1%	14.7%	76.6%

Table 5.138 Percentages of improvement in Fluency: TNU. Group/Proficiency level Interaction

5.2.4.2. Accuracy

Two measures were used to assess Accuracy: Total Number of Error Free Units (TNEFU) and the ratio between Total Number of Error Free Units in relation to Total Number of Units (%TNEFU/TNU).

5.2.4.2.1 Total Number of Error Free Units

At T0, the CLIL group statistically outperformed the Control Group ($F= 4.87$ $p=0.02839$). However, at T1, the Control Group and the CLIL group attained a very similar result. At T2 and at T3, the mean results of the intergroup comparison showed statistically significant advantages in favour of the CLIL group (T2 $F= 16.85$ $p<.0001$, T3 $F= 13.70$ $p= 0.0003$) (see Table 5.139).

	Group	Mean	StdErr	Lower	Upper	F Value	P Value
T0	CLIL	5.03	0.60	3.98	6.35	4.87	0.0283
	Control	3.49	0.40	2.78	4.38		
T1	CLIL	4.82	0.58	3.80	6.11	0.29	0.5912
	Control	4.42	0.48	3.56	5.48		
T2	CLIL	10.37	1.02	8.53	12.61	16.85	<.0001
	Control	5.81	0.59	4.77	7.09		
T3	CLIL	11.35	1.13	9.32	13.82	13.70	0.0003
	Control	6.78	0.66	5.59	8.22		

Table 5.139 Accuracy achievement results: TNEFU

In terms of progress, the Control group showed a very linear progress although not statistically significant at any of the time periods tested. As for the CLIL group, its progress decreased slightly during the first time period, from T0 to T1; however, the increase from T1 to T2 was statistically significant ($p= <.0001$). From T2 to T3, even though both groups kept progressing, neither of the groups showed statistically significant differences in their improvement. Both groups improved significantly throughout the study and this improvement was statistically significant: Control Group $p= <.0001$, CLIL group $p= <.0001$ (see Table 5.139).

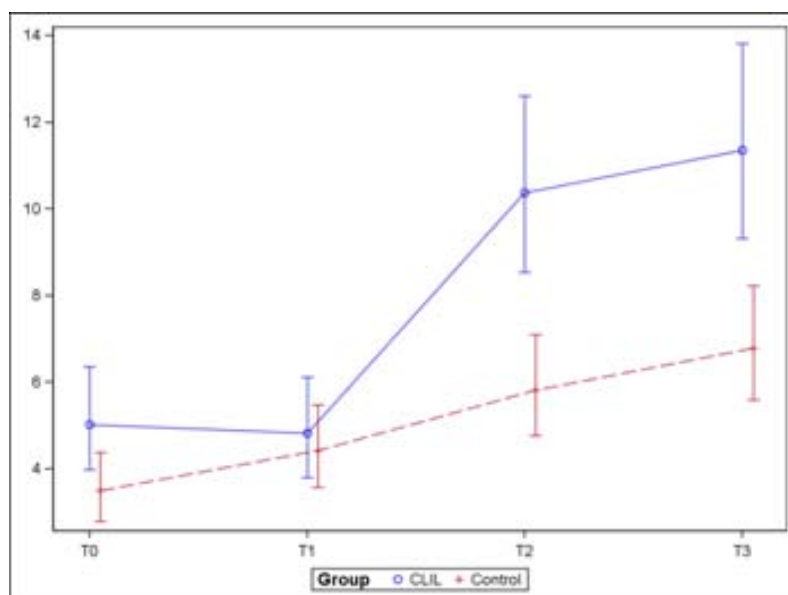


Figure 5.45 Accuracy improvement results: TNEFU

	I0_1	I1_2	I2_3	I0_3
CLIL	-4.1%	115.2%	9.4%	125%
Control	26.6%	31.6%	16.6%	94.2%

Table 5.140 Percentage of improvement in Accuracy: TNEFU

Table 5.141 below shows the achievement results as for TNEFU when the interaction Group/Proficiency level was taken into account. The comparison between High achievers in the CLIL and the Control groups revealed statistically significant differences in favour of the CLIL group at T2 ($F= 8.18$ $p= 0.0049$) and at T3 ($F= 5.98$ $p= 0.0159$). At T2 and at T3, High achievers in the CLIL group wrote a higher number of correct units than their counterparts in the Control Groups.

	Proficiency	Group	Mean	StdErr	Lower	Upper	F Value	P Value
T0	High	CLIL	6.16	0.75	4.84	7.85	1.14	0.2878
	High	Control	4.89	0.87	3.45	6.93		
T1	High	CLIL	6.10	0.75	4.79	7.77	0.96	0.3277
	High	Control	4.95	0.86	3.52	6.97		
T2	High	CLIL	11.37	1.22	9.19	14.06	8.18	0.0049
	High	Control	6.52	1.05	4.74	8.96		

	Proficiency	Group	Mean	StdErr	Lower	Upper	F Value	P Value
T3	High	CLIL	13.24	1.40	10.73	16.32	5.98	0.0159
	High	Control	8.36	1.29	6.16	11.34		

Table 5.141 Accuracy achievement results: TNEFU. Group/High achievers interaction

The results for Low achievers showed that at T0, Low achievers in the CLIL group significantly outperformed Low achievers in the Control Group ($F= 3.92$ $p= 0.0488$). Results at T1 were almost the same for both groups. However, at T2 and at T3, there were significant differences in favour of Low achievers in the CLIL group ($T2$ $p= 0.0041$ / $T3$ $p= 0.0065$). The mean number of correct units written by Low achievers in the CLIL group is much higher (9.73) than those written by Low achievers in the Control Group (5.50).

	Proficiency	Group	Mean	StdErr	Lower	Upper	F Value	P Value
T0	Low	CLIL	4.10	0.86	2.71	6.21	3.92	0.0488
	Low	Control	2.49	0.35	1.89	3.28		
T1	Low	CLIL	3.81	0.82	2.49	5.81	0.02	0.8873
	Low	Control	3.94	0.48	3.10	5.01		
T2	Low	CLIL	9.46	1.63	6.72	13.31	8.53	0.0041
	Low	Control	5.18	0.59	4.15	6.48		
T3	Low	CLIL	9.73	1.70	6.88	13.74	7.64	0.0065
	Low	Control	5.50	0.61	4.42	6.85		

Table 5.142 Accuracy achievement results: TNEFU. Group/Low achievers interaction

In terms of improvement, from T0 to T1, CLIL High and Low achievers as well as High achievers in the Control Group did not improve significantly whereas Low achievers in the Control Group progressed significantly ($p= 0.0108$). Although all the groups seemed to progress during the second time period, improvement was only significant for the CLIL High ($p= <.0001$) and Low achievers ($p= <.0001$), but not for

the students in the Control group. The results of the final time period, T2 to T3, were not statistically significant for any of the groups. Nevertheless, all the groups improved significantly throughout the study, T0-T3: Control High $p= 0.0070$, Control Low $p=<.0001$, CLIL High $p=<.0001$, CLIL Low $p=<.0001$) and as has already been reported, CLIL High and Low achievers are the students with the highest significant achievement results at T3 (see Table 5. 143 and Figure 5.46).

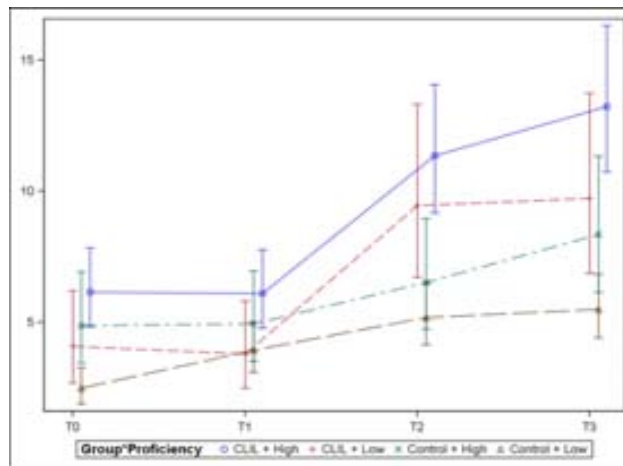


Figure 5.46 Accuracy improvement results: TNEFU. Group/Proficiency level interaction

Proficiency	Group	I0_1	I1_2	I2_3	I0_3
Low	CLIL	-7.1%	148.4%	2.8%	137.2%
	Control	58.2%	31.5%	6.1%	120.6%
High	CLIL	-0.9%	86.3%	16.4%	114.9%
	Control	1.3%	31.7%	28.2%	71.0%

Table 5.143 Percentage of improvement in Accuracy: TNEFU. Group/Proficiency level interaction

5.2.4.2.2 Ratio: Total Number of Error Free Units/Total Number of Units

As can be seen in Table 5.144, the ratio TNEFU/TNU did not show statistically significant differences between the groups at T0 and at T1. However, the achievement results obtained by students in the CLIL group at T2 and T3 were significantly higher than the ones in the Control group (T2 $F= 6.27$ $p= 0.0130$; T3 $F= 4.26$ $p= 0.0402$).

	Group	Mean	StdErr	Lower	Upper	F Value	P Value
T0	CLIL	61.8%	5.1%	52.5%	72.7%	2.43	0.1202
	Control	51.6%	4.2%	44.0%	60.6%		
T1	CLIL	53.9%	4.5%	45.7%	63.6%	2.27	0.1329
	Control	45.5%	3.4%	39.3%	52.7%		
T2	CLIL	66.0%	4.0%	58.4%	74.4%	6.27	0.0130
	Control	52.6%	3.5%	46.1%	60.0%		
T3	CLIL	67.3%	4.2%	59.5%	76.0%	4.26	0.0402
	Control	56.0%	3.5%	49.5%	63.5%		

Table 5.144 Accuracy improvement results: % TNEFU/TNU

The intragroup analysis showed a very similar pattern of progress for both groups. Despite the similarity of the pattern, none of the results obtained by the groups were significant. From T0 to T1, both groups experienced a decrease in the percentage of EFU in relation to the TNU. However, from T1 to T2, the progress of the CLIL group seemed to be slightly higher than that of the Control Group and, from T2 to T3, the CLIL group improved slightly less than the Control Group. Improvement throughout the study (T0 to T3) was very similar for both groups. However, in terms of achievement, as has already been reported, the CLIL group significantly outperformed the Control group at T2 and T3.

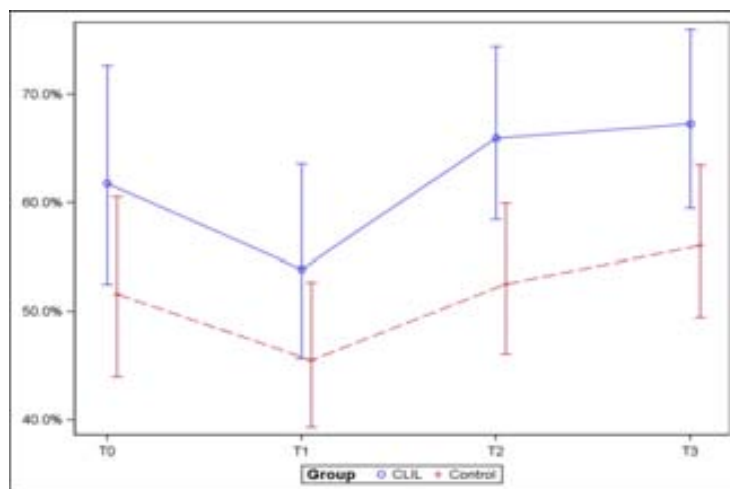


Figure 5.47 Accuracy improvement results: %TNEFU/TNU

	I0_1	I1_2	I2_3	I0_3
CLIL	-12.8%	22.4%	2.0%	8.9%
Control	-11.8%	15.5%	6.6%	8.6%

Table 5.145 Percentage of improvement in Accuracy: %TNEFU/TNU.

5.2.4.3 Complexity

5.2.4.3.1 Lexical Complexity

In order to assess the lexical complexity of CLIL and EFL writings, four measures were taken into account: Total Number of Lexical Verbs (TNLV), Total Number of Adjectives (TNAdj), and the ratio between TNLV in relation to TNWE as well as the ratio between TNAdj in relation to TNWE.

5.2.4.3.1.1 Total Number of Lexical Verbs

Statistically significant differences were found at T1 (F Value 10.78, $p=0.0012$) and T2 (F Value 4.07, $p=0.0456$) in favour of the CLIL group. Although not significantly different, the mean number of lexical verbs used by the CLIL group at T3 was higher (4.201) than the mean number of lexical verbs used by the Control group (3.27).

	Group	Mean	StdErr	Lower	Upper	F Value	P Value
T0	CLIL	0.78	0.17	0.51	1.21	1.52	0.2189
	Control	0.54	0.11	0.35	0.82		
T1	CLIL	2.28	0.31	1.75	2.98	10.78	0.0012
	Control	1.15	0.19	0.84	1.58		
T2	CLIL	4.20	0.48	3.35	5.27	4.07	0.0456
	Control	3.03	0.35	2.41	3.80		
T3	CLIL	4.01	0.47	3.18	5.07	1.59	0.2098
	Control	3.27	0.37	2.62	4.08		

Table 5.146: Lexical Complexity achievement results: TNLV

The results of the intragroup comparison showed that both groups progressed significantly from T0 to T1: Control group $p= 0.0110$, CLIL group $p= <.0001$ and from T1 to T2: Control group $p= <.0001$, CLIL group $p= <.0001$. The progress from T2 to T3 was not significant for any of the groups. The percentage of improvement throughout the study as for Lexical verbs was statistically significant for both groups: Control group $p= <.0001$, CLIL group $p= <.0001$ (see Figure 5.48 and Table 5.147).

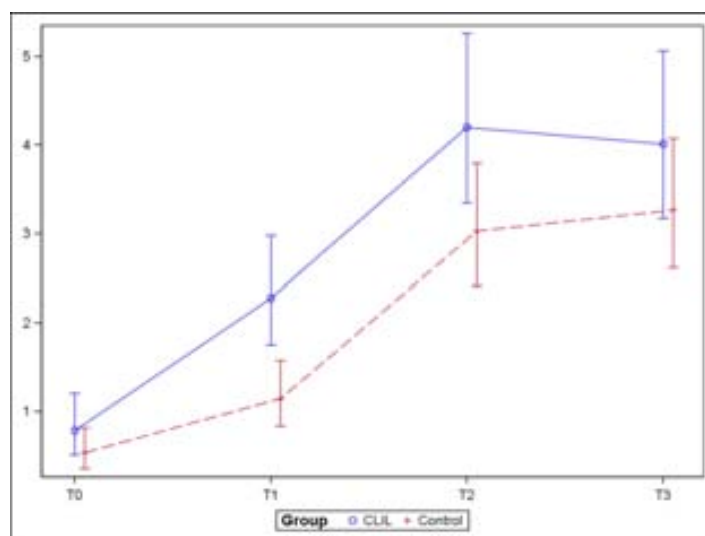


Figure 5.48 Lexical Complexity improvement results: TNLV

	I0_1	I1_2	I2_3	I0_3
CLIL	191.4%	83.9%	-4.5%	411.7%
Control	113.2%	163.7%	7.9%	506.9%

Table 5.147 Percentage of improvement in Lexical Complexity: TNLV

Table 5.148 below shows the results of the intergroup comparison when the interaction group proficiency level was taken into account. The difference in the mean scores obtained by CLIL High achievers was only statistically significant at T1 ($F= 10.03$ $p= 0.0017$) in favour of the CLIL students. At all the other times tested, even though differences between the mean scores were not significantly different, the means were higher for the CLIL High achievers.

	Proficiency	Group	Mean	StdErr	Lower	Upper	F Value	P Value
T0	High	CLIL	1.28	0.23	0.90	1.83	2.91	0.0889
	High	Control	0.70	0.21	0.38	1.28		
T1	High	CLIL	2.34	0.35	1.75	3.14	10.03	0.0017
	High	Control	0.92	0.24	0.55	1.52		
T2	High	CLIL	4.04	0.52	3.13	5.21	0.08	0.7755
	High	Control	3.80	0.66	2.69	5.37		
T3	High	CLIL	4.15	0.54	3.21	5.35	0.78	0.3778
	High	Control	3.41	0.61	2.40	4.85		

Table 5.148 Lexical Complexity achievement results: TNLV. Group/High achievers interaction

The intergroup comparison taking Low achievers into consideration revealed that there were only statistically significant differences at T2 in favour of the CLIL Low achievers ($F= 6.15$ $p= 0.0143$). As in the case of High achievers, even though differences were not significant, the mean scores were higher for Low achievers in the CLIL group (see Table 5.149).

	Proficiency	Group	Mean	StdErr	Lower	Upper	F Value	P Value
T0	Low	CLIL	0.48	0.20	0.21	1.07	0.09	0.7691
	Low	Control	0.42	0.11	0.25	0.69		
T1	Low	CLIL	2.23	0.53	1.40	3.55	2.39	0.1236
	Low	Control	1.44	0.23	1.05	1.97		
T2	Low	CLIL	4.36	0.87	2.94	6.47	6.15	0.0143
	Low	Control	2.41	0.32	1.86	3.14		
T3	Low	CLIL	3.88	0.80	2.58	5.83	0.81	0.3707
	Low	Control	3.13	0.38	2.46	3.99		

Table 5.149 Lexical Complexity achievement results: TNLV. Group/Low achievers interaction

The results of the intragroup comparison show that from T0 to T1 and T1 to T2, the CLIL High achievers ($p= 0.0097$), CLIL Low achievers ($p= 0.0019$) and Control Low achievers ($p= <.0001$) progressed significantly. Control High achievers also

improved significantly during the second time period T1 to T2 ($p = <.0001$). During the third time period, T2 to T3, none of the groups progressed significantly. However, improvement from T0 to T3 was significant for all the groups ($p = <.0001$). As has already been mentioned in the achievement report, the highest scores at T3 were for High and Low achievers in the CLIL groups. Students that obtained the lowest scores in TNLV were the Low achievers in the Control group (see Figure 5.49 and Table 5.150).

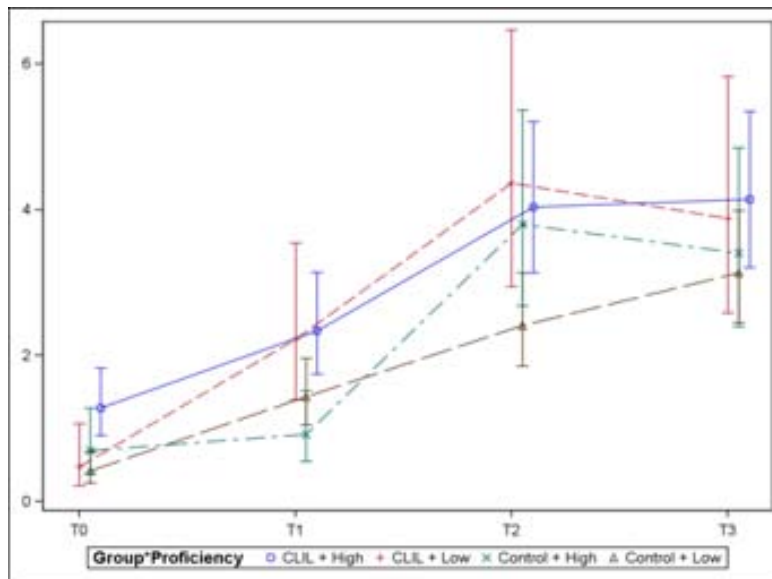


Figure 5.49 Lexical Complexity improvement results: TNLV. Group/Proficiency level interaction

Proficiency	Group	I0_1	I1_2	I2_3	I0_3
Low	CLIL	364.8%	96.0%	-11.0%	710.4%
	Control	245.3%	68.0%	29.7%	652.5%
High	CLIL	82.6%	72.5%	2.6%	223.1%
	Control	31.7%	314.0%	-10.2%	389.5%

Table 5.150 Percentage of improvement: TNLV. Group/Proficiency level interaction

5.2.4.3.1.2 Ratio: Total Number of Lexical Verbs/Total Number of Words in English

The difference in the ratio between TNLV in relation to TNWE was statistically significant at T1 (F Value 13.55, $p= 0.0003$) in favour of the CLIL group. At the rest of the times tested, although the results were no significantly different, the CLIL Group obtained slightly better mean percentages than those of the Control Group (see Table 5.151).

	Group	Mean	StdErr	Lower	Upper	F Value	P Value
T0	CLIL	1.5%	0.3%	1.0%	2.3%	0.89	0.3453
	Control	1.2%	0.2%	0.8%	1.7%		
T1	CLIL	3.6%	0.4%	2.9%	4.5%	13.55	0.0003
	Control	1.8%	0.3%	1.4%	2.5%		
T2	CLIL	5.3%	0.5%	4.4%	6.3%	1.91	0.1680
	Control	4.4%	0.4%	3.7%	5.3%		
T3	CLIL	4.3%	0.4%	3.5%	5.1%	0.07	0.7930
	Control	4.1%	0.4%	3.5%	4.9%		

Table 5.151 Lexical Complexity achievement results: %TNLV/TNWE

Figure 5.50 and Table 5.152 below show the improvement of both groups in this measure, the ratio TNWE/TNW. From T0 to T1, there were no significant differences for the Control Group ($p= 0.2185$), but the CLIL group improved significantly ($p= 0.0009$). From T1 to T2, both groups improved significantly: Control Group $p= <.0001$, CLIL group $p= 0.0212$. However, from T2 to T3, improvement was not significant for any of the groups and, in fact, descriptively, both groups slightly decreased to reach a very similar achievement percentage at the end of the study. The progression T0 to T3 was significant for both groups ($p= <.0001$).

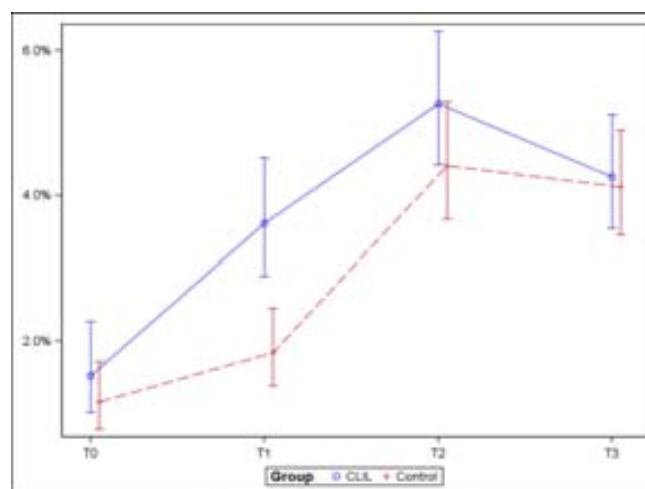


Figure 5.50 Lexical Complexity improvement results: %TNLV/TNWE

	I0_1	I1_2	I2_3	I0_3
CLIL	138.1%	45.5%	-19.1%	180.2%
Control	58.9%	139.2%	-6.7%	245.9%

Table 5.152 Percentage of improvement in Lexical Complexity: % TNLV/TNWE

5.2.4.3.1.3 Total Number of Adjectives

As for the number of adjectives used, Table 5.153 below shows the results obtained by the CLIL and the Control Groups. Mean scores were in favour of the CLIL Group at all the times tested; however, they were significantly different at T0, T2 and T3. The comparison at T1 did not show significant differences.

	Group	Mean	StdErr	Lower	Upper	F Value	P Value
T0	CLIL	2.82	0.33	2.24	3.54	15.71	<.0001
	Control	1.38	0.19	1.05	1.81		
T1	CLIL	1.83	0.25	1.40	2.39	3.30	0.0701
	Control	1.28	0.19	0.96	1.70		
T2	CLIL	3.41	0.37	2.76	4.22	16.27	<.0001
	Control	1.75	0.22	1.37	2.24		
T3	CLIL	4.24	0.46	3.43	5.25	49.31	<.0001
	Control	1.18	0.17	0.88	1.57		

Table 5.153 Lexical Complexity achievement results: TNAdj

The intragroup comparison shows that both groups decreased during the first time period, T0 to T1. The decrease for the CLIL group was statistically significant ($p=0.0206$). However, from T1 to T2, the CLIL group progressed significantly ($p=0.0206$) whereas the improvement of the Control group was not statistically significant. From T2 to T3, the Control group decreased again whereas the CLIL group progressed although not significantly. Nevertheless, as has been explained in the achievement report, the CLIL group attained a significantly higher score as for the number of adjectives than that of the Control group. The progress throughout the study, T0 to T3, was significant for the CLIL group ($p=0.0052$) but it was not significant for the Control Group ($p=0.8247$).

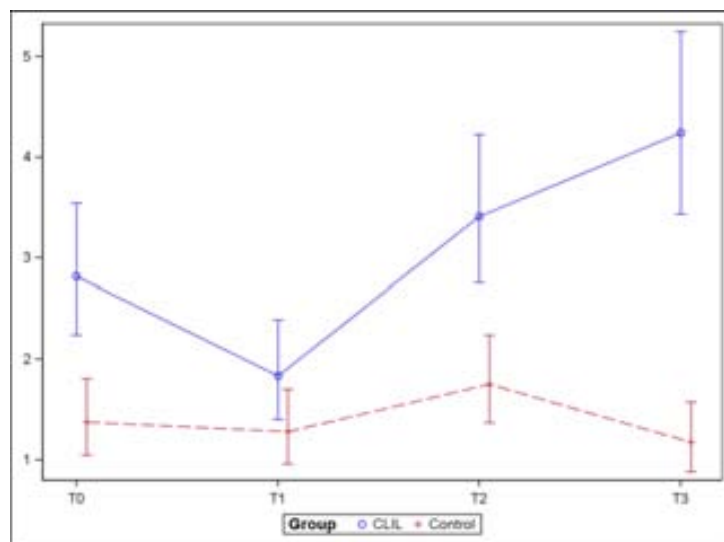


Figure 5.51 Lexical Complexity improvement results: TNAdj

	I0_1	I1_2	I2_3	I0_3
CLIL	-34.9%	86.2%	24.4%	50.8%
Control	-7.1%	37.2%	-32.8%	-14.4%

Table 5.154: Percentage of improvement in Lexical Complexity: TNAdj

The results of the interaction Group/Proficiency level can be seen in Table 5.155 below. Although there were statistically significant differences at T0 in favour of the

High achievers in the CLIL group ($F= 3.90$ $p= 0.0493$), at T1 the differences between the CLIL and the Control groups were not significant. However, at T2 and T3 there were statistically significant differences in favour of the CLIL group (T2 $F= 4.09$ $p= 0.0442$ / T3 $F= 21.43$ $p= <.0001$). The mean score of the CLIL group at the end of the study (5.16) was more than three times the mean score of the Control group (1.65).

	Proficiency	Group	Mean	StdErr	Lower	Upper	F Value	P Value
T0	High	CLIL	2.70	0.36	2.08	3.50	3.90	0.0493
	High	Control	1.63	0.35	1.07	2.50		
T1	High	CLIL	1.86	0.28	1.38	2.50	2.27	0.1329
	High	Control	1.23	0.28	0.78	1.93		
T2	High	CLIL	3.01	0.38	2.34	3.86	4.09	0.0442
	High	Control	1.85	0.38	1.24	2.76		
T3	High	CLIL	5.16	0.57	4.14	6.43	21.43	<.0001
	High	Control	1.65	0.36	1.07	2.53		

Table 5.155 Lexical Complexity achievement results: TNAdj. Group/High achievers interaction

The picture for Low achievers is identical to that of High achievers. There were significant differences in favour of the CLIL Group at T0 ($F= 13.56$ $p= 0.0003$), T2 ($F= 13.61$ $p= 0.0003$) and T3 ($F= 29.92$ $p=<.0001$). As in the case of High achievers, the mean score at T3 as for TNAdj was almost four times the mean score of the Control group.

	Proficiency	Group	Mean	StdErr	Lower	Upper	F Value	P=
T0	Low	CLIL	2.94	0.59	1.98	4.36	13.56	0.0003
	Low	Control	1.16	0.18	0.85	1.58		
T1	Low	CLIL	1.81	0.42	1.14	2.87	1.20	0.2734
	Low	Control	1.33	0.20	0.99	1.79		
T2	Low	CLIL	3.87	0.71	2.69	5.57	13.61	0.0003
	Low	Control	1.66	0.23	1.27	2.18		
T3	Low	CLIL	3.49	0.66	2.40	5.08	29.92	<.0001

Proficiency	Group	Mean	StdErr	Lower	Upper	F Value	P=
Low	Control	0.84	0.15	0.59	1.20		

Table 5.156 Lexical Complexity achievement results: TNAdj. Group/Low achievers interaction

The improvement of the groups when the interaction Group/Proficiency level was taken into account can be seen in Table 5.157 below. From T0 to T1, only Low achievers in the Control Group improved, although the progress was not significant. The mean scores of the rest of the groups decreased. However, from T1 to T2, all four groups improve even though the progress is only significant for CLIL High achievers ($p= 0.0169$) as well as for CLIL Low achievers ($p= 0.0098$). During the final time period, T2 to T3, there were significant differences in favour of CLIL High achievers ($p= 0.0002$) and Control Low achievers ($p= 0.0047$). Improvement from T0 to T3 was only statistically significant for High achievers in the CLIL group ($p= <.0001$).

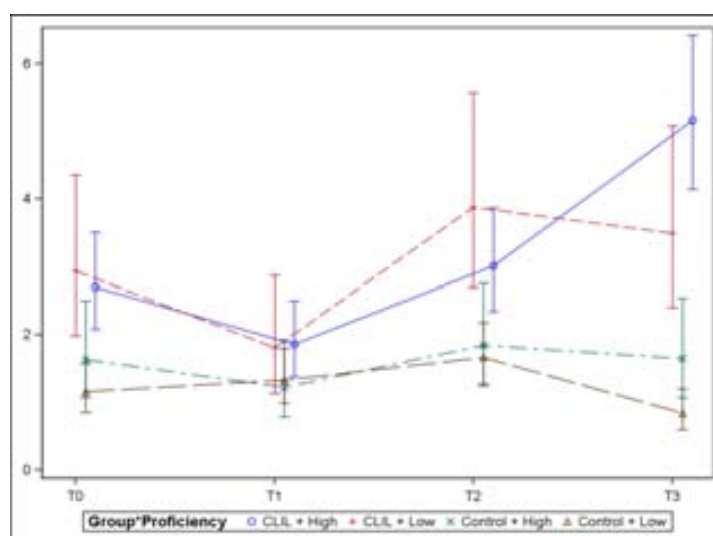


Figure 5.52 Lexical Complexity improvement results:TNAdj. Group/Proficiency level interaction

Proficiency	Group	I0_1	I1_2	I2_3	I0_3
Low	CLIL	-38.5%	114.1%	-9.8%	18.9%
	Control	14.9%	25.0%	-49.4%	-27.3%

Proficiency	Group	I0_1	I1_2	I2_3	I0_3
High	CLIL	-31.1%	61.9%	71.5%	91.3%
	Control	-24.8%	50.5%	-10.9%	0.8%

Table 5.157 Percentage of improvement in Lexical Complexity improvement results: TNAdj. Group/Proficiency level interaction.

5.2.4.3.1.4 Ratio: Total Number of Adjectives/Total Number of Words in English

As can be seen in Table 5.158, at T0, there were statistically significant differences in favour of the CLIL group ($F= 11.20$ $p= 0.0009$). At T1, the mean difference in the ratio TNAdj/TNWE between the two groups was shorter and, although not significantly, the CLIL group outperformed the Control one. However, at T2 the mean difference was statistically significant in favour of the CLIL Group ($F= 8.33$ $p= 0.0043$) which almost doubled the mean result of the Control one (4.1% vs 2.4%). The mean percentage at T3 was three times bigger than the percentage of the Control Group and the difference was significant ($F= 33.22$ $p= <.0001$).

	Group	Mean	StdErr	Lower	Upper	F Value	P Value
T0	CLIL	5.5%	0.7%	4.3%	7.1%	11.20	0.0009
	Control	2.9%	0.4%	2.1%	3.8%		
T1	CLIL	2.9%	0.4%	2.2%	3.9%	3.20	0.0746
	Control	2.0%	0.3%	1.5%	2.7%		
T2	CLIL	4.1%	0.5%	3.3%	5.2%	8.33	0.0043
	Control	2.4%	0.3%	1.9%	3.2%		
T3	CLIL	4.5%	0.5%	3.5%	5.6%	33.22	<.0001
	Control	1.4%	0.2%	1.1%	2.0%		

Table 5.158: Lexical Complexity achievement results; %TNAdj/TNWE

Figure 5.53 below shows the progress made by both, the CLIL and the Control Groups as for the ratio TNAdj /TNWE. From T0 to T1 both groups decreased and the

decrease was significant for the CLIL Group ($p= 0.0003$). From T1 to T2 both groups progressed although the progression was not significant for any of the groups. However, during the final time period analysed, from T2 to T3, the Control Group decreased significantly ($p= 0.0178$) whereas the CLIL group continued to increase. As has already been mentioned, at T3, the achievement percentage in the ratio TNAdj/TNWE of the CLIL group was significantly higher than that of the Control group. From T0 to T3 none of the groups increases significantly.

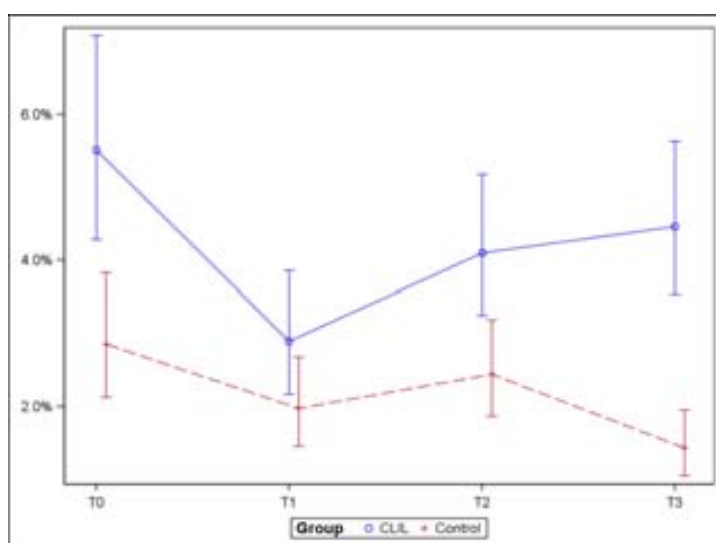


Figure 5.53 Lexical Complexity improvement results: % TNAdj/TNWE

	I0_1	I1_2	I2_3	I0_3
CLIL	-47.4%	41.6%	8.8%	-19.0%
Control	-30.9%	23.5%	-41.3%	-49.9%

Table 5.159 Percentage of improvement in Lexical Complexity: % TNAdj/TNWE

5.2.4.3.2. Syntactic Complexity

For the analysis of syntactic complexity, two measures were taken into account:

Instances of Coordinated Units (ICU) and Instances of Subordinate Units (ISU).

5.2.4.3.2.1 Instances of Coordinated Units

The intergroup analysis showed statistically significant differences in favour of the CLIL group at T0 ($F= 10.64$ $p= 0.0012$). However, the mean results in achievement were not significant at any of the other times tested, even though descriptive statistics showed a slight advantage for the CLIL group at all times (see Table 5.160).

	Group	Mean	StdErr	Lower	Upper	F Value	P Value
T0	CLIL	0.88	0.20	0.56	1.38	10.64	0.0012
	Control	0.11	0.07	0.03	0.36		
T1	CLIL	1.23	0.23	0.85	1.79	0.02	0.8824
	Control	1.19	0.21	0.84	1.68		
T2	CLIL	1.90	0.31	1.38	2.62	0.74	0.3906
	Control	1.56	0.25	1.15	2.14		
T3	CLIL	2.06	0.34	1.49	2.84	0.98	0.3228
	Control	1.65	0.25	1.22	2.24		

Table 5.160 Syntactic Complexity achievement results: ICU

The intragroup analysis, however, showed different patterns of progress for both groups. From T0 to T1, the progress was significant for the Control Group ($p= 0.0007$) but not for the CLIL group. Nevertheless, achievement scores at T1 were almost the same for the CLIL and Control groups. From T1 to T2 and from T2 to T3, although both groups progressed, their progress was not significant. The progress made by the groups throughout the study, T0 to T3, was significant for both of them: Control group $p<.0001$, CLIL group $p=0.0035$.

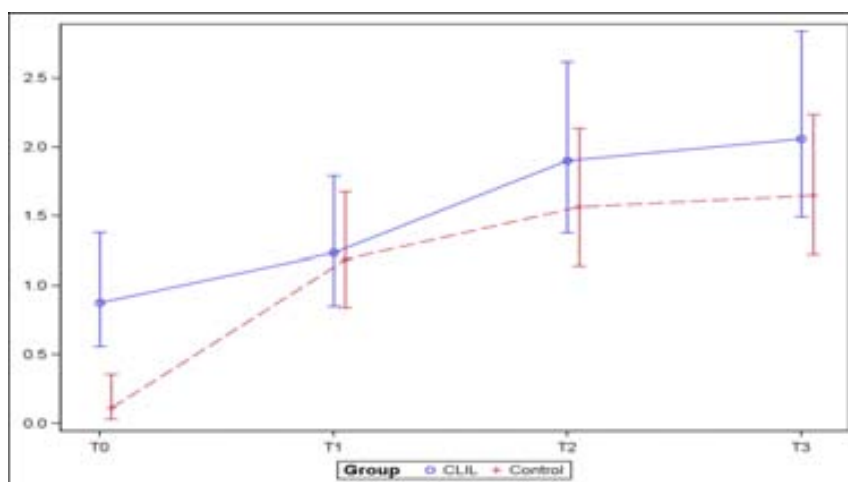


Figure 5.54 Syntactic Complexity improvement results: ICU

	I0_1	I1_2	I2_3	I0-3
CLIL	40.5%	54.0%	8.4%	134.5%
Control	96.5%	31.6%	5.5%	137.7%

Table 5.161 Percentage of improvement in Syntactic Complexity: ICU

Table 5.162 below shows the students' achievement when the variables Group/Proficiency level were considered. At T0, there were statistically significant differences in favour of High achievers in the CLIL group ($F= 9.17$ $p= 0.0027$). However, at any of the other times tested during the study, there were no significant differences for any of the groups. Descriptive statistics showed a slight advantage at T1, T2 and T3 in favour of High achievers in the Control Group.

	Proficiency	Group	Mean	StdErr	Lower	Upper	F Value	P Value
T0	High	CLIL	1.20	0.26	0.79	1.83	9.17	0.0027
	High	Control	0.04	0.05	0.01	0.36		
T1	High	CLIL	1.23	0.26	0.81	1.87	0.01	0.9144
	High	Control	1.28	0.35	0.74	2.21		
T2	High	CLIL	1.88	0.34	1.32	2.69	0.01	0.9268
	High	Control	1.94	0.47	1.20	3.12		
T3	High	CLIL	2.26	0.39	1.61	3.17	0.03	0.8724

Proficiency	Group	Mean	StdErr	Lower	Upper	F Value	P Value
High	Control	2.37	0.55	1.49	3.75		

Table 5.162 Syntactic Complexity achievement results: ICU. Group/High achievers

No significant differences were found when the results of Low achievers in the CLIL and the Control groups were compared. However, it is interesting to notice that CLIL Low achievers slightly advantaged Control Low achievers at all the times tested (see Table 5.163).

	Proficiency	Group	Mean	StdErr	Lower	Upper	F Value	P Value
T0	Low	CLIL	0.64	0.27	0.28	1.46	2.13	0.1455
	Low	Control	0.29	0.10	0.14	0.59		
T1	Low	CLIL	1.24	0.41	0.65	2.37	0.09	0.7655
	Low	Control	1.10	0.21	0.75	1.62		
T2	Low	CLIL	1.92	0.54	1.10	3.35	1.54	0.2154
	Low	Control	1.26	0.23	0.88	1.82		
T3	Low	CLIL	1.88	0.54	1.07	3.30	2.09	0.1494
	Low	Control	1.15	0.22	0.79	1.67		

Table 5.163 Syntactic Complexity achievement results: ICU. Group/Low achievers

As for ICU, the progress was only significant for High and Low achievers in the Control group from T0 to T1: High achievers $p=0.0117$ / Low achievers $p=0.0033$. No other significant differences were found at any of the other time periods for any of the groups. As has already been explained in the achievement results, Control High achievers slightly outperformed CLIL High level students and the group that attained the lowest results is Control Low achievers. The progress throughout the study (T0-T3) was significant for the students in the Control Group: High achievers ($p=0.0016$), Low achievers ($p=0.0020$) as well as for High achievers in the CLIL group ($p=0.0301$). It was not significant for Low achievers in CLIL ($p=0.0755$) even though their scores at T3 were higher than those of Low achievers in the Control Group.

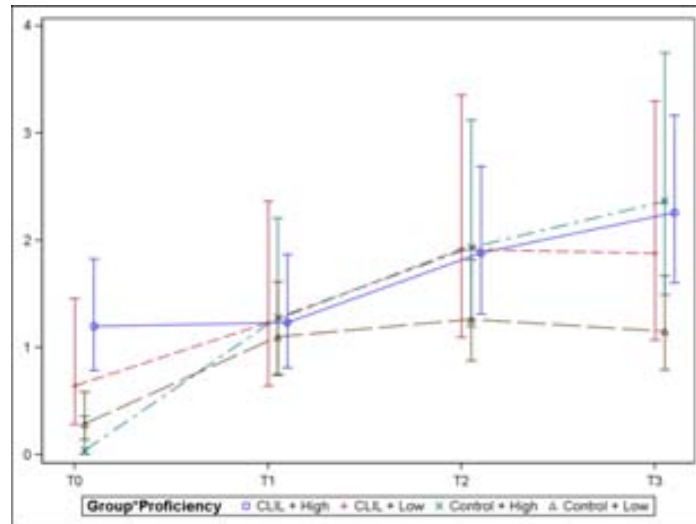


Figure 5.55 Syntactic Complexity improvement results: ICU. Group/Proficiency level interaction

Proficiency	Group	I0_1	I1_2	I2_3	I0_3
Low	CLIL	92.3%	55.3%	-2.0%	192.6%
	Control	280.0%	14.6%	-8.9%	296.7%
High	CLIL	2.7%	52.7%	19.9%	88.0%
	Control	288.6%	51.2%	22.2%	54.19%

Table 5.164 Percentage of improvement results: ICU. Group/Proficiency level interaction.

5.2.4.3.2.2 Instances of Subordinate Units

As for the percentage in the number of examples of subordination, the intergroup analysis showed that at T0 none of the groups used subordination. At T1, T2 and T3, although the writings showed several instances of subordinate units, the groups’ mean scores were not statistically different (see Table 5.165).

	Group	Mean	StdErr	Lower	Upper	F Value	P Value
T1	CLIL	0.33	0.11	0.17	0.63	1.80	0.1814
	Control	0.57	0.14	0.35	0.91		
T2	CLIL	1.50	0.30	1.01	2.22	2.00	0.1599
	Control	1.00	0.20	0.67	1.49		

	Group	Mean	StdErr	Lower	Upper	F Value	P Value
T3	CLIL	1.07	0.25	0.68	1.68	0.19	0.6606
	Control	1.22	0.24	0.83	1.78		

Table 5.165 Syntactic Complexity achievement results: ISU

The intragroup analysis, however, showed different patterns of progress for both groups. From T1 to T2, progress was significant for the CLIL group ($p = <.0001$) but not for the Control Group. The CLIL Group reached the highest mean score at T2. Nevertheless, the progress from T2 to T3 was not significant for any of the groups. The Control group continued to progress but the CLIL group decreased. The progression T1-T3 was significant for both groups: Control Group ($p = 0.0025$) and CLIL group (0.0063) (see Figure 5.56 and Table 5.166).

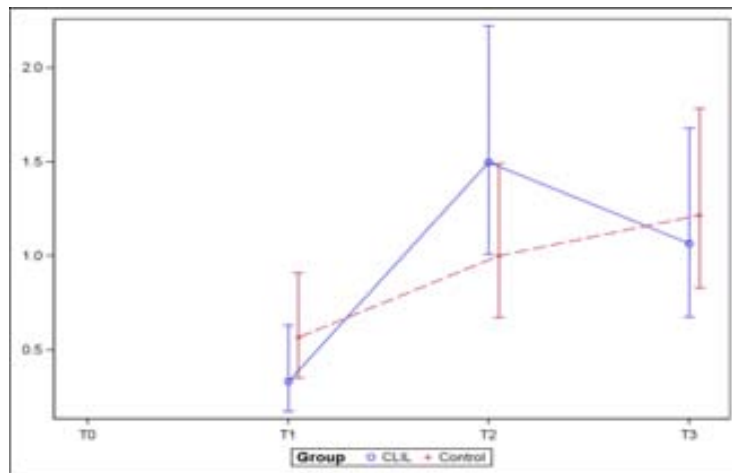


Figure 5.56: Syntactic Complexity improvement results: ISU

	I1_2	I2_3	I1_3
CLIL	351.2%	-28.8%	221.2%
Control	77.2%	21.3%	115.0%

Table 5.166: Percentage of improvement in syntactic complexity: ISU

As described in Table 5.167 below, the intergroup comparison taking into account the proficiency level of the students showed no significant differences between High achievers in the CLIL and Control groups at any of the times tested.

	Proficiency	Group	Mean	StdErr	Lower	Upper	F Value	P Value
T1	High	CLIL	0.41	0.13	0.22	0.77	2.89	0.0906
	High	Control	0.93	0.34	0.46	1.89		
T2	High	CLIL	1.55	0.33	1.01	2.38	0.21	0.6455
	High	Control	1.31	0.40	0.71	2.40		
T3	High	CLIL	1.47	0.32	0.95	2.27	0.06	0.8076
	High	Control	1.61	0.47	0.90	2.88		

Table 5.167 Syntactic Complexity achievement results: ISU. Group/High achievers interaction

As in the case of High achievers, the comparison of Low achievers did not reveal significant differences either. The mean number of subordinate units was generally very low (see Table 5.168).

	Proficiency	Group	Mean	StdErr	Lower	Upper	F Value	P Value
T1	Low	CLIL	0.27	0.16	0.09	0.84	0.15	0.6956
	Low	Control	0.34	0.10	0.19	0.62		
T2	Low	CLIL	1.44	0.50	0.73	2.87	2.26	0.1347
	Low	Control	0.77	0.18	0.48	1.22		
T3	Low	CLIL	0.77	0.32	0.34	1.74	0.14	0.7083
	Low	Control	0.92	0.21	0.59	1.43		

Table 5.168 Syntactic Complexity achievement results: ISU. Group/Low achievers interaction

The intragroup comparison when the proficiency level of the students was taken into account showed no statistically significant differences during the second and the third time periods analysed. The group with the highest score at T3 was Control High achievers and the group with the lowest score at T3 was CLIL Low achievers. The progress T1-T3 was significant for the CLIL High group ($p= 0.0002$) but not for the

CLIL Low group ($p= 0.1996$). Conversely, it was not significant for the Control High group ($p= 0.2712$) but it was for the Control Low group ($p=0.0046$) (see Table 5.169).

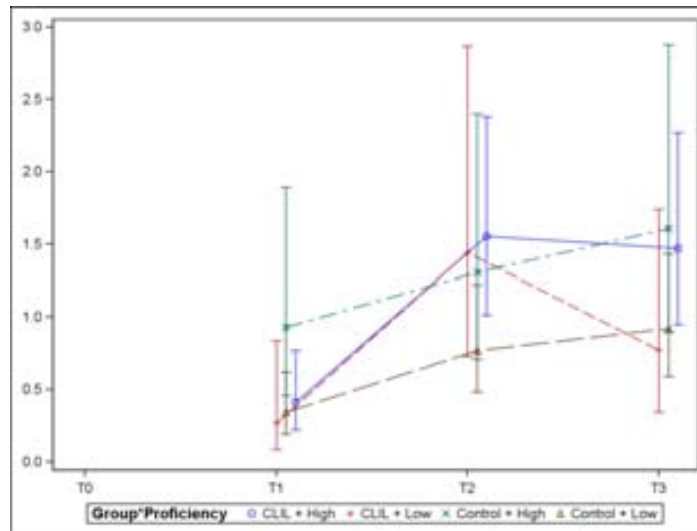


Figure 5.57 Syntactic Complexity improvement results: ISU. Group/Proficiency level interaction.

Proficiency	Group	I1_2	I2_3	I1-3
Low	CLIL	439.0%	-46.4%	188.8%
	Control	123.8%	19.6%	176.7%
High	CLIL	277.7%	-5.4%	257.1%
	Control	40.3%	23.1%	72.7%

Table 5.169 Percentage of improvement in Syntactic Complexity: ISU. Group/Proficiency level interaction

5.2.5 Summary of CLIL Arts & Crafts Writing Results

This section will present a summary of the writing results attained by students exposed to Arts & Crafts in English compared to those obtained by their Control group.

5.2.5.1 Summary of Writing Achievement Results

Table 5.170 below shows a summary of the achievement results in the different measures taken to analyse Fluency.

Language Competence of Young Learners Exposed to EFL and CLIL

FLUENCY	TNW		TNWE		%TNWE/TNW		TNU	
	Control	CLIL	Control	CLIL	Control	CLIL	Control	CLIL
T0	63.47	65.26	49.77	55.31	77.4%	83.8%	7.26	8.75
	p= 0.8047		p= 0.4101		p= 0.0425		p= 0.2211	
T1	82.37	76.23	67.85	67.16	81.2%	86.9%	10.46	9.53
	p= 0.3981		p= 0.9186		p= 0.0598		p= 0.4464	
T2	86.42	98.85	73.72	84.73	83.9%	85.1%	11.85	16.39
	p= 0.0883		p= 0.1027		p= 0.6755		p= 0.0002	
T3	99.31	110.8	86.03	100	85.2%	89.8%	13.22	18.07
	p= 0.1199		p= 0.0409		p= 0.1208		p= 0.0001	

Table 5.170 Summary of Fluency achievement results

The results of the intergroup comparison when the interaction Group/Proficiency level was taken into account are displayed in Table 5.171 below.

FLUENCY Group/proficiency level interaction	TNWE				TNU			
	HIGH		LOW		HIGH		LOW	
	Control	CLIL	Control	CLIL	Control	CLIL	Control	CLIL
T0	56.68	58.74	42.86	51.87	8.39	9.86	6.14	7.64
	p= 0.8250		p= 0.3649		p= 0.3826		p= 0.4028	
T1	70.00	69.07	65.70	65.25	10.75	11.02	10.17	8.05
	p= 0.9209		p= 0.9642		p= 0.8718		p= 0.2383	
T2	78.87	93.45	68.58	76.02	12.91	16.77	10.80	16.00
	p= 0.1196		p= 0.4546		p= 0.0229		p= 0.0042	
T3	95.42	107.4	76.64	92.66	16.43	19.72	11.63	14.81
	p= 0.2021		p= 0.1145		p= 0.0042		p= 0.0093	

Table 5.171 Summary of Fluency achievement results: TNWE, TNU. Group/Proficiency level interaction.

Table 5.172 summarises the accuracy achievement results of the CLIL and the Control Groups.

ACCURACY	TNEFU		%TNEFU/TNU	
	Control	CLIL	Control	CLIL
T0	3.49	5.03	51.6%	61.8%
	p= 0.0283		p= 0.1202	
T1	4.42	4.82	45.5%	53.9%

Language Competence of Young Learners Exposed to EFL and CLIL

	p= 0.0512		p= 0.1329	
T2	5.81	<.0001	52.6%	66.0%
	p= 0.5556		p= 0.0130	
T3	6.78	11.35	56.0%	67.3%
	p= 0.0003		p= 0.0402	

Table 5.172 Summary of Accuracy achievement results.

Table 5.173 below describes the accuracy achievement results as for the TNEFU when the interaction Group/ proficiency level was taken into account.

ACCURACY Group/proficiency level interaction	TNEFU			
	HIGH		LOW	
	Control	CLIL	Control	CLIL
T0	4.89	6.16	2.49	4.10
	P= 0.2878		P= 0.0488	
T1	4.95	6.10	3.94	3.81
	P= 0.3277		P= 0.8873	
T2	6.52	11.37	5.18	9.46
	P= 0.0049		P= 0.0041	
T3	8.36	13.24	5.50	9.73
	P= 0.0159		P= 0.0065	

Table 5.173 Summary of Accuracy achievement results: TNEFU. Group/Proficiency level interaction

A summary of the achievement results as for Lexical Complexity is displayed in table 5.174 below.

LEXICAL COMPLEXITY	TNLV		%TNLV/TNWE		TNAdj		TNAdj/TNWE	
	Control	CLIL	Control	CLIL	Control	CLIL	Control	CLIL
T0	0.54	0.78	1.2%	1.5%	1.38	2.82	2.9%	5.5%
	p= 0.2189		p= 0.3453		p= <.0001		p= 0.0009	
T1	1.15	2.28	1.8%	3.6%	1.28	1.83	2.0%	2.9%
	p= 0.0012		p= 0.0003		p= 0.0701		p= 0.0746	
T2	3.03	4.20	4.4%	5.3%	1.75	3.41	2.4%	4.1%
	p= 0.0456		p= 0.1680		p= <.0001		p= 0.0043	

T3	3.27	4.01	4.1%	4.3%	1.18	4.24	1.4%	4.5%
	p= 0.2098		p= 0.7930		p= < .0001		p= < .0001	

Table 5.174 Summary of Lexical Complexity achievement results

Table 5.175 below shows the results for two of the measures used to analyse lexical complexity (TNLV and TNAdj) when the interaction Group/Proficiency level was taken into consideration.

LEXICAL COMPLEXITY Group/proficiency level interaction	TNLV				TNAdj			
	HIGH		LOW		HIGH		LOW	
	Control	CLIL	Control	CLIL	Control	CLIL	Control	CLIL
T0	0.70	1.28	0.42	0.48	1.63	2.70	1.16	2.94
	p= 0.0889		p= 0.7691		p= 0.0493		p= 0.0003	
T1	0.92	2.34	1.44	2.23	1.23	1.86	1.33	1.81
	pP= 0.0017		p= 0.1236		p= 0.1329		p= 0.2734	
T2	3.80	4.04	2.41	4.36	1.85	3.01	1.66	3.87
	p= 0.7755		p= 0.0143		p= 0.0442		p= 0.0003	
T3	3.41	4.15	3.13	3.88	1.65	5.16	0.84	3.49
	p= 0.3778		p= 0.3703		p= <.0001		p= <.0001	

Table 5.175 Summary of Lexical Complexity achievement results: TNLV, TNAdj

Group/Proficiency level interaction

Table 5.176 summarises the results of two measures taken in order to measure the syntactic complexity of the writings: Instances of Coordinated Units and Instances of Subordinate Units.

SYNTACTIC COMPLEXITY	ICU		ISU	
	Control	CLIL	Control	CLIL
T0	0.11	0.88		
	p= 0.0012			
T1	1.19	1.23	0.57	0.33
	p= 0.8824		p= 0.1814	
T2	1.56	1.90	1.00	1.50
	p= 0.39065		p= 0.1599	

T3	1.65	2.06	1.22	1.07
	p= 0.3228		p= 0.6606	

Table 5.176 Summary of Syntactic Complexity achievement results: ICU, ISU

Table 5.177 below shows a summary of the results attained as for the ICU and ISU when the interaction Group/Proficiency level was taken into account.

SYNTACTIC COMPLEXITY Group/proficiency level interaction	ICU				ISU			
	HIGH		LOW		HIGH		LOW	
	Control	CLIL	Control	CLIL	Control	CLIL	Control	CLIL
T0	0.04	1.20	0.29	0.64				
	p = 0.0027		p = 0.1455					
T1	1.28	1.23	1.10	1.24	0.93	0.41	0.34	0.27
	p = 9144		p = 0.7655		p = 0.0906		p = 0.6965	
T2	1.94	1.88	1.26	1.92	1.31	1.55	0.77	1.44
	p = 0.9268		p = 0.2154		p = 0.6455		p = 0.1347	
T3	2.37	2.26	1.15	1.88	1.61	1.47	0.92	0.77
	p = 8724		p = 0.1494		p = 0.8076		p = 0.7083	

Table 5.177 Summary of Syntactic Complexity achievement results: ICU, ISU Group/Proficiency level interaction.

5.2.5.2 Summary of Writing Improvement Results

Table 5.178 below displays a summary of the fluency improvement results

FLUENCY	TNW		TNWE		%TNWE/TNW		TNU	
	Control	CLIL	Control	CLIL	Control	CLIL	Control	CLIL
T0-T1	16.8% p=0.0009	29.8% p=0.1746	21.4% p=0.0002	36.3% p=0.0593	4.8% p=0.3508	3.6% p=0.6545	44.0% p=0.0004	8.9% p=0.7966
T1-T2	29.7% p=0.8439	4.9% p=0.0002	26.2% p=0.5256	8.7% p=0.0013	3.3% p=0.5945	-2.1% p=0.8730	13.3% p=0.2879	71.9% p=<.0001
T2-T3	12.1% p=0.0459	14.9% p=0.1341	18.1% p=0.0244	16.7% p=0.0094	5.5% p=0.9106	1.6% p=0.1603	11.5% p=0.3037	10.3% p=0.2229
T0-T3	69.8% p=<.0001	56.5% p=<.0001	72.9% p=<.0001	80.9% p=<.0001	7.1% p=0.0028	10.1% p=0.0926	82.0% p=<.0001	106.5% p=<.0001

Table 5.178 Summary of Fluency improvement results

The description of the improvement results for the TNWE, when the variable group interacted with proficiency level is shown in the table below.

FLUENCY Group/Proficiencylevel Interaction	TNWE				TNU			
	HIGH		LOW		HIGH		LOW	
	Control	CLIL	Control	CLIL	Control	CLIL	Control	CLIL
T0-T1	23.5% p=0.2411	17.6% p=0.2046	53.3% p<.0001	25.8% p=0.3642	28.1% p=0.2631	11.7% p=0.6206	28.1% p<.0001	11.7% p=0.9929
T1-T2	12.7% p=0.9283	35.3% p<.0001	4.4% p=0.9283	16.5% p=0.5568	20.1% p=0.3395	52.2% p<.0001	20.1% p=0.8847	52.2% p<.0001
T2-T3	21.0% p=0.3188	14.9% p=0.0465	11.8% p=0.3188	21.9% p=0.1992	14.7% p=0.4565	17.5% p=0.0143	14.7% p=0.7617	15.5% p=0.9922
T0-T3	68.3% p<.0001	82.8% p<.0001	78.6% p<.0001	78.8% p<.0001	76.6% p<.0001	99.9% p<.0001	89.6% p<.0001	115.1% p<.0001

Table 5.179 Summary of Fluency improvement results TNWE, TNU. Group/Proficiency Level interaction

Table 5.180 summarises the results of the intragroup comparison as for the measures chosen to test accuracy: TNEFU and the ratio TNEFU/TNU.

ACCURACY	TNEFU		%TNEFU/TNU	
	Control	CLIL	Control	CLIL
T0-T1	26.6% p= 0.2098	-4.1% p= 0.9880	-12.8% p= 0.5774	-11.8% p= 0.5676
T1-T2	31.6% p= 0.0534	115.2% p<.0001	15.5% p= 0.3525	22.4% p= 0.1166
T2-T3	16.6% p= 0.3754	9.4% p= 0.7235	6.6% p= 0.8459	2.0% p= 0.9926
T0-T3	92.4% p<.0001	125.8% p<.0001	8,3% p=0.8014	8.9% p= 0.7770

Table 5.180 Summary of Accuracy improvement results

The findings from the intragroup comparison when the variables Group/Proficiency level were considered can be seen in Table 5.181 below.

Language Competence of Young Learners Exposed to EFL and CLIL

ACCURACY Group/Proficiency level Interaction	TNEFU			
	HIGH		LOW	
	Control	CLIL	Control	CLIL
T0-T1	1.3% p=0.9999	-0.9% p=0.9998	21.0% p=0.0108	20.2% p=0.9885
T1-T2	86.3% p=0.3534	31.7% p=<.0001	61.1% p=0.1162	46.9% p=<.0001
T2-T3	28.2% p=0.3329	16.4% p=0.2992	20.6% p=0.9526	72.8% p=0.9978
T0- T3	71.0% p=<.0001	114.9% p=<.0001	120.6% p=<.0001	137.2% p=<.0001

Table 5.181 Summary of Accuracy improvement results: TNWE. Group/Proficiency level interaction

Table 5.182 below summarises the results of the measures used to analyse lexical complexity.

	LEXICAL COMPLEXITY							
	TNLV		% TNLV/TNWE		TNAdj		% TNAdj/TNWE	
	CLIL	CONTROL	CLIL	CONTROL	CLIL	CONTROL	CLIL	CONTROL
T0-T1	191.4% p= <.0001	113.2% p= 0.0110	138.1% p= 0.0009	58.9% p= 0.2185	-34.9% p= 0.0206	-7.1% p= 0.9766	-47.4% p= 0.0003	-30.9% p=0.2086
T1-T2	83.9% p= <.0001	163.7% p= <.0001	45.5% p= 0.0212	139.2% p= <.0001	86.2% p=<.0001	37.2% p= 0.2388	41.6% p= 0.0971	23.5% p= 0.0653
T2-T3	-4.5% p=0.9773	7.9% p=0.9105	-19.1% p=0.2276	-6.7% p=0.9297	24.4% p=0.2301	-32.4% p=0.0913	8.8% p=0.8994	-41.3% p=0.0178
T0-T3	506.9% p= <.0001	411.7% p= <.0001	180.2% p= <.0001	254.9% p= <.0001	50.8% p=0.0052	-14.4% p= 0.8247	-19.0% p= 0.3815	-49.9% p=0.0021

Table 5.182: Summary of Lexical Complexity improvement results

The intragroup comparison when the variables Group/Proficiency level were considered are shown in Table 5.183.

LEXICAL COMPREXITY Group/Proficiency level Interaction	TNLV				TNAdj			
	HIGH		LOW		HIGH		LOW	
	Control	CLIL	Control	CLIL	Control	CLIL	Control	CLIL
T0-T1	31.7% p= 0.8675	82.6% p= 0.0097	245.3% p=<.0001	364.8% p= 0.0019	-24.8% p= 0.7305	-31.1% p= 0.1160	14.9% P= 0.8899	-38.5% p= 0.2230
T1-T2	314% p=<.0001	72.5% p=0.0011	68.0% p=0.0115	96.0% p=0.0177	50.5% p=0.4223	61.9% p=0.0169	25.0% p=0.5916	114.1% p= 0.0098
T2-T3	-10.2% p=0.9207	2.6% p=0.9969	29.7% p=0.2076	-11.0% p=0.9318	-10.9% p=0.9701	71.5% p=0.0002	-49.4% p=0.0047	-9.8% p= 0.9540
T0-T3	389.5% p=<.0001	223.1% p=<.0001	652.5% p=<.0001	710.4% p=<.0001	0.8% p=1.0000	91.3% p=<.0001	-27.3% p=0.4494	18.9% p=8479

Table 5.183 Summary of Lexical Complexity improvement results: TNLV, TNAdj. Group/Proficiency level interaction

Table 5.184 below shows the results obtained by the CLIL Arts & Crafts group and the Control Group as for Syntactic Complexity.

SYNTACTIC COMPLEXITY				
	ICU		ISU	
	CLIL	CONTROL	CLIL	CONTROL
T0-T1	40.5% p= 0.5724	965.4% p= 0.0007		
T1-T2	54% p= 0.1664	31.6% p= 0.5120	351.2% p=<.0001	77.2% p= 0.5120
T2-T3	8.4% p= 0.9722	5.5% p=0.9908	-28.8% p= 0.3041	21.3% p= 0.6209
T1-T3	134.5% p= 0.0035	137.9% p=<.0001	221.2% p= 0.0025	115.0% p=0.0063

Table 5.184: Summary of Syntactic Complexity improvement results

Table 5.185 below shows a summary of the improvement results in Syntactic Complexity when the variables Group/Proficiency level were taken into consideration for the intragroup comparison

SYNTACTIC COMPLEXITY								
	ICU				ISU			
	HIGH		LOW		HIGH		LOW	
	CLIL	CONTROL	CLIL	CONTROL	CLIL	CONTROL	CLIL	CONTROL
T0-T1	2.7% p=0.9996	28.8% p= 0.0117	92.3% p=0.5158	280.0% p=0.0033				
T1-T2	52.7% p=0.2678	51.2% p= 0.5308	55.3% p=0.6211	14.6% p=0.9322	277.7% p=0.0001	40.3% p= 0.6247	439.0% p=0.0114	123.8% p=0.0317
T2-T3	19.9% p=0	22.2% p=0.8766	-2.0% p=0.9999	-8.9% p=0.9751	-5.4% p=0.9622	23.1% p=0.7700	-46.4% p=0.2854	19.6% p=0.7468
T0-T3	88.0% p=0.0301	54.1% p=0.0016	192.6% p=0.0755	296.7% p=0.0020	257.1% p=0.0002	72.7% p=0.2712	188.8% p=0.1996	167.7% p=0.0046

Table 5.185: Summary of Syntactic Complexity improvement results: ICU, ISU. Group/Proficiency level interaction

5.3 CLIL Science results vs CLIL Arts & Crafts results

5.3.1 Listening test results

In sections 5.3.1 and 5.3.2, a description of the results obtained in listening and reading by CLIL Science students compared to those obtained by CLIL Arts & Crafts students will be presented in order to answer the following research questions and subquestions: RQ 3.1: Are there differences in achievement²⁹ in listening and reading between CLIL Science students and CLIL Arts & Crafts students at different times (T0, T1, T2, T3)? RQ 3.3: How does the initial level of English proficiency affect the students' performance in listening and reading in the CLIL Science and the CLIL Arts & Crafts group?

The intergroup comparisons between CLIL Science and the CLIL Arts & Crafts students did not yield the same results for all the times. The comparison of the scores at T0 showed that, despite the slightly higher percentage obtained by CLIL Science

²⁹ This section will only report the achievement results of the intergroup comparison CLIL Science vs. CLIL Arts & Crafts. The progress made by each individual group has already been explained in the previous sections.

students (38.4%), the differences between the CLIL Science and CLIL Arts & Crafts groups were not significant ($F=0.02$ $p=0.8950$); nor were the differences significant between the same groups at T1 ($F= 1.01$ $p= 0.3147$) and T2 ($F= 0.001$, $p= 0.9365$). The mean results at T2 were almost the same for both groups. The results at T3, however, showed a statistically significant advantage in favour of the CLIL Science group ($F=9.94$, $p=0.0018$) (see Table 5.186).

	Type	Mean	StdErr	Lower	Upper	F Value	P Value
T0	Arts & Crafts	38.4%	3.0%	32.6%	44.5%	0.02	0.8950
	Science	37.8%	3.5%	31.1%	44.9%		
T1	Arts & Crafts	45.9%	3.2%	39.7%	52.3%	1.01	0.3147
	Science	51.0%	3.8%	43.6%	58.4%		
T2	Arts & Crafts	54.4%	3.2%	48.0%	60.6%	0.001	0.9365
	Science	54.8%	3.5%	47.9%	61.5%		
T3	Arts & Crafts	57.5%	3.2%	51.2%	63.6%	9.94	0.0018
	Science	71.8%	3.0%	65.5%	77.3%		

Table 5.186 Listening achievement results

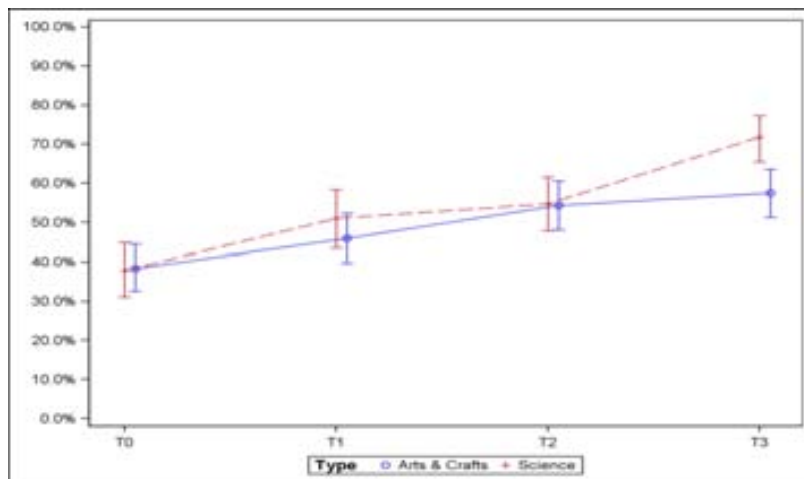


Figure 5.58 Listening achievement results

The intergroup comparison, when the variables Group/High achievers were taken into account, shows that there were no statistically significant differences between the two groups, CLIL Science and CLIL Arts & Crafts, at any of the times tested. Descriptive

statistics, however, indicate that, even though High achievers in Arts & Crafts attained better mean scores at T0 and at T2, at the end of the study, at T3, High achievers in Science obtained a better mean percentage than High achievers in the CLIL Arts& Crafts group (see Table 5.187 below).

	Proficiency	Type	Mean	StdErr	Lower	Upper	F Value	P Value
T0	High	Arts & Crafts	54.9%	5.1%	44.8%	64.5%	0.08	0.7760
	High	Science	53.0%	3.8%	45.5%	60.4%		
T1	High	Arts & Crafts	50.8%	5.4%	40.4%	61.2%	0.71	0.4006
	High	Science	56.5%	4.0%	48.6%	64.1%		
T2	High	Arts & Crafts	67.3%	4.7%	57.4%	75.8%	1.47	0.2267
	High	Science	59.8%	3.7%	52.3%	66.8%		
T3	High	Arts & Crafts	68.9%	4.6%	59.1%	77.2%	1.90	0.1686
	High	Science	76.5%	3.1%	70.0%	82.0%		

Table 5.187 Listening achievement results: Group/High achievers interaction

Table 5.188 below shows that the difference in achievement scores between CLIL Science Low achievers compared to CLIL Arts & Crafts Low achievers was only statistically significant in favour of the CLIL Science Low achievers at T3 (F=11.30 p 0.0009). At this time, the mean difference between the two groups was 21.2%.

	Proficiency	Type	Mean	StdErr	Lower	Upper	F Value	P Value
T0	Low	Arts & Crafts	24.2%	2.9%	19.0%	30.3%	0.01	0.9400
	Low	Science	24.6%	4.2%	17.3%	33.7%		
T1	Low	Arts & Crafts	41.1%	3.6%	34.3%	48.3%	0.46	0.5002
	Low	Science	45.5%	5.4%	35.2%	56.2%		
T2	Low	Arts & Crafts	40.9%	3.6%	34.1%	48.0%	1.96	0.1628
	Low	Science	49.7%	5.1%	39.7%	59.6%		
T3	Low	Arts & Crafts	45.3%	3.5%	38.5%	52.3%	11.30	0.0009
	Low	Science	66.5%	4.8%	56.4%	75.3%		

Table 5.188 Listening achievement results: Group/Low achievers interaction

As can be seen in Figure 5.59, at the end of the study, at T3, the highest mean percentage score was attained by High achievers in Science (76.5%) and the second mean percentage score was for High achievers in Arts & Crafts (68.9%). However, it is interesting to notice that Low achievers in Science attained a very similar mean percentage (66.5%) to the percentage of High achievers in Arts & Crafts. The lowest score was for Low achievers in Arts & Crafts.

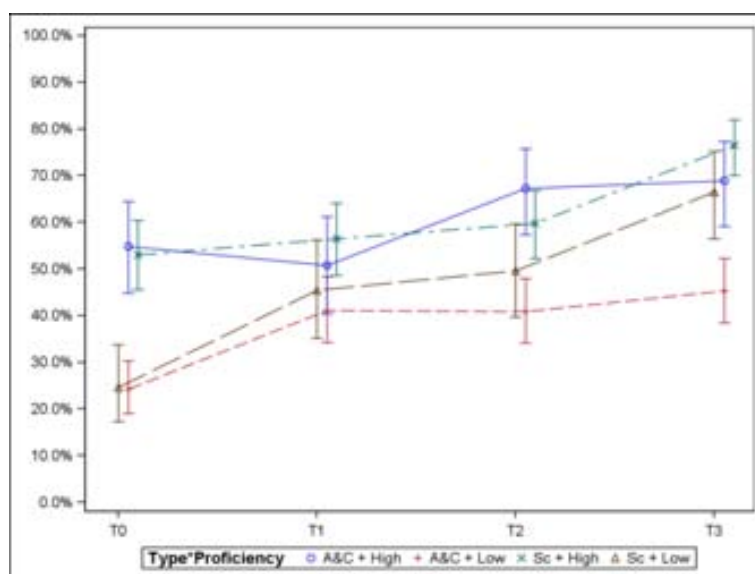


Figure 5.59 Listening achievement results: Group/Proficiency level interaction

5.3.2 Reading test results

In this section, the results for the reading tests will be presented. In the first place the total results for reading will be reported and then the results for each of the individual reading questions will also be presented. As in previous sections, reading question 4 has been analysed for accuracy and comprehension. The analysis of the interaction Group/Proficiency level will only be carried out for the general results, but not for the results of each individual question.

The intergroup comparisons carried out using the results of the reading tests administered to the CLIL Science and to the CLIL Arts & Crafts students did not show significant differences in achievement at any of the times tested. The comparison of the

scores obtained at T0 showed that the differences between the CLIL Science and the CLIL Arts & Crafts were not significant ($F = 0.39$ $p= 0.5318$); nor were the differences between the same groups at T1 ($F= 0.37$ $p= 0.5457$), T2 ($F=1.45$ $p= 0.2305$) and T3 ($F= 1.85$ $p=0.1761$). Descriptive statistics showed that, the Arts & Crafts group scores were slightly higher than those of the CLIL Science students at all the times tested. The difference in the mean scores at T3 is 5.5% (see Table 5.189 and Figure 5.60).

	Type	Mean	StdErr	Lower	Upper	F Value	P Value
T0	Arts & Crafts	40.5%	2.7%	35.2%	46.0%	0.39	0.5318
	Science	37.8%	3.1%	32.0%	44.0%		
T1	Arts & Crafts	50.8%	2.9%	45.2%	56.5%	0.37	0.5457
	Science	48.1%	3.3%	41.7%	54.7%		
T2	Arts & Crafts	55.5%	2.8%	49.9%	60.9%	1.45	0.2305
	Science	50.3%	3.2%	44.0%	56.5%		
T3	Arts & Crafts	66.9%	2.6%	61.6%	71.7%	1.85	0.1761
	Science	61.4%	3.0%	55.2%	67.2%		

Table 5.189 Reading achievement results

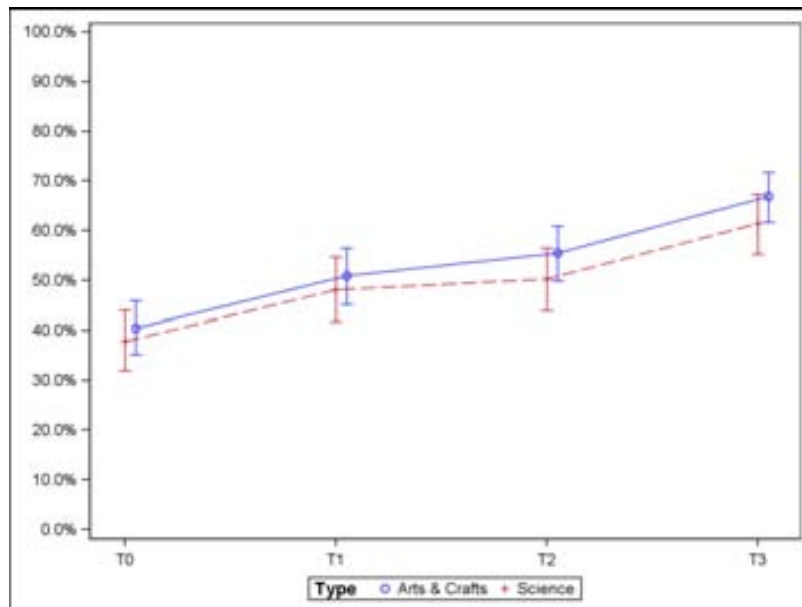


Figure 5.60 Reading achievement results

The interaction Group/High achievers showed statistically significant differences in Reading in favour of High achievers in Arts & Crafts at T2 ($F= 6.45$ $p= 0.0124$). Descriptively, High achievers in Arts & Crafts attained slightly better mean scores at all the times tested (see Table 5.190).

	Proficiency	Type	Mean	StdErr	Lower	Upper	F Value	P Value
T0	High	Arts & Crafts	54.7%	4.6%	45.5%	63.5%	0.16	0.6943
	High	Science	52.4%	3.5%	45.5%	59.1%		
T1	High	Arts & Crafts	61.7%	4.5%	52.5%	70.1%	0.03	0.8551
	High	Science	60.6%	3.4%	53.8%	67.1%		
T2	High	Arts & Crafts	66.4%	4.2%	57.7%	74.2%	6.45	0.0124
	High	Science	52.1%	3.5%	45.2%	58.9%		
T3	High	Arts & Crafts	77.8%	3.3%	70.5%	83.7%	3.54	0.0622
	High	Science	69.0%	3.0%	62.7%	74.6%		

Table 5.190 Reading achievement results: Group/High achievers interaction

As can be seen in Table 5.191 and Figure 5.61 below, no statistically significant differences in Reading were found in favour of Low achievers in any of the groups at any of the times tested.

	Proficiency	Type	Mean	StdErr	Lower	Upper	P Value	F Value
T0	Low	Arts & Crafts	27.7%	2.6%	22.8%	33.1%	0.30	0.5836
	Low	Science	25.2%	3.6%	18.7%	33.0%		
T1	Low	Arts & Crafts	39.9%	3.1%	33.9%	46.2%	0.054	0.4629
	Low	Science	35.9%	4.4%	27.6%	45.0%		
T2	Low	Arts & Crafts	43.9%	3.2%	37.8%	50.3%	0.63	0.4305
	Low	Science	48.4%	4.7%	39.3%	57.6%		
T3	Low	Arts & Crafts	53.8%	3.2%	47.5%	60.0%	0.01	0.9213
	Low	Science	53.2%	4.7%	44.0%	62.3%		

Table 5.191 Reading achievement results: Group/Low achievers interaction

High and Low achievers progressed in a parallel way, except for High achievers in Science at T2, whose score was similar to the one attained by Science Low achievers. At the end of the study, the mean scores of the Low achievers in Science (53.2%) and in Arts & Crafts (53.8%) were almost the same.

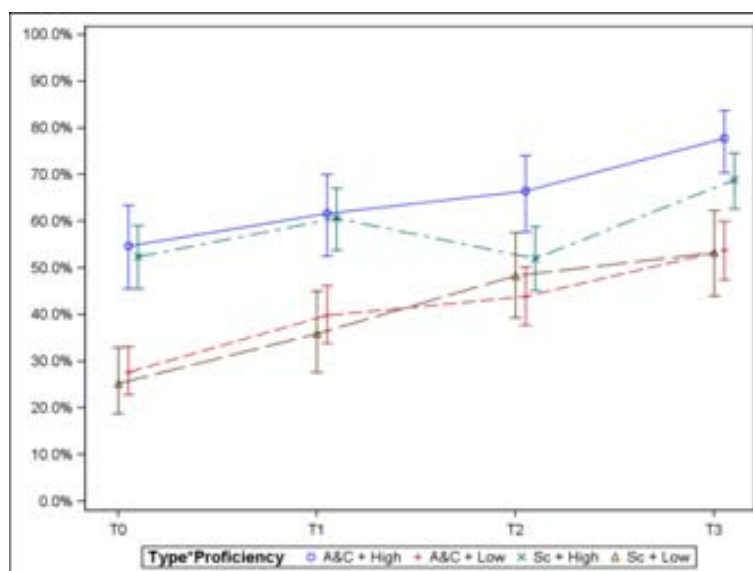


Figure 5.61 Reading achievement results. Group/Proficiency level interaction

5.3.2.1 Reading question 1

In reading question 1, students were asked to read and match a definition with the corresponding word. The intergroup comparison scores obtained at T0 show that, despite the slightly higher percentage obtained by the CLIL Science group (47.1%), the difference between the CLIL Science and the CLIL Arts & Crafts groups was not statistically significant ($F= 0.35$ $p= 0.5538$), which indicates that both groups started the study roughly with the same proficiency level as far as reading question 1 was concerned. The differences between the same groups at T1 ($F= 0.176$ $p= 0.1867$) and T2 ($F= 1.00$, $p= 0.3189$) were not significant either, although descriptive statistics showed that the CLIL Science group slightly outperforms the CLIL Arts & Crafts group. The results at T3 did not show a statistically significant advantage for any of the

groups in this question. Nevertheless, descriptively, the CLIL Science students obtained a better mean percentage score (74.6%) (see Table 5.192).

	Type	Mean	StdErr	Lower	Upper	F Value	P Value
T0	Arts & Crafts	43.2%	4.3%	35.0%	51.8%	0.35	0.5538
	Science	47.1%	5.0%	37.5%	57.0%		
T1	Arts & Crafts	57.7%	4.5%	48.8%	66.2%	1.76	0.1867
	Science	66.5%	4.7%	56.7%	75.1%		
T2	Arts & Crafts	51.4%	4.4%	42.8%	59.9%	1.00	0.3189
	Science	58.0%	4.8%	48.4%	67.1%		
T3	Arts & Crafts	69.1%	4.0%	60.7%	76.3%	0.95	0.3307
	Science	74.6%	3.8%	66.3%	81.4%		

Table 5.192 Reading achievement results: Reading Question1

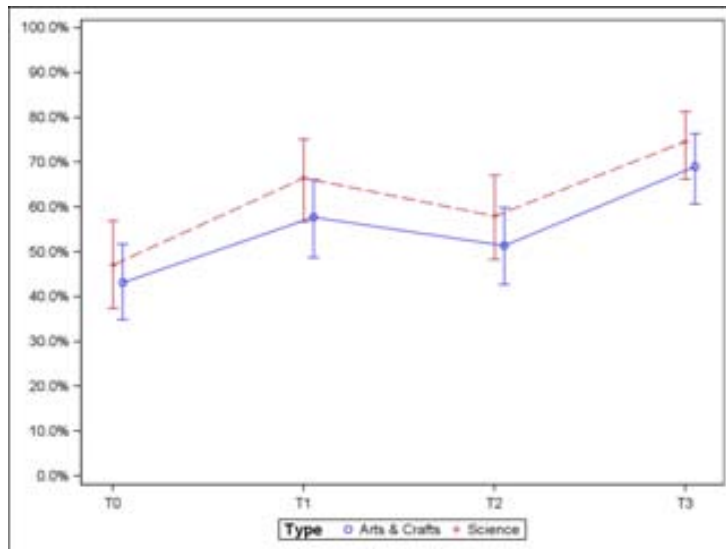


Figure 5.62 Reading achievement results: Reading Question1

5.3.2.2 Reading question 2

Reading question 2 was a True/False reading comprehension task. Students were asked to read five statements and to write Yes or No considering what they could see in the picture in the task. The intergroup analysis showed no statistically significant differences between the CLIL Science and the CLIL Arts & Crafts students at any of the times. Descriptive statistics revealed that at T0 and at T1 the achievement percentages

were higher for the CLIL Science group; however, at T2 and at T3 the mean percentages were slightly higher for the CLIL Arts & Crafts group. The percentage obtained by Science students at T3 (65.5%) was actually lower than the percentage at T0 (74.1%) (see Table 5.193 and Figure 5.63).

	Type	Mean	StdErr	Lower	Upper	F Value	P Value
T0	Arts & Crafts	66.2%	3.7%	56.7%	74.6%	2.21	0.2126
	Science	74.1%	3.7%	60.3%	84.4%		
T1	Arts & Crafts	69.4%	3.6%	60.3%	77.3%	1.73	0.2394
	Science	76.6%	3.9%	64.9%	85.3%		
T2	Arts & Crafts	60.3%	3.8%	50.4%	69.4%	0.29	0.6239
	Science	57.1%	4.3%	39.0%	73.5%		
T3	Arts & Crafts	68.6%	3.6%	59.4%	76.6%	0.31	0.6082
	Science	65.5%	4.1%	48.9%	79.0%		

Table 5.193 Reading achievement results: Reading Question 2

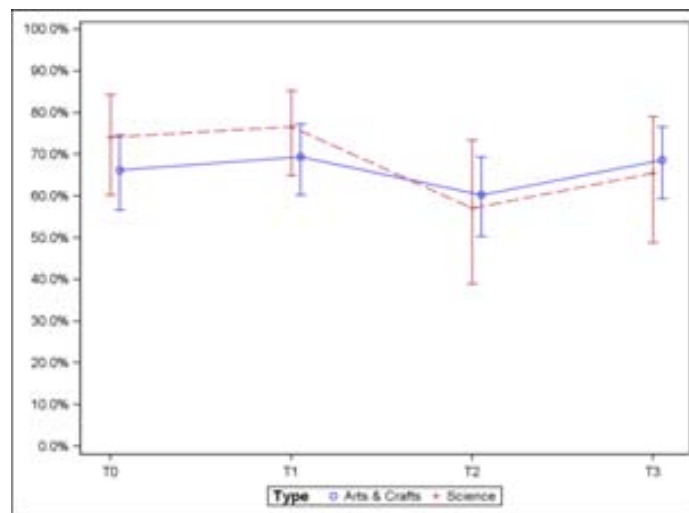


Figure 5.63 Reading achievement results: Reading Question 2

5.3.2.3 Reading question 3

Reading Question 3 was a gap filling exercise consisting of a text and some small pictures and words. Children had to decide which word to use and copy it into the gap. This exercise was a Cloze type test designed to test the students' use of English.

As can be seen in Table 5.194 and Figure 5.64, there were no statistically significant differences at any of the times tested. Descriptive statistics showed a slight advantage for the Arts & Crafts group at T0, T1 and at T2. The results at T3 are very similar, even though the mean percentage result was again slightly higher for the Arts & crafts group (1.1%).

	Type	Mean	StdErr	Lower	Upper	F Value	P Value
T0	Arts & Crafts	34.4%	4.2%	26.8%	43.0%	0.13	0.7229
	Science	32.1%	5.0%	23.2%	42.5%		
T1	Arts & Crafts	44.1%	4.6%	35.3%	53.3%	0.08	0.7737
	Science	42.0%	5.6%	31.6%	53.1%		
T2	Arts & Crafts	56.0%	4.6%	47.0%	64.7%	0.72	0.3966
	Science	50.1%	5.1%	40.1%	60.0%		
T3	Arts & Crafts	61.9%	4.4%	52.9%	70.1%	0.02	0.8757
	Science	60.8%	5.0%	50.7%	70.1%		

Table 5.194 Reading achievement results: Reading Question 3

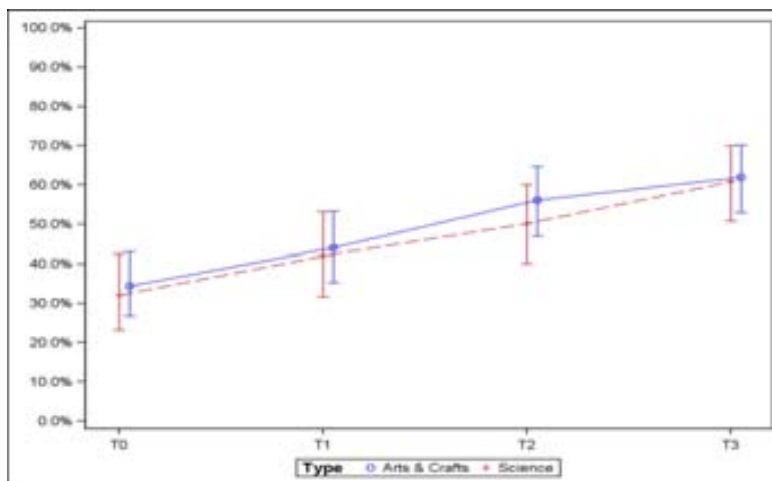


Figure 5.64 Reading achievement results: Reading Question 3

5.3.2.4 Reading question 4 (Comprehension)

This test was a story in three parts. Each part of the story had a matching picture. After each part, the students had to complete 10 sentences about the story using one,

two or three words. This task was scored under two different categories: comprehension and accuracy. This section will deal with the results in comprehension and the next section will address the results in accuracy.

As shown in Table 5.195 and Figure 5.65, no significant differences were found in the intergroup comparisons between CLIL Science and CLIL Arts & Crafts learners in this particular task. Descriptive statistics showed, however, an advantage for the CLIL Arts & Crafts Group at all the times tested. The difference between the scores at T3 was around 12%.

	Type	Mean	StdErr	Lower	Upper	F Value	P Value
T0	Arts & Crafts	33.1%	3.9%	26.0%	41.2%	2.17	0.1420
	Science	24.6%	4.1%	17.5%	33.4%		
T1	Arts & Crafts	50.2%	4.3%	41.7%	58.7%	3.48	0.0635
	Science	37.5%	5.0%	28.3%	47.6%		
T2	Arts & Crafts	59.7%	4.2%	51.2%	67.6%	3.35	0.0688
	Science	47.6%	4.9%	38.2%	57.2%		
T3	Arts & Crafts	69.0%	3.8%	61.0%	76.0%	3.30	0.0711
	Science	57.8%	4.7%	48.3%	66.7%		

Table 5.195 Reading achievement results: Reading Question 4 (Comprehension)

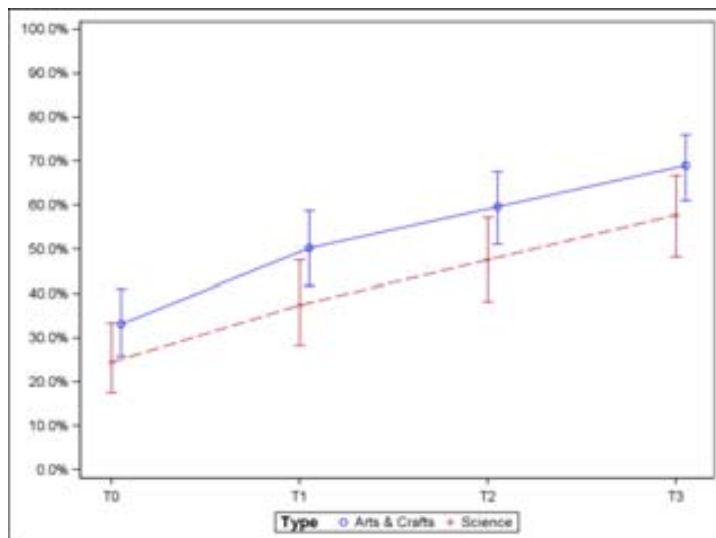


Figure 5.65 Reading achievement results: Reading Question 4 (Comprehension)

5.3.2.5. Reading question 4 (Accuracy)

As has been mentioned in the previous section, the answers provided by students in Reading Question 4 were analysed for accuracy, which means that only accurate answers were taken into consideration.

Table 5.196 shows significant differences in favour of the Arts & Crafts students at T0 (F= 8.57 p= 0.0037) and at T3 (F= 6.31 p= 0.0127).

	Type	Mean	StdErr	Lower	Upper	F Value	P Value
T0	Arts & Crafts	25.9%	3.2%	20.1%	32.7%	8.57	0.0037
	Science	12.7%	2.7%	8.3%	19.1%		
T1	Arts & Crafts	32.6%	3.6%	26.0%	40.0%	2.11	0.1477
	Science	24.7%	3.9%	17.9%	33.0%		
T2	Arts & Crafts	51.4%	3.9%	43.8%	59.0%	3.82	0.0520
	Science	39.7%	4.4%	31.5%	48.5%		
T3	Arts & Crafts	64.1%	3.7%	56.4%	71.0%	6.31	0.0127
	Science	49.0%	4.5%	40.3%	57.8%		

Table 5.196 Reading achievement results: Reading Question 4 (Accuracy)

Figure 5.66 shows that the groups developed in a parallel way and they did not seem to catch each other up at any of the times. The difference in the mean scores was almost always between 8% and 14%.

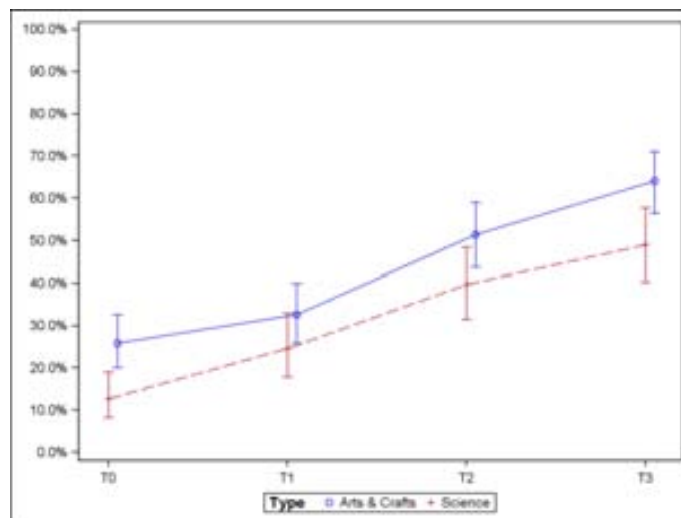


Figure 5.66 Reading achievement results: Reading Question 4 (Accuracy)

5.3.3 Summary of Listening and Reading Results

Tables 5.197 and 5.198 below display a summary of the results attained by CLIL Science and CLIL Arts & Crafts students in the Listening and Reading Tests.

	Listening (total)		Reading (Total)	
	Science	Arts&Crafts	Science	Arts&Crafts
T0	37.8%	38.4%	37.8%	40.5%
	p= 0.8950		p= 0.5318	
T1	51.0%	45.9%	48.1%	50.8%
	P Value0.3147		p= 0.5457	
T2	54.8%	54.4%	50.3%	55.5%
	P Value0.9365		p= 0.2305	
T3	71.8%	57.5%	61.4%	66.9%
	P Value 0.0018		p= 0.1761	

Table 5.197 Summary of Listening and Reading achievement results

Reading	Question1		Question2		Question 3		Question 4 (Comprehension)		Question 4 (Accuracy)	
	Science	Arts& Crafts	Science	Arts& Crafts	Science	Arts& Crafts	Science	Arts& Crafts	Science	Arts& Crafts
T0	47.1%	43.2%	74.1%	62.2%	32.1%	34.4%	24.6%	33.1%	12.7%	25.9%
	p= 0.5538		p= 2126		p= 0.7229		p= 0.1420		p= 0.0037	
T1	66.5%	57.7%	76.6%	69.4%	42.0%	44.1%	37.5%	50.2%	24.7%	36.2%
	p= 0.1867		p= 0.2394		p= 7737		p= 0.0635		p= 0.1477	
T2	58.0%	51.4%	57.1%	60.3%	50.1%	56.0%	47.6%	59.7%	39.7%	51.4%
	p= 0.3189		p= 0.6239		p= 3966		p= 0. 0688		p= 0.0520	
T3	74.6%	69.1%	65.5%	68.6%	60.8%	61.9%	57.8%	69.0%	49.0%	64.1%
	p= 0.3307		p= 0.6082		p=8757		P Value 0.0711		p= 0.0127	

Table 5.198 Summary of Reading achievement results: individual reading questions

Tables 5.199 and 5.200 below display a summary of the results attained by CLIL Science and CLIL Arts & Crafts students in Listening and Reading when the interaction Group/Proficiency level was taken into account.

Listening (Total)	Group/Proficiency level			
	High		Low	
	Science	Arts & Crafts	Science	Arts & Crafts
T0	53.0%	54.9%	24.6%	24.2%
	p= 0.7760		p= 0.9400	
T1	56.5%	50.8%	45.5%	41.1%
	p= 0.4006		p= 0.5002	
T2	59.8%	67.3%	49.7%	40.9%
	p= 0.2267		p= 0.1628	
T3	76.5%	68.9%	66.5%	45.3%
	p= 0.1686		p= 0.0009	

Table 5.199 Summary of Listening achievement results. Group/Proficiency level interaction

Reading (Total)	Group/Proficiency level			
	High		Low	
	Science	Arts&Crafts	Science	Arts&Crafts
T0	52.4%	54.7%	25.2%	27.7%
	p= 0.6943		p= 0.5836	
T1	60.6%	61.7%	35.9%	39.9%
	p= 0.8551		p= 0.4629	
T2	52.1%	66.4%	48.4%	43.9%
	p= 0.0124		p= 0.4305	
T3	69.0%	77.8%	53.2%	53.8%
	p= 0.0622		p= 0.9213	

Table 5.200 Summary of Reading achievement results. Group/Proficiency level interaction

5.3.4 Writing

In this section a description of the achievement³⁰ results obtained by CLIL Science students compared to those obtained by the CLIL Arts & Crafts group will be presented in order to answer the following research questions: RQ 3.2 In terms of writing measured as complexity, accuracy and fluency, are there differences in achievement between CLIL Science students and CLIL Arts & Crafts students at different times (T0, T1, T2 T3)? RQ 4.5 How does the initial level of English proficiency affect the students' performance in writing in the CLIL Science and the CLIL Arts & Crafts groups?

5.3.4.1 Fluency

This section presents the Fluency achievement results obtained by the CLIL Science and the CLIL Arts & Crafts groups. Four different measures were used to analyse Fluency: Total Number of Words (TNW), Total Number of Words in English (TNWE), Total Number of Units (TNU), and the Ratio between the Total Number of Words in English and the Total Number of Words (TNWE/TNW). For the TNWE and the TNU, as well as for the ratio TNWE/TNW, the results of the interaction between Group/and Proficiency Level (High and Low achievers) will also be reported.

5.3.4.1.1 Total Number of Words

Table 5.201 shows the results in achievement for each group regarding the TNW written by the students at different times. There were only statistically significant differences at T2 ($F= 6.63$ $p= 0.0107$) in favour of the Arts & Crafts group. Descriptive results show that at T0 and T1, the CLIL Arts & Crafts group wrote, on average, more

³⁰ As in the case of reading and listening, the improvement results for each of the groups, CLIL Science and CLIL Arts & crafts, have already been reported in the previous sections.

words than the CLIL Science students. However, at T3, the mean number of words written by the students in the CLIL Science group (120.4) was slightly higher than the mean for the Arts & Crafts group (117.7).

	Type	Mean	StdErr	Lower	Upper	F Value	P Value
T0	Arts & Crafts	71.69	5.39	61.06	82.32	1.32	0.2523
	Science	62.09	6.25	49.77	74.41		
T1	Arts & Crafts	78.95	5.39	68.32	89.58	3.13	0.0785
	Science	64.16	6.25	51.85	76.48		
T2	Arts & Crafts	105.2	5.39	94.59	115.8	6.63	0.0107
	Science	83.68	6.25	71.37	96.00		
T3	Arts & Crafts	117.7	5.47	106.9	128.5	0.10	0.7512
	Science	120.4	6.25	108.0	132.7		

Table 5.201 Fluency achievement results: TNW

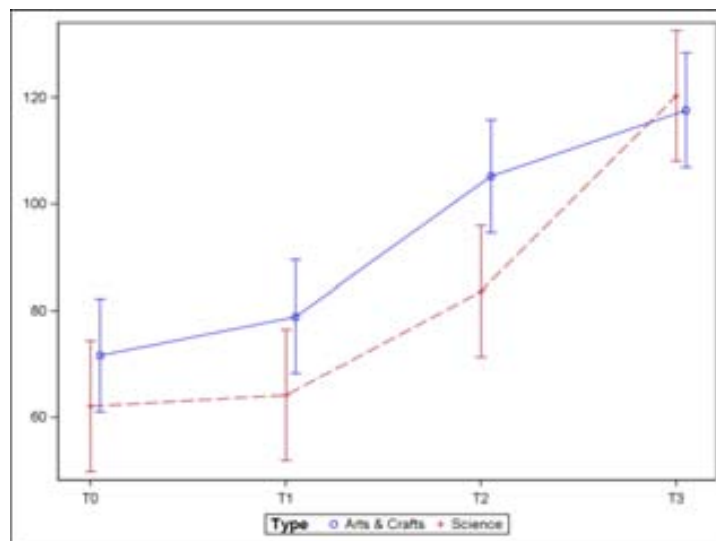


Figure 5.67 Fluency achievement results: TNW

5.3.4.1.2 Total Number of Words in English

As for TNWE, significant differences were found at T1 ($F= 4.24$ $p= 0.0408$) and at T2 ($F= 7.17$ $p= 0.0081$). Descriptive statistics reveal that, at T3, there are almost no differences in the mean number of words in English written by the groups. The CLIL

Science group wrote, on average, 107.8 words in English and the CLIL Arts & Crafts 106.7. The differences at T3 were not significant (see Table 5.202).

	Type	Mean	StdErr	Lower	Upper	F Value	P Value
T0	Arts & Crafts	61.03	4.94	51.28	70.78	2.42	0.1218
	Science	49.11	5.73	37.81	60.41		
T1	Arts & Crafts	70.77	4.94	61.02	80.52	4.24	0.0408
	Science	54.97	5.73	43.68	66.27		
T2	Arts & Crafts	91.71	4.94	81.96	101.5	7.17	0.0081
	Science	71.17	5.73	59.88	82.47		
T3	Arts & Crafts	106.7	5.01	96.86	116.6	0.02	0.8906
	Science	107.8	5.73	96.50	119.1		

Table 5.202 Fluency achievement results: TNWE

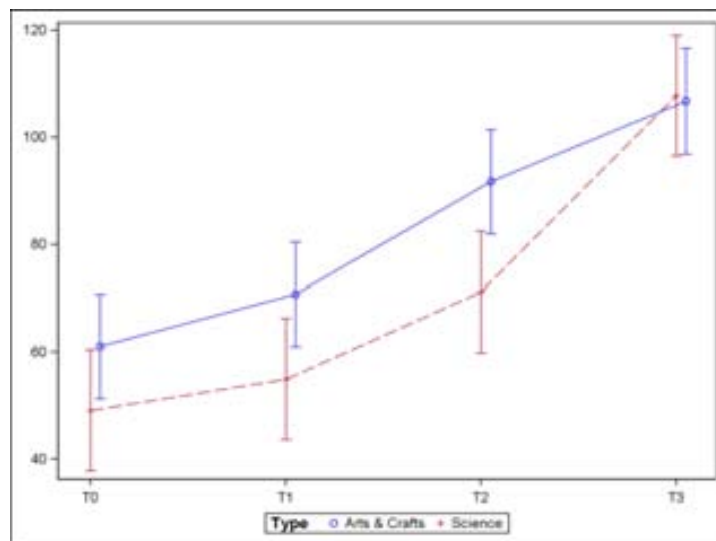


Figure 5.68 Fluency achievement results: TNWE

The interaction Group/High achievers in terms of TNWE can be seen in Table 5.203 and Figure 5.67 below. No significant differences were found between High achievers in the CLIL Science group and High achievers in the Arts & Crafts group. Even though results were not significant, descriptive statistics showed that High achievers in Arts & Crafts wrote on average more words in English than their counterparts in the Science group at T0, T1 and T2. At T3, however, CLIL Science

students seemed to catch up with CLIL Arts & Crafts student, and the differences between the groups become shorter: the mean difference was only 3.1.

	Proficiency	Type	Mean	StdErr	Lower	Upper	F Value	P Value
T0	High	Arts & Crafts	74.28	8.00	58.51	90.06	3.87	0.0507
	High	Science	54.13	6.24	41.81	66.44		
T1	High	Arts & Crafts	79.74	8.00	63.96	95.51	3.25	0.0731
	High	Science	61.27	6.24	48.95	73.59		
T2	High	Arts & Crafts	102.4	8.00	86.61	118.2	1.48	0.2255
	High	Science	89.92	6.24	77.61	102.2		
T3	High	Arts & Crafts	121.5	8.01	105.7	137.2	0.09	0.7679
	High	Science	118.4	6.25	106.1	130.7		

Table 5.203 Fluency achievement results: TNWE. Group/High achievers interaction

Table 5.204 and Figure 5.69 show the achievement results of the interaction Group/Low achievers for TNWE. As can be seen, there are only significant results at T2 ($F= 8.00$ $p= 0.0052$). Descriptive statistics, however, show that although at T0 and at T1 of data collection Low achievers in the Arts & Crafts group performed slightly better than their peers in the Science group, at T3 Low achievers in the CLIL Science group slightly outperformed their counterparts (97.17) in the Arts & Crafts group (92.02).

	Proficiency	Type	Mean	StdErr	Lower	Upper	F Value	P Value
T0	Low	Arts & Crafts	47.77	5.80	36.33	59.21	0.13	0.7161
	Low	Science	44.09	8.31	27.71	60.47		
T1	Low	Arts & Crafts	61.80	5.80	50.36	73.24	1.68	0.1959
	Low	Science	48.68	8.31	32.30	65.06		
T2	Low	Arts & Crafts	81.03	5.80	69.59	92.46	8.00	0.0052
	Low	Science	52.43	8.31	36.04	68.81		
T3	Low	Arts & Crafts	92.02	6.07	80.06	104.0	0.25	0.6157
	Low	Science	97.17	8.31	80.78	113.6		

Table 5.204 Fluency achievement results: TNWE. Group/Low achievers interaction

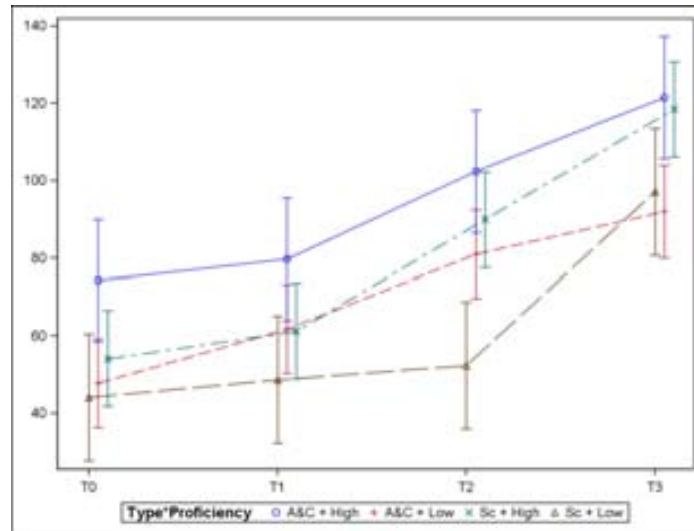


Figure 5.69 Fluency achievement results: TNWE. Group/Proficiency level interaction

5.3.4.1.3 Ratio: Total Number of Words in English/Total Number of Words

As a measure of fluency, the ratio between the Total Number of Words in English and the Total Number of Words was also calculated. The results at T0 were significant in favour of the CLIL Arts & Crafts group ($F= 5.27$ $p= 0.0224$). At T1 and T2, even though it was not significant, the ratio was still higher for the CLIL Arts & Crafts group. However, at T3, students in the CLIL Science group seemed to catch up with those in the CLIL Arts & Crafts group. The mean percentage difference at T3 was only 2% (see Table 5.205 and Figure 5.70).

	Type	Mean	StdErr	Lower	Upper	F Value	P Value
T0	Arts & Crafts	84.1%	1.9%	80.6%	87.9%	5.27	0.0224
	Science	77.0%	2.4%	72.4%	81.8%		
T1	Arts & Crafts	88.7%	1.8%	85.2%	92.4%	1.80	0.1809
	Science	84.6%	2.4%	79.9%	89.5%		
T2	Arts & Crafts	86.9%	1.6%	83.9%	90.1%	2.81	0.0949
	Science	82.3%	2.2%	78.1%	86.7%		
T3	Arts & Crafts	90.7%	1.6%	87.7%	93.9%	0.68	0.4110
	Science	88.7%	1.9%	85.1%	92.4%		

Table 5.205 Fluency achievement results: % TNWE/TNW.

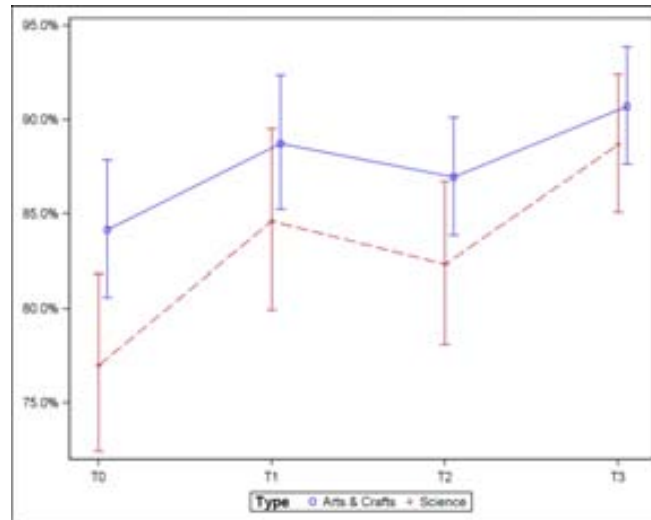


Figure 5.70 Fluency achievement results: % TNWE/TNW

5.3.4.1.4 Total Number of Units

Table 5.206 and Figure 5.71 below describe the results of the intergroup analyses as for the TNU. The table shows significant differences between the CLIL Science and the CLIL Arts & Crafts groups at T2 ($F= 13.96$ $p= 0.0002$) in favour of the Arts & Crafts group. Descriptive statistics revealed that, although the groups at T0 started at very similar points, achievement results for the Arts & Crafts group at T1 and at T3 were higher (19.46) than those of the CLIL Science Group (17.65).

	Type	Mean	StdErr	Lower	Upper	F Value	P Value
T0	Arts & Crafts	9.84	0.84	8.18	11.51	0.06	0.7993
	Science	9.51	0.98	7.58	11.44		
T1	Arts & Crafts	10.67	0.84	9.01	12.34	1.67	0.1977
	Science	8.98	0.98	7.05	10.91		
T2	Arts & Crafts	16.98	0.84	15.32	18.64	13.96	0.0002
	Science	12.09	0.98	10.16	14.02		
T3	Arts & Crafts	19.46	0.86	17.77	21.15	1.89	0.1703
	Science	17.65	0.98	15.72	19.58		

Table 5.206 Fluency achievement results: TNU

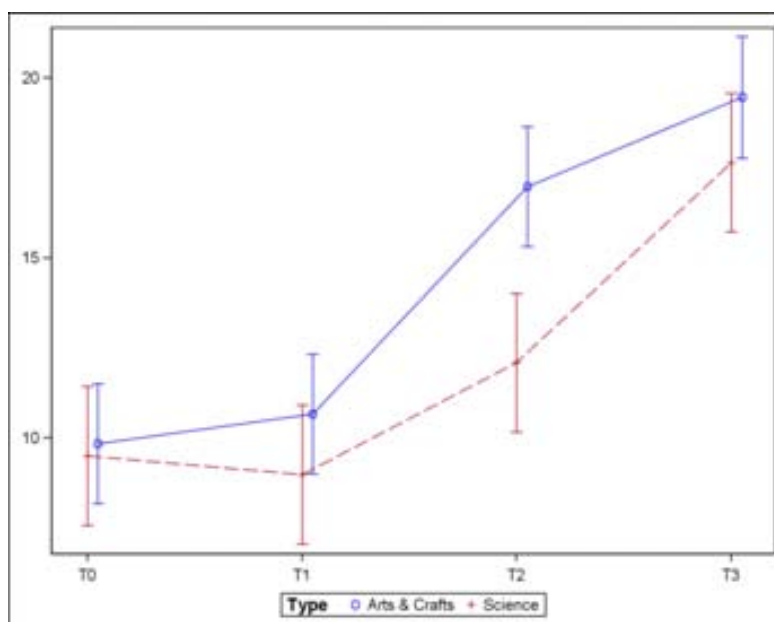


Figure 5.71 Fluency achievement results: TNU

The achievement results of the interaction Group/High achievers as for TNU can be seen in Table 5.207 and Figure 5.72 below. There were no statistically significant differences for High achievers when the interaction between Group/Proficiency level was taken into account. Descriptive statistics for the High achievers group show that, at T0, the mean score for the CLIL Arts & Crafts group was slightly higher than that of the CLIL Science Group. At T1 and at T2, the mean score was also higher for the Arts & Crafts group. At the end of the study, the mean number of units was still higher for the students in the CLIL Arts & Crafts group (22.14) than it was for the High achievers in the CLIL Science group (19.11).

	Proficiency	Type	Mean	StdErr	Lower	Upper	F Value	P Value
T0	High	Arts & Crafts	11.92	1.37	9.23	14.61	0.82	0.3667
	High	Science	10.34	1.07	8.24	12.44		
T1	High	Arts & Crafts	12.36	1.37	9.66	15.05	2.31	0.1304
	High	Science	9.70	1.07	7.60	11.80		
T2	High	Arts & Crafts	18.30	1.37	15.61	21.00	3.26	0.0723
	High	Science	15.14	1.07	13.04	17.24		
T3	High	Arts & Crafts	22.14	1.37	19.45	24.84	3.00	0.0849

Proficiency	Type	Mean	StdErr	Lower	Upper	F Value	P Value
High	Science	19.11	1.07	17.01	21.21		

Table 5.207 Fluency achievement results: TNU. Group/High achievers interaction

As in the case of High achievers, no statistically significant differences were found when the interaction Group/Low achievers was taken into account, except for T2 ($F= 14.71$ $p= 0.0002$). At T0, Low achievers in the CLIL Science group performed slightly better in terms of TNU than their counterparts in the CLIL Science group. results at T1 were almost the same for both groups. At T2, the mean number of units was statistically significant for the students in Arts & Crafts ($F= 14.71$ $p= 0.0002$). However, at the end of the study (T3), the average number of units written by the CLIL Science group (16.19) was almost the same as the number of units written by the CLIL Arts & Crafts group (16.19) (see Table 5.208).

	Proficiency	Type	Mean	StdErr	Lower	Upper	F Value	P Value
T0	Low	Arts & Crafts	7.76	0.99	5.81	9.71	0.28	0.5960
	Low	Science	8.68	1.42	5.88	11.47		
T1	Low	Arts & Crafts	8.99	0.99	7.04	10.94	0.18	0.6742
	Low	Science	8.26	1.42	5.47	11.06		
T2	Low	Arts & Crafts	15.66	0.99	13.71	17.61	14.71	0.0002
	Low	Science	9.04	1.42	6.24	11.83		
T3	Low	Arts & Crafts	16.78	1.04	14.73	18.83	0.11	0.7369
	Low	Science	16.19	1.42	13.39	18.99		

Table 5.208 Fluency achievement results: TNU. Group/Low achievers interaction

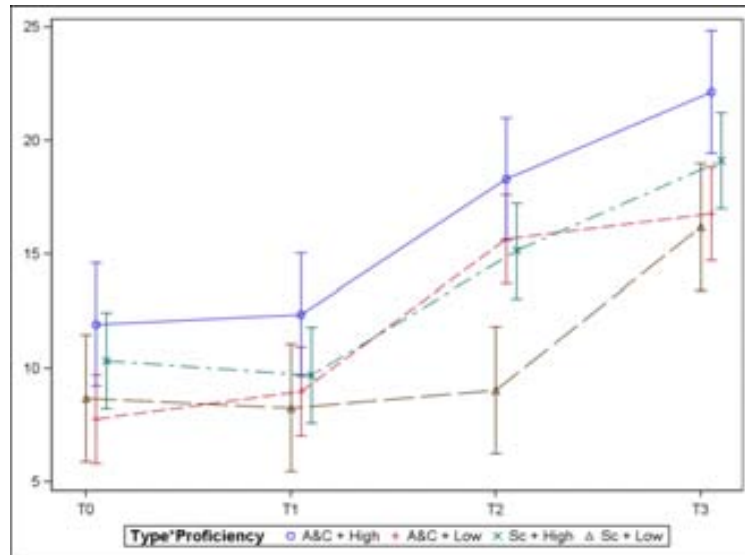


Figure 5.72 Fluency achievement results: TNU. Group/Proficiency level interaction

5.3.4.2. Accuracy

The results presented in this section show the differences between the CLIL Science students and the CLIL Arts & Crafts groups in the area of Accuracy. Two different measures were used to analyse Accuracy: Total Number of Error Free Units (TNEFU) and the ratio between the Total number of Error Free Units in relation to the Total Number of Units (TNEFU/TNU). As for TNEFU, the results of the interaction Group/Proficiency level will also be presented.

5.3.4.2.1 Total Number of Error Free Units

The intergroup comparison showed statistically significant differences between the groups at all the times tested: T0 ($F= 22.53$ $p= <.0001$), T1 ($F= 10.00$ $p= 0.0017$), T2 ($F= 29.65$ $p= <.0001$), T3 ($F= 12.24$ $p= 0.0006$). The scores obtained by CLIL Arts & Crafts students were much higher than the ones attained by the students in the CLIL Science group (see Table 5.209 and Figure 5.73). The difference between the mean scores at T3 was 5.5, which was even higher than the difference at T0 (3.71).

	Type	Mean	Stand Err	Lower	Upper	F Value	P Value
T0	Arts & Crafts	5.80	0.60	4.73	7.12	22.53	<.0001

	Type	Mean	Stand Err	Lower	Upper	F Value	P Value
T1	Science	2.09	0.39	1.45	3.02	10.00	0.0017
	Arts & Crafts	5.70	0.59	4.65	6.99		
T2	Science	3.07	0.50	2.23	4.24	29.65	<.0001
	Arts & Crafts	11.16	0.96	9.42	13.23		
T3	Science	4.45	0.64	3.35	5.90	12.24	0.0006
	Arts & Crafts	12.74	1.07	10.78	15.05		
	Science	7.69	0.88	6.13	9.65		

Table 5.209 Accuracy achievement results: TNEFU

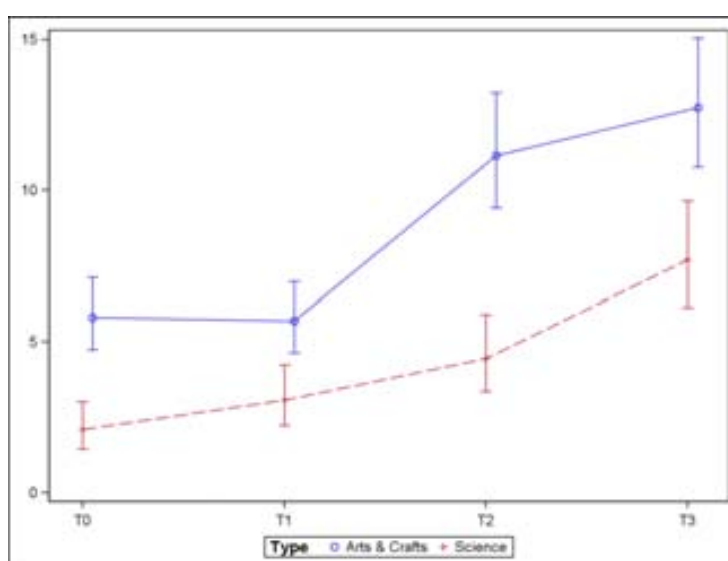


Figure 5.73 Accuracy achievement results: TNEFU

As can be seen from Table 5.210 below, there were statistically significant differences in favour of High achievers in CLIL Arts & Crafts classes at all the times tested: T0 (F= 25.70 p= <.0001), T1 (F= 5.08 p= 0.0251), T2 (F= 11.12 p= 0.0010), T3 (F=8.48 p= 0.0043). The mean scores for High achievers in Arts & Crafts (15.17) were much higher than the scores for High achievers in the Science group (9.02).

	Proficiency	Type	Mean	Std Err	Lower	Upper	F Value	P Value
T0	High	Arts & Crafts	7.26	1.13	5.33	9.88	25.70	<.0001
	High	Science	1.94	0.40	1.29	2.91		
T1	High	Arts & Crafts	6.69	1.06	4.89	9.15	5.08	0.0251
	High	Science	3.98	0.65	2.88	5.50		

	Proficiency	Type	Mean	Std Err	Lower	Upper	F Value	P Value
T2	High	Arts & Crafts	12.22	1.66	9.33	15.99	11.12	0.0010
	High	Science	6.30	0.90	4.76	8.34		
T3	High	Arts & Crafts	15.17	1.97	11.73	19.63	8.48	0.0043
	High	Science	9.02	1.09	7.11	11.45		

Table 5.210 Accuracy achievement results: TNEFU. Group/High achievers interaction

As in the case of High achievers, the results of the intergroup comparison showed significant differences in favour of Low achievers in the Arts & Crafts classes at all the times tested: T0 (F= 6.67 p= 0.0103), T1 (F= 7.41 p= 0.0069), T2 (F= 21.87 p= <.0001), T3 (F= 6.18 p= 0.0139).

	Proficiency	Type	Mean	Std Err	Lower	Upper	F Value	P Value
T0	Low	Arts & Crafts	4.64	0.60	3.59	6.00	6.67	0.0103
	Low	Science	2.26	0.56	1.39	3.66		
T1	Low	Arts & Crafts	4.85	0.62	3.77	6.25	7.41	0.0069
	Low	Science	2.37	0.54	1.51	3.72		
T2	Low	Arts & Crafts	10.20	1.06	8.31	12.51	26.87	<.0001
	Low	Science	3.14	0.63	2.11	4.67		
T3	Low	Arts & Crafts	10.69	1.14	8.67	13.19	6.18	0.0139
	Low	Science	6.55	1.08	4.73	9.08		

Table 5.211 Accuracy achievement results: TNEFU. Group/Low achievers interaction

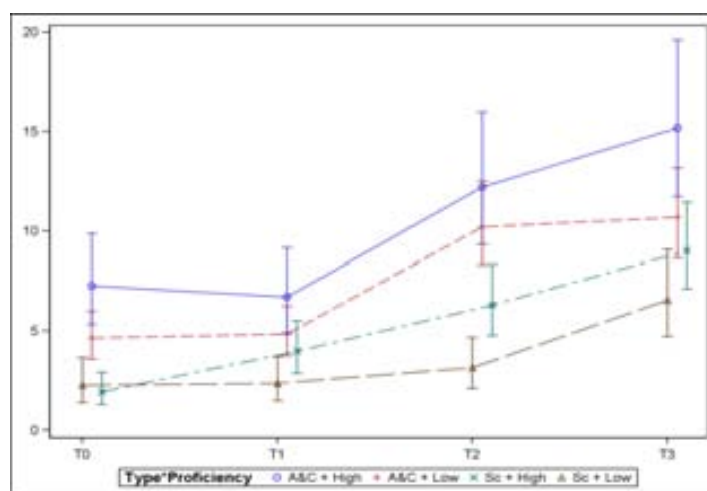


Figure 5.74 Accuracy achievement results: TNEFU. Group Proficiency level interaction

5.3.4.2.2 Ratio: Total Number of Error Free Units/Total Number of Units

As can be seen in Table 5.212, the ratio TNEFU/TNU shows statistically significant differences in favour of the CLIL Arts & Crafts group at all the times tested: T0 (F=34.03 p= <.0001), T1 (F=7.33 p=0.0072), T2 (F= 16.02 p=<.0001), T3 (F= 15.45 p=0.0001). Even though the ratio was much higher for the Arts & Crafts students at the beginning of the study, the Science students seemed to catch up with the Arts & Crafts ones at T3. The difference between the mean percentages at T3 (27.7%) is shorter than the difference at T0 (40%).

	Type	Mean	StdErr	Lower	Upper	F Value	P Value
T0	Arts & Crafts	63.6%	4.9%	54.6%	74.1%	34.03	<.0001
	Science	23.5%	3.6%	17.5%	31.6%		
T1	Arts & Crafts	55.9%	4.4%	48.0%	65.2%	7.33	0.0072
	Science	37.0%	4.8%	28.6%	47.7%		
T2	Arts & Crafts	67.9%	4.0%	60.4%	76.4%	16.02	<.0001
	Science	40.2%	4.6%	32.0%	50.4%		
T3	Arts & Crafts	69.2%	4.0%	61.8%	77.6%	15.45	0.0001
	Science	45.8%	4.0%	38.7%	54.3%		

Table 5.212 Accuracy achievement results: % TNEFU/TNU

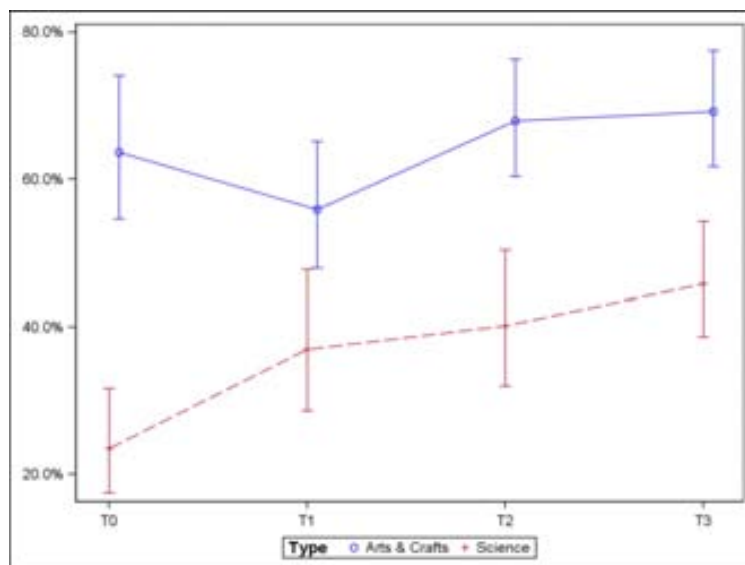


Figure 5.75 Accuracy achievement results: %TNEFU/TNU

The results of the intergroup comparison, when the interaction Group/High achievers was taken into account, shows significant differences at three different times. At T0, there were significant differences in favour of High achievers in the Arts & Crafts group ($F= 33.44$ $p<.0001$). At T1, no significant differences were found. However, at T2 and T3, the differences were again significant in favour of High achievers in the Arts & Crafts group: T2 ($F= 9.09$ $p= 0.0028$), T3 ($F= 9.04$ $p= 0.0030$). The mean percentage difference at T3 in the ratio TNWE/TNW is 22.1% (see Table 5.213 and Figure 5.76).

	Proficiency	Type	Mean	StdErr	Lower	Upper	F Value	P Value
T0	High	Arts & Crafts	65.1%	7.3%	52.2%	81.2%	33.44	<.0001
	High	Science	19.8%	3.4%	14.2%	27.6%		
T1	High	Arts & Crafts	55.8%	6.5%	44.4%	70.2%	2.03	0.1549
	High	Science	43.4%	5.7%	33.6%	56.1%		
T2	High	Arts & Crafts	68.7%	6.4%	57.1%	82.6%	9.09	0.0028
	High	Science	44.6%	4.8%	36.1%	55.1%		
T3	High	Arts & Crafts	71.1%	6.1%	60.0%	84.4%	9.04	0.0030
	High	Science	49.0%	4.3%	41.2%	58.2%		

Table 5.213 Accuracy achievement results: % TNEFU/TNU Group/High achievers interaction

The results of the comparison, taking into account the interaction Group/Low achievers, revealed significant advantages in favour of Low achievers in the Arts & Crafts group throughout the study: T0 ($F=13.09$ $p= 0.0003$), T1 ($F=7.80$ $p=0.00756$), T2 ($F= 11.95$ $p= 0.0006$), T3 ($F= 9.71$ $p=0.0021$). The mean difference at T3 for High achievers was around 22%, and the mean difference at T3 for Low achievers was around 25 % (see Table 5.214 and Figure 5.76).

	Proficiency	Type	Mean	StdErr	Lower	Upper	F Value	P Value
T0	Low	Arts & Crafts	62.2%	6.3%	51.0%	75.8%	13.09	0.0003
	Low	Science	27.9%	5.5%	18.9%	41.0%		

	Proficiency	Type	Mean	StdErr	Lower	Upper	F Value	P Value
T1	Low	Arts & Crafts	56.0%	5.5%	46.2%	67.9%	7.80	0.0056
	Low	Science	31.5%	5.7%	22.0%	44.9%		
T2	Low	Arts & Crafts	67.1%	4.9%	58.2%	77.5%	11.95	0.0006
	Low	Science	36.2%	5.9%	26.3%	49.9%		
T3	Low	Arts & Crafts	67.4%	5.0%	58.1%	78.1%	9.71	0.0021
	Low	Science	42.9%	5.3%	33.6%	54.7%		

Table 5.214 Accuracy achievement results: % TNEFU/TNU. Group/Low achievers interaction

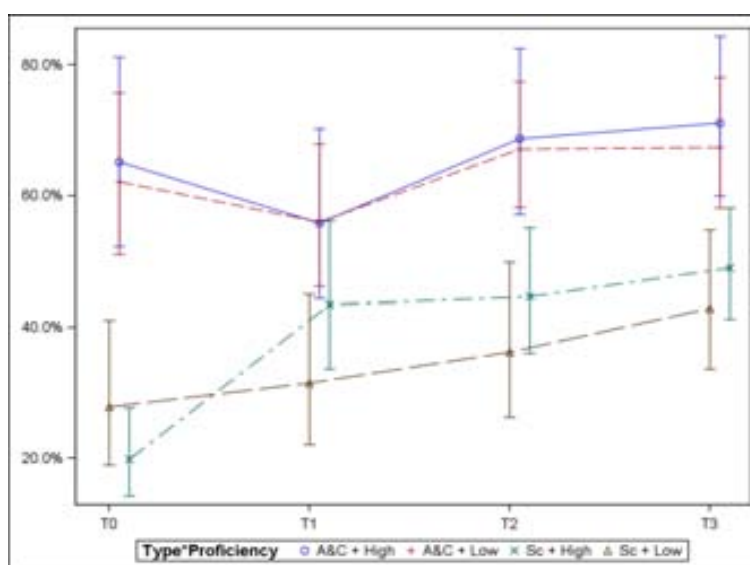


Figure 5.76 Accuracy achievement results: % TNEFU/TNU. Group/Proficiency level interaction

5.3.4.3 Complexity

5.3.4.3.1 Lexical Complexity

Two measures were used to determine lexical complexity: Total Number of Lexical Verbs (TNLV) and Total Number of Adjectives (TNAdj). The ratio between TNLV in relation to TNWE and the ratio between TNAdj in relation to TNWE will also be presented. The differences between High and Low achievers will be reported as well.

5.3.4.3.1.1 Total Number of Lexical Verbs

The comparison of the scores obtained showed that there were statistically significant differences at T2 ($F=9.98$ $p= 0.0018$) between the CLIL Science and the CLIL Arts & Crafts students. Even though descriptive statistics showed a slight advantage for CLIL Science students at T0, there were no differences in the mean number of lexical verbs at T1. As has been reported, at T2, the students in Arts & Crafts showed a significant advantage. However, at the end of the study, at T3, the CLIL Science students seemed to catch up with the Arts & Crafts students once again. The difference at T3 as for the mean number of lexical verbs was only 0.23 (see Table 5.215 and Figure 5.77).

	Type	Mean	Std Err	Lower	Upper	F Value	P Value
T0	Arts & Crafts	1.09	0.16	0.82	1.46	0.67	0.4123
	Science	1.32	0.23	0.94	1.87		
T1	Arts & Crafts	2.43	0.27	1.94	3.03	0.00	0.9957
	Science	2.43	0.34	1.85	3.19		
T2	Arts & Crafts	4.20	0.41	3.46	5.09	9.98	0.0018
	Science	2.43	0.34	1.85	3.21		
T3	Arts & Crafts	4.30	0.42	3.55	5.22	0.14	0.7126
	Science	4.07	0.48	3.22	5.13		

Table 5.215 Lexical Complexity achievement results: TNLV

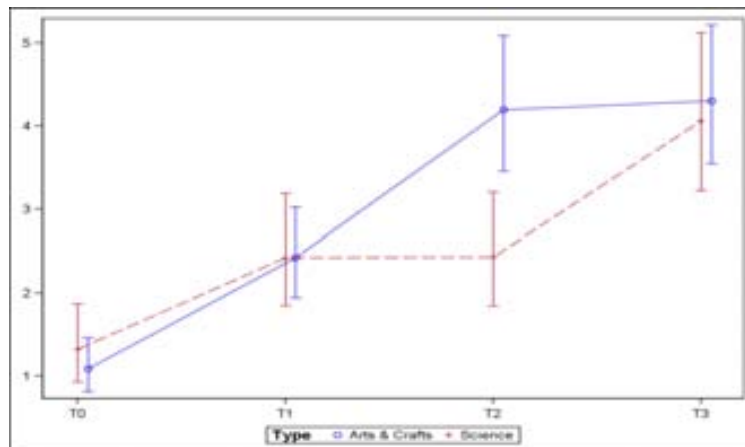


Figure 5.77 Lexical Complexity achievement results: TNLV

The intergroup comparison, when the interaction Group/High was taken into account, revealed no significant differences for High achievers. Table 5.216 shows that, even though High achievers in the CLIL Science group scored slightly better than their counterparts in Arts & crafts at T0 and at T1, CLIL High achievers in the Arts & Crafts group outperformed their peers at T2 in the mean number of adjectives used. However, at the end of the study, High achievers in the CLIL Science Group progressed and almost caught up with the students in Arts & Crafts. At the end of the study, the mean number of lexical verbs used by High achievers in Science (5.00) and High achievers in Arts & Crafts (5.21) was almost the same (see Table 5.216 and Figure 5.74).

	Proficiency	Type	Mean	Std Err	Lower	Upper	F Value	P Value
T0	High	Arts & Crafts	1.71	0.35	1.14	2.55	0.67	0.4135
	High	Science	2.12	0.35	1.54	2.94		
T1	High	Arts & Crafts	2.95	0.51	2.11	4.14	0.28	0.5958
	High	Science	3.32	0.45	2.54	4.34		
T2	High	Arts & Crafts	4.40	0.68	3.24	5.98	1.49	0.2246
	High	Science	3.41	0.47	2.60	4.47		
T3	High	Arts & Crafts	5.21	0.78	3.88	7.01	0.04	0.8335
	High	Science	5.00	0.60	3.94	6.35		

Table 5.216 Lexical Complexity achievement results: TNLV. Group/High achievers interaction

In the case of Low achievers, there were significant differences at T2 ($F= 12.64$ $p=0.0005$). At T0, there was a slight advantage for Low achievers in Science, but at T1 the advantage was for the students in Arts & Crafts. As has been reported, the difference at T2 was significantly different for the students in Arts & Crafts. At the end of the study, at T3, the difference was only 0.25 in favour of the students in CLIL Arts & Crafts.

	Proficiency	Type	Mean	Std Err	Lower	Upper	F Value	P Value
T0	Low	Arts & Crafts	0.70	0.15	0.46	1.07	0.24	0.6234
	Low	Science	0.82	0.21	0.50	1.37		
T1	Low	Arts & Crafts	1.99	0.29	1.50	2.64	0.21	0.6476
	Low	Science	1.78	0.37	1.18	2.67		
T2	Low	Arts & Crafts	4.01	0.47	3.18	5.05	12.64	0.0005
	Low	Science	1.74	0.35	1.16	2.60		
T3	Low	Arts & Crafts	3.55	0.45	2.77	4.55	0.12	0.7336
	Low	Science	3.30	0.58	2.34	4.66		

Table 5.217 Lexical Complexity achievement results: TNLV. Group/Low achievers interaction

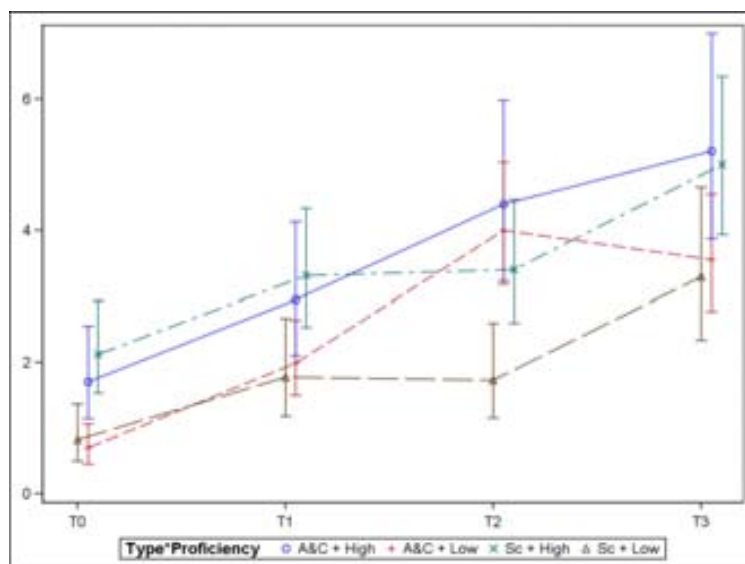


Figure 5.78 Lexical Complexity achievement results: TNLV. Group/Proficiency level interaction

5.3.4.3.1.2 Ratio: Total Number of Lexical Verbs/Total Number of Words in English

As can be seen in Table 5.218, the achievement results in the ratio TNLV/TNWE were not significant at any of the times. Descriptive statistics show that the mean percentage at T0 was much lower for the CLIL Arts & Crafts group than it was for the CLIL Science Group. However, even though at T2 the mean percentage of

the CLIL Arts & Crafts group was higher, the mean percentages of both groups were almost the same at T3.

	Type	Mean	StdErr	Lower	Upper	F Value	P Value
T0	Arts & Crafts	1.9%	0.3%	1.4%	2.6%	3.30	0.0704
	Science	3.0%	0.5%	2.1%	4.2%		
T1	Arts & Crafts	3.6%	0.4%	3.0%	4.5%	3.41	0.0659
	Science	4.9%	0.6%	3.8%	6.4%		
T2	Arts & Crafts	4.8%	0.4%	4.1%	5.6%	1.62	0.2036
	Science	3.9%	0.5%	3.0%	5.1%		
T3	Arts & Crafts	4.3%	0.4%	3.6%	5.0%	0.06	0.8087
	Science	4.1%	0.4%	3.4%	5.0%		

Table 5.218 Lexical Complexity achievement results: %TNLV/TNWE

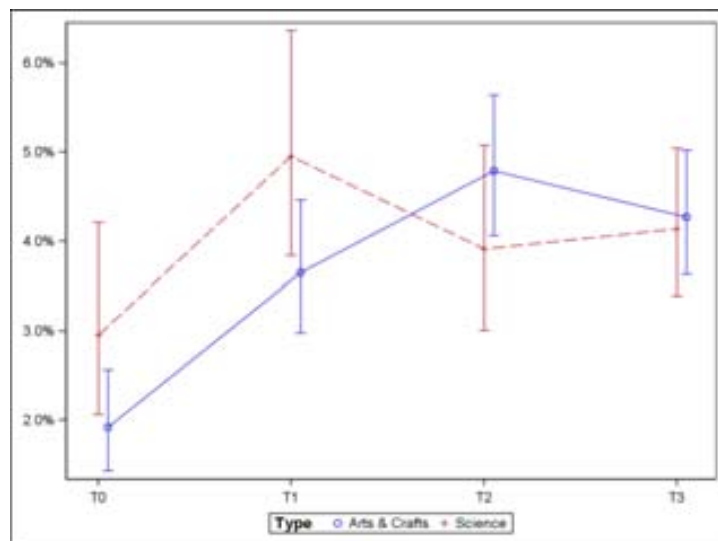


Figure 5.79 Lexical Complexity achievement results: %TNLV/TNWE

The intergroup comparison, when the interaction Group/High achievers was taken into consideration, showed significant differences at T0 ($F= 4.91$ $p= 0.0275$) and T1 ($F= 3.98$ $p= 0.0469$) in favour of the High achievers in the CLIL Science Group. Descriptive statistics revealed that at T2 and T3, the High achievers in Arts & Crafts seemed to catch up with their counterparts in the Science group. At T3, there was no

difference in the mean percentage between the High achievers in Science and the High achievers in Arts & Crafts (4.5%)

	Proficiency	Type	Mean	StdErr	Lower	Upper	F Value	P Value
T0	High	Arts & Crafts	2.4%	0.5%	1.6%	3.5%	4.91	0.0275
	High	Science	4.2%	0.7%	3.1%	5.7%		
T1	High	Arts & Crafts	4.0%	0.6%	2.9%	5.3%	3.98	0.0469
	High	Science	5.8%	0.7%	4.6%	7.4%		
T2	High	Arts & Crafts	4.5%	0.6%	3.4%	5.8%	0.13	0.7202
	High	Science	4.2%	0.5%	3.3%	5.3%		
T3	High	Arts & Crafts	4.5%	0.5%	3.5%	5.7%	0.00	0.9961
	High	Science	4.5%	0.4%	3.7%	5.5%		

Table 5.219 Lexical Complexity achievement results: %TNLV/TNWE. Group/High achievers interaction

As for Low achievers, even though there were no significant differences, at T0 and T1, the mean percentages of the Low achievers in Arts & Crafts were lower than the mean percentages of the Low achievers in Science. However, the mean percentages of the Low achievers in Arts & crafts at T2 and T3 were slightly higher than those of the Low achievers in the CLIL Science group.

	Proficiency	Type	Mean	StdErr	Lower	Upper	F Value	P Value
T0	Low	Arts & Crafts	1.5%	0.3%	1.0%	2.4%	0.75	0.3864
	Low	Science	2.1%	0.6%	1.2%	3.5%		
T1	Low	Arts & Crafts	3.4%	0.5%	2.6%	4.4%	0.88	0.3499
	Low	Science	4.2%	0.8%	2.9%	6.1%		
T2	Low	Arts & Crafts	5.1%	0.5%	4.2%	6.3%	2.40	0.1222
	Low	Science	3.7%	0.7%	2.5%	5.4%		
T3	Low	Arts & Crafts	4.1%	0.5%	3.3%	5.0%	0.09	0.7617
	Low	Science	3.8%	0.6%	2.8%	5.1%		

Table 5.220 Lexical Complexity achievement results: %TNLV/TNWE. Group/Low achievers interaction

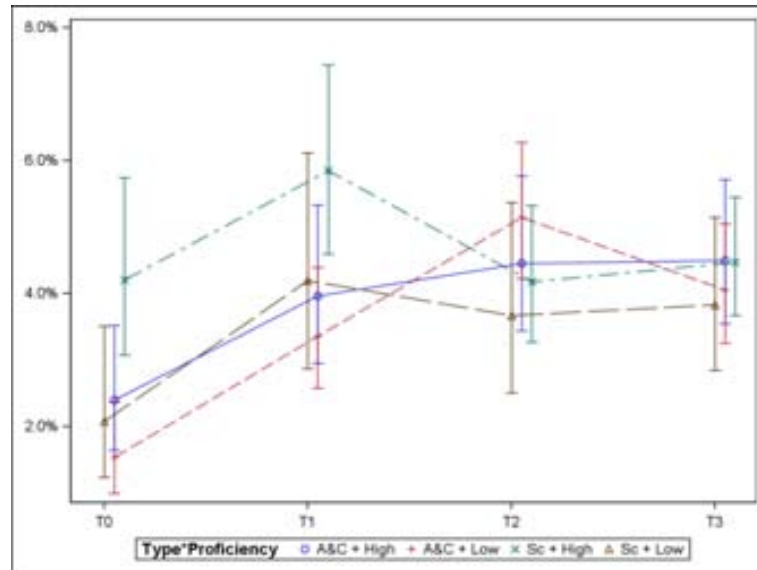


Figure 5.80 Lexical Complexity achievement results: %TNLV/TNWE. Group/Proficiency level interaction

5.3.4.3.1.3 Total Number of Adjectives

As can be seen in Table 5.221, there were statistically significant differences at all times tested in favour of the CLIL Arts & Crafts group: T0 (F= 16.35 p= <.0001), T1 (F= 10.74 p= 0.0012), T2 (F= 10.70 p= 0.0012), T3 (F= .5.01 p= 0.02649).

Type	Mean	Std Err	Lower	Upper	F Value	P Value
T0 Arts & Crafts	2.95	0.33	2.36	3.69	16.35	<.0001
T0 Science	1.11	0.23	0.74	1.68		
T1 Arts & Crafts	1.83	0.25	1.40	2.41	10.74	0.0012
T1 Science	0.68	0.18	0.41	1.14		
T2 Arts & Crafts	3.37	0.36	2.72	4.16	10.70	0.0012
T2 Science	1.73	0.29	1.24	2.41		
T3 Arts & Crafts	4.83	0.46	4.00	5.83	5.01	0.0264
T3 Science	3.37	0.42	2.63	4.32		

Table 5.221 Lexical Complexity achievement results: TNAdj

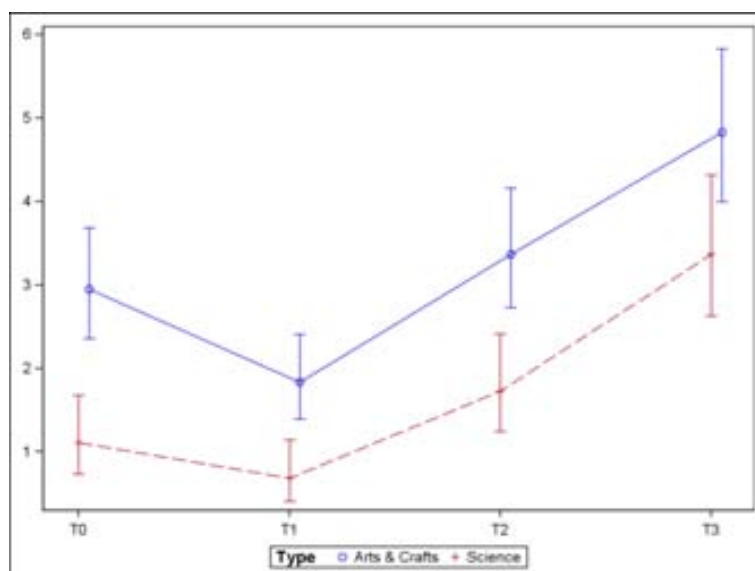


Figure 5.81 Lexical Complexity achievement results: TNAdj

Table 5.222 and Figure 5.82 below show the results of the intergroup comparison when the interaction Group/High achievers was taken into account. Except for T0, at which the mean number of adjectives was significantly in favour of the High achievers in the Arts & Crafts group ($F= 11.66$ $p= 0.0007$), there were no statistically significant differences between the groups the rest of the times. Descriptively, the High achievers in the CLIL Arts & Crafts group used more adjectives on average than the High achievers in the CLIL Science Group.

	Proficiency	Type	Mean	StdErr	Lower	Upper	F value	P Value
T0	High	Arts & Crafts	3.35	0.59	2.37	4.74	11.66	0.0007
	High	Science	1.30	0.27	0.86	1.96		
T1	High	Arts & Crafts	1.70	0.38	1.09	2.65	3.53	0.0606
	High	Science	0.89	0.23	0.54	1.47		
T2	High	Arts & Crafts	3.44	0.60	2.45	4.85	1.88	0.1714
	High	Science	2.48	0.39	1.82	3.40		
T3	High	Arts & Crafts	5.42	0.81	4.04	7.27	3.10	0.0802
	High	Science	3.81	0.50	2.94	4.93		

Table 5.222 Lexical Complexity achievement results: TNAdj. Group/High achievers interaction

The Low achievers in Arts & Crafts scored significantly higher at three of the times tested: T0 (F= 9.43 p= 0.0023), T1 (F= 10.72 p= 0.0012), T2 (F= 12.19 p= 0.0005) (see Table 5.223 and Figure 5.82).

	Proficiency	Type	Mean	Std Err	Lower	Upper	F Value	P Value
T0	Low	Arts & Crafts	2.59	0.36	1.98	3.40	9.43	0.0023
	Low	Science	0.95	0.28	0.53	1.70		
T1	Low	Arts & Crafts	1.98	0.30	1.46	2.67	10.72	0.0012
	Low	Science	0.52	0.20	0.25	1.10		
T2	Low	Arts & Crafts	3.29	0.41	2.57	4.22	12.29	0.0005
	Low	Science	1.21	0.31	0.73	2.00		
T3	Low	Arts & Crafts	4.30	0.52	3.39	5.45	2.65	0.1048
	Low	Science	2.98	0.56	2.05	4.33		

Table 5.223 Lexical Complexity achievement results: TNAdj. Group/Low achievers interaction

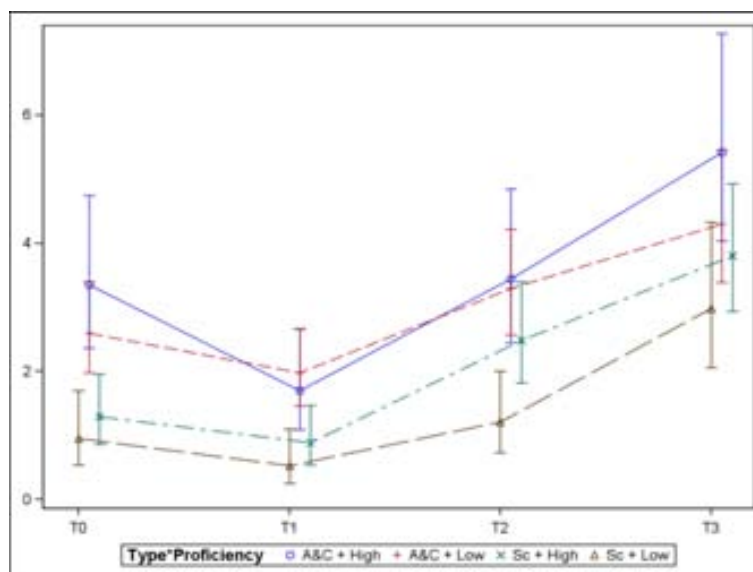


Figure 5.82 Lexical Complexity achievement results: TNAdj. Group/Proficiency level interaction

5.3.4.3.1.4 Ratio: Total Number of Adjectives/Total Number of Words in English

Table 5.224 and Figure 5.83 show the results of the ratio TNAdj/TNWE. There were significant differences in favour of the CLIL Arts & Crafts group at T0 (F= 11.47

p= 0.0008), T1 (F= 6-13 p= 0.0138) and T3 (F= 6.64 p= 0.0106) but no differences at T2. It is interesting to notice that the mean percentage of the CLIL Arts & crafts groups was slightly higher at T0 (5.3%) than at T3 (4.8%), whereas the mean percentage for the Science students was higher at T3 (3.2%) than at T0.

	Type	Mean	StdErr	Lower	Upper	F Value	P Value
T0	Arts & Crafts	5.3%	0.6%	4.3%	6.5%	11.47	0.0008
	Science	2.3%	0.5%	1.5%	3.5%		
T1	Arts & Crafts	2.8%	0.4%	2.1%	3.6%	6.13	0.0138
	Science	1.3%	0.3%	0.8%	2.2%		
T2	Arts & Crafts	3.8%	0.4%	3.1%	4.6%	3.39	0.0668
	Science	2.6%	0.4%	1.9%	3.7%		
T3	Arts & Crafts	4.8%	0.4%	4.0%	5.7%	6.64	0.0106
	Science	3.2%	0.4%	2.6%	4.1%		

Table 5.224 Lexical Complexity achievement results: %TNAdj/TNWE

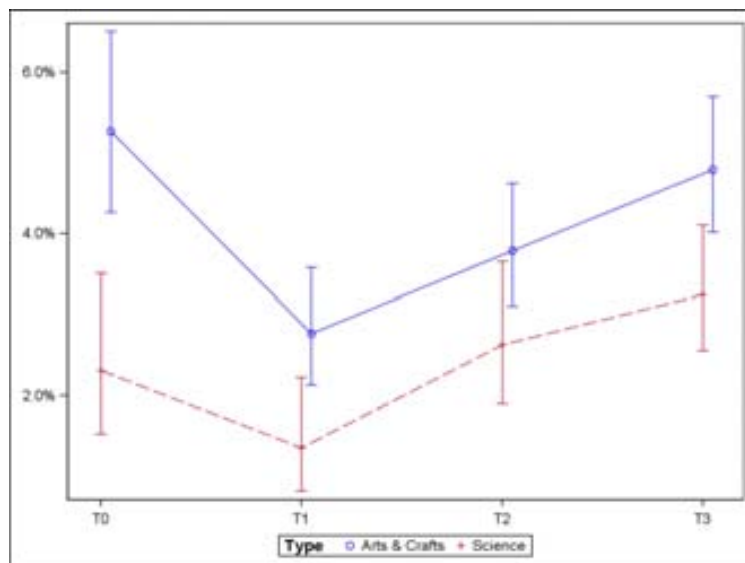


Figure 5.83 Lexical Complexity achievement results: %TNAdj/TNWE

The results of the comparison when the interaction Group/High achievers was taken into account showed significant differences at T1 and T3 in favour of the CLIL Arts & Crafts students: T1 (F= 7.01 p= 0.0085), T3 (F= 3.44 p= 0.0653). Even though at T0 the students in Arts & Crafts doubled the mean percentage of the students in

Science, at T3, the differences in the mean percentages were not very important, although they were significant for the Arts & Crafts group.

	Proficiency	Type	Mean	StdErr	Lower	Upper	F Value	P Value
T0	High	Arts & Crafts	5.0%	0.8%	3.6%	6.8%	7.01	0.0085
	High	Science	2.5%	0.5%	1.7%	3.7%		
T1	High	Arts & Crafts	2.3%	0.5%	1.5%	3.5%	1.48	0.2253
	High	Science	1.5%	0.4%	0.9%	2.5%		
T2	High	Arts & Crafts	3.5%	0.6%	2.6%	4.9%	0.58	0.4475
	High	Science	3.0%	0.4%	2.2%	4.0%		
T3	High	Arts & Crafts	4.6%	0.6%	3.5%	6.1%	3.44	0.0653
	High	Science	3.3%	0.4%	2.6%	4.2%		

Table 5.225 Lexical Complexity achievement results: %TNAdj/TNWE. Group/High achievers interaction

The results of the comparison for the Low achievers were significantly different at all times tested in favour of the CLIL Arts & Crafts group: T0 (F= 8.12 p= 0.0047), T1 (F= 6.97 p= 0.0087), T2 (F= 3.92 p= 0.0486), T3 (F= 4.14 p= 0.0431).

	Proficiency	Type	Mean	StdErr	Lower	Upper	F Value	P Value
T0	Low	Arts & Crafts	5.6%	0.7%	4.3%	7.3%	8.12	0.0047
	Low	Science	2.2%	0.7%	1.2%	3.9%		
T1	Low	Arts & Crafts	3.3%	0.5%	2.5%	4.5%	6.97	0.0087
	Low	Science	1.2%	0.4%	0.6%	2.4%		
T2	Low	Arts & Crafts	4.1%	0.5%	3.2%	5.2%	3.92	0.0486
	Low	Science	2.3%	0.6%	1.4%	3.9%		
T3	Low	Arts & Crafts	4.9%	0.6%	4.0%	6.2%	4.14	0.0431
	Low	Science	3.2%	0.6%	2.2%	4.6%		

Table 5.226 Lexical Complexity achievement results: %TNAdj/TNWE. Group/Low achievers interaction

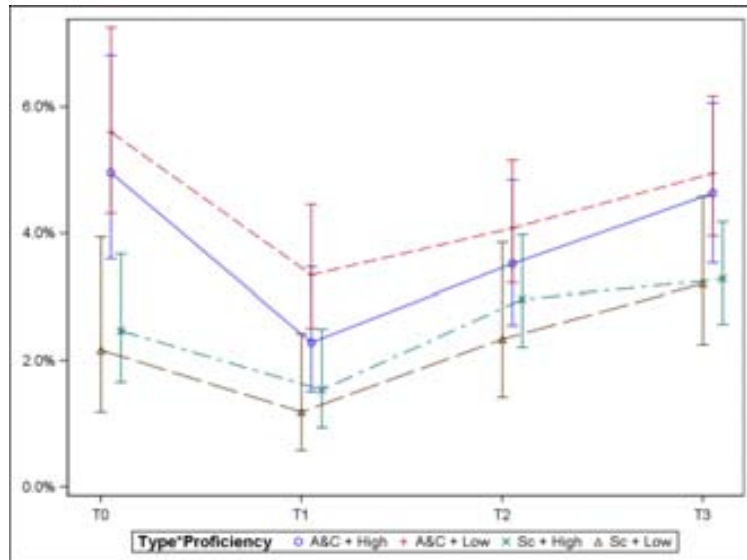


Figure 5.84 Lexical Complexity achievement results: %TNAdj/TNWE. Group/Proficiency level interaction

5.3.4.3.2 Syntactic Complexity

Two measures were taken into consideration for the analysis of syntactic complexity: Instances of Coordinated Units (ICU), that is, the number of examples of coordination, and Instances of Subordinate Units (ISU), the number of examples of subordination.

5.3.4.3.2.1 Instances of Coordinated Units

None of the achievement results for ICU was statistically significant at any of the times tested. The mean results obtained by the CLIL Arts & Crafts group were slightly higher at T0 and T1 than those obtained by the Science Group. However, at T2 and T3, it was the Science group that showed slightly more positive mean results than the Arts & Crafts group (see Table 5.227 and Figure 5.85).

	Type	Mean	Std Err	Lower	Upper	F Value	P Value
T0	Arts & Crafts	1.11	0.20	0.78	1.57	0.00	0.9826
	Science	1.12	0.24	0.73	1.71		
T1	Arts & Crafts	1.32	0.22	0.95	1.84	1.99	0.1592
	Science	0.86	0.21	0.53	1.41		

	Type	Mean	Std Err	Lower	Upper	F Value	P Value
T2	Arts & Crafts	2.07	0.29	1.57	2.73	0.09	0.7601
	Science	2.22	0.37	1.60	3.07		
T3	Arts & Crafts	2.27	0.31	1.73	2.97	1.32	0.2527
	Science	2.88	0.43	2.13	3.88		

Table 5.227 Syntactic Complexity achievement results: ICU

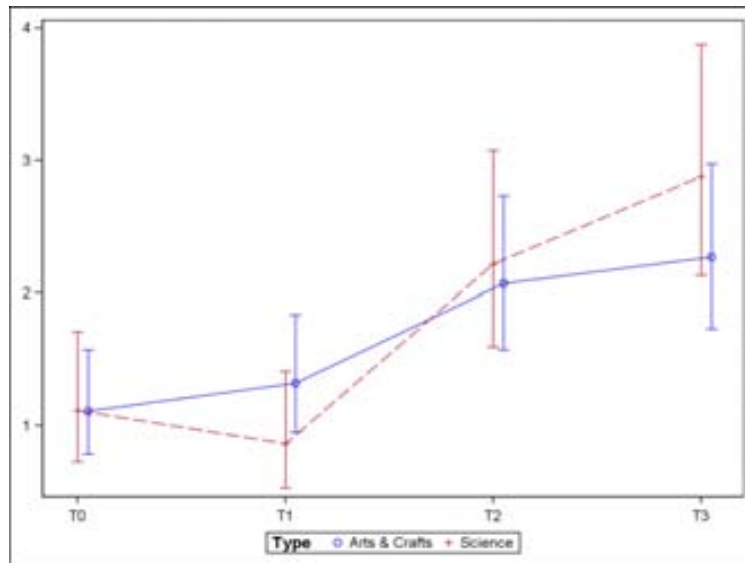


Figure 5.85 Syntactic Complexity achievement results: ICU

The results of the interaction Group/High achievers show that there were no statistically significant differences at any of the times tested in terms of ICU. Descriptive statistics reveal that at T2, the High achievers in Science obtained higher mean results (3.00) than the High achievers in the Arts & Crafts group (2.69). Results at T3, however, were almost exactly the same as results at T2 for both groups (see Table 5.228).

	Proficiency	Type	Mean	Std Err	Lower	Upper	F Value	P Value
T0	High	Arts & Crafts	1.73	0.42	1.07	2.80	0.00	0.9887
	High	Science	1.72	0.33	1.18	2.53		
T1	High	Arts & Crafts	1.55	0.39	0.95	2.56	0.03	0.8592
	High	Science	1.46	0.32	0.95	2.26		

	Proficiency	Type	Mean	Std Err	Lower	Upper	F Value	P Value
T2	High	Arts & Crafts	2.69	0.56	1.78	4.07	0.17	0.6799
	High	Science	3.00	0.48	2.19	4.12		
T3	High	Arts & Crafts	2.68	0.56	1.78	4.04	0.19	0.6602
	High	Science	3.01	0.48	2.20	4.12		

Table 5.228 Lexical Complexity achievement results: ICU. Group/High achievers interaction

As in the case of High achievers, no significant differences were found between the groups. Even though the Low achievers in the Control Group obtained the same mean differences at T0 and at T2, the Low achievers in CLIL Science performed better than the Low achievers in Arts & crafts at T3 (see Table 5.229).

	Proficiency	Type	Mean	StdErr	Lower	Upper	F Value	P Value
T0	Low	Arts & Crafts	0.71	0.18	0.43	1.17	0.00	0.9689
	Low	Science	0.72	0.25	0.37	1.43		
T1	Low	Arts & Crafts	1.13	0.23	0.75	1.70	3.37	0.0674
	Low	Science	0.51	0.19	0.24	1.07		
T2	Low	Arts & Crafts	1.60	0.29	1.11	2.29	0.01	0.9379
	Low	Science	1.64	0.41	1.00	2.69		
T3	Low	Arts & Crafts	1.92	0.35	1.35	2.74	1.56	0.2134
	Low	Science	2.75	0.61	1.78	4.25		

Table 5.229 Syntactic Complexity achievement results: ICU. Group/Low achievers interaction

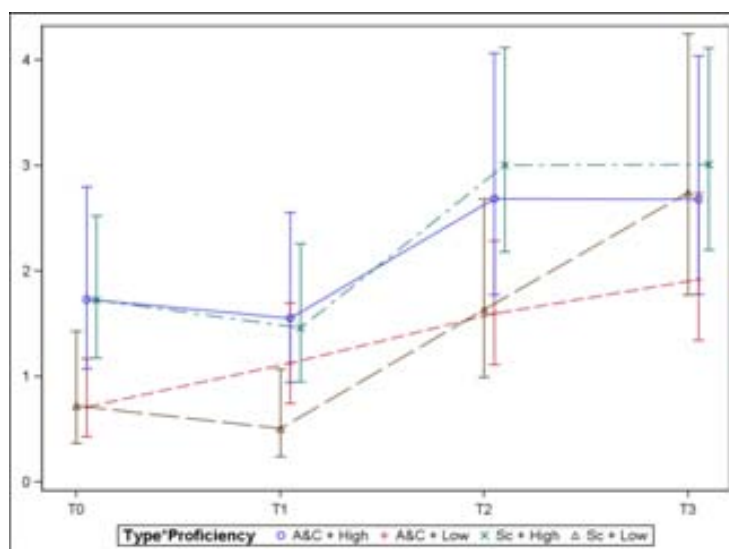


Figure 5.86 Syntactic Complexity achievement results: ICU. Group/Proficiency level interaction

5.3.4.3.2.2 Instances of Subordinate Units

There were no instances of subordination at T0. Significant differences were found at T1 in favour of the Arts & Crafts group. The CLIL Science group seemed to slightly outperform the CLIL Arts & crafts group at T3, although the differences were not statistically significant. It is important to notice that, at T3, mean scores for the CLIL Science and the CLIL Arts & Crafts groups were lower than the score obtained by the Arts & Crafts group at T2. However, the scores of both groups at T3 were very similar (see Table 5.230 and Figure 5.87).

	Type	Mean	Std Err	Lower	Upper	F Value	P Value
T1	Arts & Crafts	0.34	0.10	0.19	0.60	1.80	0.1813
	Science	0.15	0.08	0.05	0.44		
T2	Arts & Crafts	1.48	0.29	1.00	2.20	6.66	0.0108
	Science	0.57	0.18	0.31	1.05		
T3	Arts & Crafts	1.19	0.25	0.78	1.81	0.01	0.9386
	Science	1.22	0.31	0.74	2.02		

Table 5.230 Syntactic Complexity achievement results: ISU

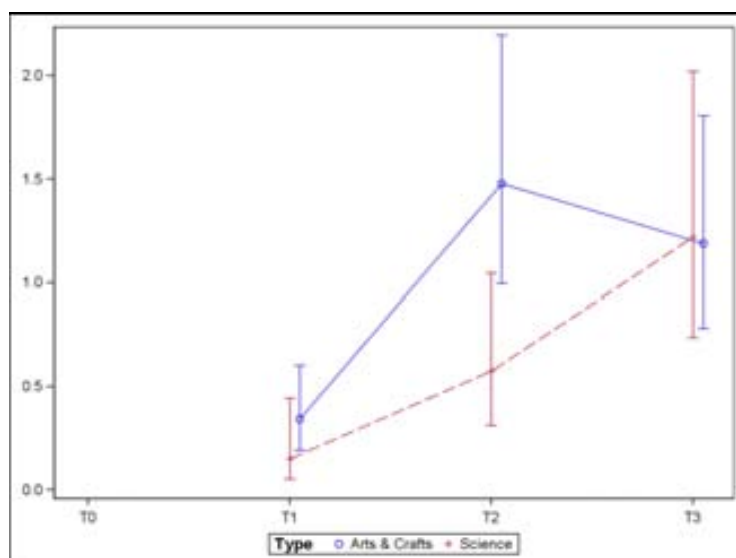


Figure 5.87 Syntactic Complexity achievement results: ISU

As can be seen in Table 5.231 below, there were no statistically significant differences between the High achievers in the Science and in the Arts & crafts groups. Descriptive statistics show that the High achievers in the Arts & Crafts group outperformed the High achievers in the Science group at all the times tested.

	Proficiency	Type	Mean	Std Err	Lower	Upper	F Value	P Value
T1	High	Arts & Crafts	0.47	0.19	0.21	1.05	1.87	0.1725
	High	Science	0.19	0.10	0.07	0.55		
T2	High	Arts & Crafts	1.54	0.48	0.82	2.88	1.07	0.3043
	High	Science	0.99	0.28	0.56	1.74		
T3	High	Arts & Crafts	2.24	0.67	1.23	4.06	0.58	0.4468
	High	Science	1.65	0.42	1.00	2.74		

Table 5.231 Syntactic Complexity achievement results: ISU. Group/ High achievers interaction

Significant differences were found in the number of instances of subordinate units at T1 in favour of the Low achievers in Arts & Crafts ($F_0 8.16$ $p= 0.0048$). Generally, the mean number of subordinate units used by the Low achievers in Science and in Arts & Crafts was lower than the mean number used by their High achievers counterparts.

	Proficiency	Type	Mean	Std Err	Lower	Upper	F Value	P Value
T1	Low	Arts & Crafts	0.25	0.09	0.12	0.53	0.88	0.3500
	Low	Science	0.12	0.08	0.03	0.45		
T2	Low	Arts & Crafts	1.42	0.34	0.89	2.29	8.16	0.0048
	Low	Science	0.33	0.15	0.14	0.81		
T3	Low	Arts & Crafts	0.63	0.19	0.35	1.14	0.54	0.4634
	Low	Science	0.90	0.34	0.43	1.90		

Table 5.232 Syntactic Complexity achievement results: ISU. Group/Low achievers interaction

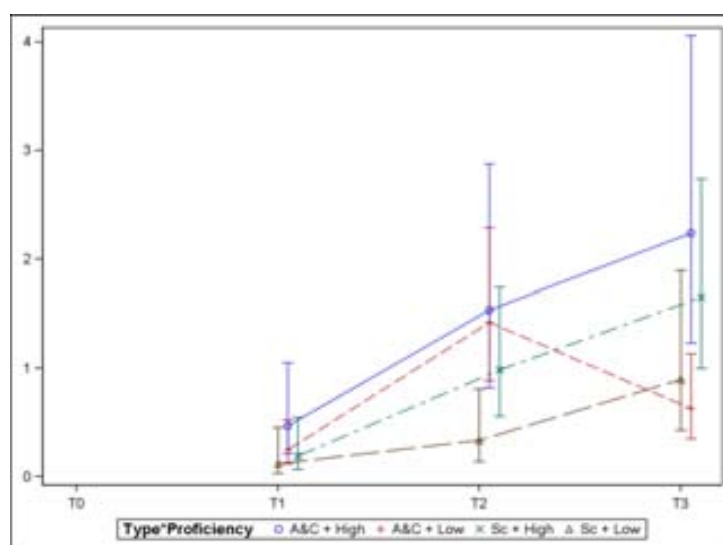


Figure 5.88 Syntactic Complexity achievement results: ISU. Group/Proficiency level interaction

5.3.5 Summary of CLIL Science vs. CLIL Arts & Crafts Achievement

Results

This section presents a summary of the results obtained by CLIL Science students compared to CLIL Arts & Crafts students. Table 5.233 below displays the fluency achievement results of the intergroup comparison between CLIL Science and CLIL Arts & Crafts students.

FLUENCY	TNW		TNWE		%TNWE/TNW		TNU	
	CLIL Science	CLIL Arts & Crafts	CLIL Science	CLIL Arts & Crafts	CLIL Science	CLIL Arts & Crafts	CLIL Science	CLIL Arts & Crafts
T0	62.09	71.69	49.11	61.03	77.0%	84.1%	9.51	9.84

Language Competence of Young Learners Exposed to EFL and CLIL

	p= 0.2523		p= 0.1218		p= 0.0042		p= 0.7993	
T1	64.16	78.95	54.97	70.77	84.6%	88.7%	8.98	10.67
	p= 0.0785		p= 0.0408		p= 0.0095		p= 0.1977	
T2	83.68	105.2	70.17	91.71	82.3%	86.9%	12.09	16.98
	p= 0.0107		p= 0.0081		p= 0.5179		p= 0.0002	
T3	120.4	117.7	107.8	106.7	88.7%	90.7%	17.65	19.46
	p= 0.7512		p= 0.8906		p= 0.7886		p= 0.1703	

Table 5.233 Summary of Fluency achievement results. Science vs. Arts & Crafts

The achievement results of the intergroup comparison CLIL Science vs. CLIL Arts & Crafts students for two of the fluency measures: TNWE and TNU when the interaction Group/Proficiency level was taken into account can be seen in Table 5.234 below.

FLUENCY Group/proficiency	TNWE				TNU			
	HIGH		LOW		HIGH		LOW	
	Science	Arts & Crafts	Science	Arts & Crafts	Science	Arts & Crafts	Science	Arts & Crafts
T0	54.13	74.28	44.09	47.77	10.34	11.92	8.68	7.76
	p= 0.0507		p= 0.7161		p= 0.3667		p= 0.5960	
T1	61.80	79.74	48.68	61.80	9.70	12.36	8.26	8.99
	p= 0.0731		p= 0.1959		p= 0.1304		p= 0.6742	
T2	89.92	102.4	52.43	81.03	15.14	18.30	9.04	15.66
	p= 0.2255		p= 0.0052		p= 0.0723		p= 0.0002	
T3	118.4	121.5	97.17	9202	19.11	22.14	16.19	16.78
	p= 0.7679		p= 0.6157		p= 0.7369		p= 0.7369	

Table 5.234 Summary of Fluency achievement results. Science vs. Arts & Crafts. Group/Proficiency level interaction

Table 5.235 sums up the Accuracy achievement results and Table 5.236 describes the accuracy achievement results as for the TNEFU when the interaction Group/Proficiency level was taken into account.

ACCURACY	TNEFU		%TNEFU/TNU	
	Science	Arts & Crafts	Science	Arts & Crafts
T0	2.09	5.80	23.5%	63.6%

	p= <.0001		p= <.0001	
T1	3.07	5.70	37.0%	55.9%
	p= 0.0017		p= 0.0072	
T2	4.45	11.16	40.2%	67.9%
	p= <.0001		p= <.0001	
T3	7.69	12.74	45.8%	69.2%
	p= 0.0006		p= 0.0001	

Table 5.235 Summary of Accuracy achievement results. Science vs. Arts & Crafts

ACCURACY Group/proficiency	TNEFU			
	HIGH		LOW	
	Science	Arts& Crafts	Science	Arts& Crafts
T0	1.94	7.26	2.26	4.64
	p= <.0001		p= 0.0103	
T1	3.98	6.69	2.57	4.85
	p= 0.0251		p= 0.0069	
T2	6.30	12.22	3.14	10.20
	p= 0.0010		p= <.0001	
T3	9.02	15.17	6.55	10.69
	p= 0.0043		p= 0.0139	

Table 5.236 Summary of Accuracy achievement results TNEFU Science vs. Arts & Crafts.

Group /Proficiency level interaction

The results of the intergroup comparison as for the different measures taken to analyse Lexical Complexity are shown in Table 5.237 below.

LEXICAL COMPLEXITY	TNLV		%TNLV/TNWE		TNAdj		TNAdj/TNWE	
	Science	Arts& Crafts	Science	Arts& Crafts	Science	Arts& Crafts	Science	Arts& Crafts
T0	1.32	1.09	3.0%	1.9%	1.11	2.95	2.3%	5.3%
	p= 0.4123		p= 0.0704		p= <.0001		p= 0.0008	
T1	1.43	2.43	4.9%	3.6%	0.68	1.83	1.3%	2.8%
	p= 0.9957		p= 0.0659		p= 0.0012		p= 0.0138	
T2	2.43	4.20	3.9%	4.8%	1.73	3.37	2.6%	3.8%
	p= 0.0018		p= 0.2036		p= 0.0012		p= 0.0668	
T3	4.07	4.30	4.1%	4.3%	3.37	4.83	3.2%	4.8%
	p= 0.7126		p= 0.8087		p= 0.0264		p= 0.0106	

Table 5.237 Summary of Lexical Complexity achievement results. Science vs. Arts & Crafts.

Table 5.238 shows a summary of the achievement results as for the TNLV and the TNAdj when the interaction Group/proficiency level was taken into account.

LEXICAL COMPLEXITY Group/proficiency	TNLV				TNAdj			
	HIGH		LOW		HIGH		LOW	
	Science	Arts & Crafts	Science	Arts & Crafts	Science	Arts & Crafts	Science	Arts & Crafts
T0	2.12	1.71	0.82	0.70	1.30	3.35	0.95	2.59
	p= 0.4135		p= 0.6234		p= 0.0007		p= 0.0023	
T1	3.32	2.95	1.78	1.99	0.89	1.70	0.52	1.98
	p= 0.5958		p= 0.6476		p= 0.0606		p= 0.0012	
T2	3.41	4.40	1.74	4.01	2.48	3.44	1.21	3.29
	p= 0.2246		p= 0.0005		p= 0.1714		p= 0.0005	
T3	5.00	5.21	3.30	3.55	3.81	5.42	2.98	4.30
	p= 0.8335		p= 0.7336		p= 0.0802		p= 0.1048	

Table 5.238 Summary of Lexical Complexity achievement results: TNLV, TNAdj. Science vs. Arts & Crafts Group/Proficiency level

The achievement results in Syntactic Complexity are summarised in Table 5.239.

SYNTACTIC COMPLEXITY	ICU		ISU	
	Science	Arts & Crafts	Science	Arts & Crafts
T0	1.12	1.11		
	p= 0.9826			
T1	0.86	1.32	0.15	0.34
	p= 0.1592		p= 0.1813	
T2	2.22	2.07	0.57	1.48
	p= 0.7601		p= 0.0108	
T3	2.88	2.27	1.22	1.19
	p= 0.2527		p= 0.9386	

Table 5.239 Summary of Syntactic Complexity achievement results

Table 5.240 shows a summary of the achievement results in Syntactic Complexity for High and Low achievers as for the ICU and ISU.

Language Competence of Young Learners Exposed to EFL and CLIL

SYNTACTIC COMPLEXITY Group/proficiency	ICU				ISU			
	HIGH		LOW		HIGH		LOW	
	Science	Arts & Crafts	Science	Arts & Crafts	Science	Arts & Crafts	Science	Arts & Crafts
T0	0.71	1.73	0.72	0.71				
	p= 0.9887		p= 0.9689		P Value		P Value	
T1	1.46	1.55	0.51	1.13	0.619	0.47	0.12	0.25
	p= 0.8592		p= 0.0674		p= 0.1725		p= 0.3500	
T2	3.00	2.69	1.64	1.60	0.99	1.54	0.33	1.42
	p= 0.6799		p= 0.9397		p= 0.3043		p= 0.0048	
T3	3.01	2.68	2.75	1.92	1.65	2.24	0.90	0.63
	p= 0.6602		p= 0.2134		p= 0.4468		p= 0.4634	

Table 5.240 Summary of Syntactic Complexity achievement results: ICU, ISU. Science vs. Arts & Crafts Group/Proficiency level

CHAPTER 6 Discussion

In the light of the main findings of this study, the research questions posed at the beginning of this thesis are now discussed here. The results of this study centred on the linguistic competence of young primary learners exposed to EFL and EFL+CLIL at school in three different skills (listening, reading and writing), keeping the number of hours of exposure constant.

The organisation of this chapter is as follows: Sections 6.1 and 6.2 discuss the results obtained by CLIL Science students and CLIL Arts & Crafts students compared to the results obtained by their control EFL groups in receptive skills: section 6.1 addresses the listening skill and section 6.2 addresses the reading skill. Both sections also deal with the results in the same skills (listening and reading) of CLIL Science students compared to the results of CLIL Arts & Crafts students. Section 6.3 deals, in the first place, with the results of CLIL Science and CLIL Arts & Crafts students and their control EFL groups in writing skills. It also addresses the results of the comparison in writing between CLIL Science and CLIL Arts & Crafts. In both cases, the discussion focuses on the results in Complexity, Accuracy, and Fluency.

Each section begins with a summary of the results followed by the discussion of the interpretation of those results. Findings in all the sections are discussed and compared with claims made in related fields.

6.1 Achievement and improvement in Listening skills

Hypothesis 1 in the study stated that with the same number of hours of exposure to the target language, CLIL learners in their final two years of primary education exposed to an EFL+CLIL programme would obtain better results in receptive skills, listening and reading, than their counterparts only exposed to an EFL programme.

Based on this hypothesis, several research questions were posed. The questions addressed the issues of achievement and progress in Listening and Reading of CLIL learners compared to their Control groups at different times and over different time periods, as well as the results of the comparison between Science and Arts & Crafts students. The questions were also concerned with the influence of the initial proficiency of the learners on the results.

An overview of the key relevant findings in listening revealed, as had been hypothesized, that in terms of achievement, the CLIL Science group significantly outperformed the Control group at the end of the study, at T3. The difference between the final scores was 9.8%. Although there were non-significant differences in the mean initial scores of the participants, it is interesting to see that, at T2, the mean scores were exactly the same for both groups. The results also revealed no significant differences in the listening comprehension tests between CLIL High achievers and High achievers in the Control Group. However, the results of the Low achievers in CLIL were significantly higher than those of the Low achievers in the Control Group. The lowest mean score was for the Low achievers in the Control group.

Both the CLIL group and its Control Group progressed significantly throughout the two-year period of this study (T0-T3). A close look at the findings in progress showed that the Control Group only improved significantly during the first year. It did not show a significant improvement during the other time periods. As for the CLIL group, no significant improvement was shown during the first year, but the group improved significantly during the final two time periods (T1-T2 and T2-T3).

CLIL High achievers progressed significantly during the final time period, as did High achievers in the Control Group. CLIL Low achievers progressed significantly

during the first and the third time periods, whereas Low achievers in the Control Group only showed significant differences during the first year.

However, contrary to what had been hypothesized, CLIL Arts & Crafts students did not outperform their counterparts in the Control Group at any of the times tested. The percentage results of the Control Group were significantly higher at T2 and T3. The mean percentage difference at the end of the study was 15% higher for the Control Group. High achievers in the Control Group attained significantly better results at all times tested except for T0. The comparison between CLIL and EFL Low achievers did not show significant differences at any of the times tested.

In terms of progress, both groups progressed significantly throughout the study, with the Control Group improving more than the CLIL group. The CLIL group progressed significantly during the first year, but its improvement during the second year was not significantly different.

The comparison CLIL Science vs. CLIL Arts & Crafts revealed that CLIL Science students significantly outperformed their CLIL Arts & Crafts counterparts. Even though the difference between High achievers in Science and in Arts & Crafts was not significantly different, descriptive statistics showed that High achievers in the CLIL Science group obtained better mean scores than High achievers in Arts & Crafts. However, Low achievers in Science significantly outperformed Low achievers in Arts & Crafts.

The results of the statistical analysis suggest that, in the case of Listening, CLIL may have had a more beneficial effect on the students exposed to Science than on the students exposed to Arts & Crafts. The results obtained by CLIL Science students are in line with the results obtained by students in Canadian immersion programmes who developed much higher levels in receptive skills than non-immersion students. This was

also true for the Canadian students who were at a linguistic and academic disadvantage with respect to other students in the same groups (Genesse, 1992; Lightbrown and Spada, 1994). However, these studies were carried out in immersion contexts or in contexts where the students received more hours of exposure to English. The context of the present study was one of a minimal input situation. The results in this study are also very similar to some studies in the European context, which have also found higher levels of Listening skills in countries where CLIL has been applied at school level (Stotz and Meuter, 2003; Dalton-Puffer, 2007; Victori and Vallbona, 2008).

The participants in the present CLIL Science and Arts & Crafts study were young learners who, at the time of data collection, had received the same number of hours of exposure to the language as the Control Group. As has already been explained in Chapter 4 (Method), the CLIL group was part of an EFL+CLIL programme, and the Control Group was part of an EFL programme. The difference between the groups was the weekly CLIL hour that the Science and Arts & Crafts students were exposed to instead of an EFL class. As the number of hours of exposure was the same for both groups, the variability in the results cannot be explained by the amount of exposure, which has been one of the reasons many studies in similar contexts to ours have used to justify the positive results in CLIL (Perez-Cañado, 2012). The variability cannot be explained either by the extra-practice outside the school walls, as the context of the study was one of a minimal-input situation, as explained in Chapter 3. The interpretation, then, should be centred on the role of comprehensible input and the type of instructional practice in the classrooms, as well as on the important role of the discourse in the Science and Arts & Crafts lessons.

As was pointed out in Chapter 3, Dekeyser (2007) claimed that not only the amount of practice but also the type of practice is crucial to language learning. Exposure

to abundant comprehensible input plays a fundamental part in language acquisition (Krashen, 1995). As Muñoz (2006:18) stated, ‘Input is the necessary catalyst through which language is processed and which results in changes in the learners’ linguistic system’. According to Muñoz, apart from comprehensibility and quantity, input also needs to fulfil other conditions: it has to be authentic, contextualized and used in real communicative situations. The CLIL classes in our study provided the ideal setting for these conditions to be fulfilled: the students were dealing with authentic contextualized curricular topics in either Science or Arts & Crafts, which had been chosen by the CLIL teachers with the agreement of specialists in each content area. Class observations carried out during the study by the researcher, as explained in Chapter 4, as well as informal exchanges with the teachers, plus analysis of the materials used, allowed us to see and evaluate the type of instruction carried out by the relevant CLIL teachers, as well as the way they introduced the different skills in their lessons. In order to ensure comprehension of subject matter, both the CLIL Science and the CLIL Arts and Crafts’ teachers, provided plenty of comprehensible input through their own speech. Listening was mostly practised through teacher talk, and the teachers’ speech was modified by speaking slowly, emphasizing key words, using restricted vocabulary, cognates, shorter phrases and plenty of repetition. The positive results of the CLIL Science group during the second year, and indeed throughout the study, which were higher than the positive results of its Control Group, may be explained by the constant exposure to contextualized, challenging and authentic linguistic input during the CLIL Science classes, which added to the skills practice in listening of the EFL sessions. However, in the case of the Arts & Crafts students, even though listening was also practised mainly through teacher talk in a very similar way to the Science classes, this did not have the same effect, as the Control group outperformed the CLIL group in terms of

achievement. As Lyster (2007) points out, even though comprehensible input lies at the heart of academic success in content-based approaches, ‘the continuous use of strategies that rely on visual and non-linguistic support may have negative effects on the students’ communicative ability’ (p.61), in the sense that the students learn to engage in comprehension strategies and ‘draw on schematic and contextual knowledge, thus partly avoiding facing the linguistic demands on the learners’ interlanguage that are fundamental to language acquisition’ (Skehan, 1998: 26). Instructional practice in Arts & Crafts may have relied too much on the visual cues and the gestural support provided by the teachers and among the students themselves, as the sessions were, in most cases, hands-on practical sessions that had as their main objective the production of artistic elements through drawing, painting, colouring or manipulating different materials. Although the sessions started with a brief explanation of the topic by the teacher, the second part of the classes was devoted to creating artistic elements, drawing or painting.

The Science lessons were also delivered using visual support. However, the challenging nature of the academic discourse in the Science classes, the activities associated with the Science lessons, and the linguistic support provided by the teacher to help students respond adequately to the task demands may have been more challenging for young Science learners thereby triggering the necessary prerequisites for language acquisition. As pointed out in section 2.3.1, learning academic language is very challenging, especially for those students who are acquiring English as a foreign language (Bailey and Butler, 2007). Following ideas suggested by Swain (1988) and Harley (1994), incidental attention drawn to language, without focusing the students’ attention on more salient features of the language, may have provided contextualized comprehensible input which, in the case of Arts & Crafts, may have been insufficient to

maximize second language learning, especially with young learners, whose limited concentration span requires good scaffolding in the instructional practice.

The different approaches to scaffolding that teachers provided for their students may be another reason for the success of CLIL Science students in listening, as opposed to the poor results in this skill obtained by Arts & Crafts learners. As Bruner (1978) suggested, and as is explained in Chapter 3, the scaffolding process turns out to be a useful tool that facilitates language learning and helps learners to make sense of new language forms in context, especially with young learners. The learners in the Science group actively participated in the classes due to the interaction created by the Science CLIL teacher, in the form of questions and answers, and group and individual activities, in order to help them in the construction of their own knowledge. This interaction was frequently preceded by a general talk by the teacher introducing the topic, and followed by the introduction of language structures which were relevant to the topic and that allowed the students to participate in the exchanges. The CLIL Arts & Crafts sessions also started with a general introduction to the topic but, as has already been pointed out earlier in this discussion, as the objective of the lesson was the production of artistic objects and pictures, the nature of the teachers' discourse was much more instructional, so that the students could start working on their own projects. The scaffolding process, especially in terms of language features, was much less salient in the Arts & Crafts sessions and less explicit than it was in the Science classes.

The lack of appropriate scaffolding may also provide an explanation as to why the Control Group exposed to much more systematic and explicit EFL classes, outperformed the EFL+CLIL Arts & Crafts group during the second year of the study and why, even though both groups progressed significantly, the achievement at T2 and

T3, as well as the progress throughout the study, was much higher for the Control group than it was for the CLIL Arts & Crafts group.

The differences in scaffolding may again provide one of the explanations for the different results obtained by High and Low achievers in the CLIL+EFL groups. The groups were far from homogeneous. Great variability was found among the groups, especially between High and Low achievers in terms of listening. Even though in terms of statistics no differences in achievement were found between CLIL and Control High achievers, descriptively, at the end of the study, the mean percentage was higher for those in CLIL Science. As for Low achievers, the fact that they obtained significant differences at T3 may indicate that the group benefitted not only from the immersion process (Genessee, 1984), and the contextualized input they received, but also from the highly-structured lessons in terms of scaffolding that the Science teacher provided. This helped them to focus on salient aspects of the language, needed to understand what was happening in the Science class. The highly structured lessons helped the students to develop their cognitive abilities more than would normally have been the case. The fact that they had to concentrate and make an important cognitive effort to understand the contents explained by the teacher may have contributed to the results obtained by the Low achievers in CLIL. As Halbach (2009) points out, understanding new concepts that are presented in a second language represents a great challenge for young learners, who necessarily increase their cognitive effort to face such complex input.

In terms of the Arts & Crafts group, their results displayed differences in favour of High achievers in the Control Group and no differences for the Low achievers. Descriptive statistics showed that the group which benefitted the least at the end of the two-year study was, in fact, the Low achievers in the CLIL+ EFL group in Arts and Crafts. Even though they may not have fully understood the teachers' explanations, as

the classes were quite practical, they could see from their peers' performance what the required work was. Thus, in this way, they may not have developed the necessary listening strategies or had to make the cognitive effort that challenging input requires in order to fully understand what is being said.

As was explained in Chapter 3, according to Schmidt's Noticing Hypothesis (1990), a crucial step in SLA is that learners in general notice target features in the input so that this input becomes intake in terms of triggering interlanguage development. For Skehan (1998), noticing plays a key role in input processing. What students manage to notice in the input depends on many factors, including prior knowledge, task demands, frequency and saliency of the features of the target language, as well as the input itself. As Swain (1996) pointed out, in order to maximize second language learning, content teachers need to manipulate and complement their teaching to draw the students' attention to those aspects that may contribute to language development and which, without the help of the teacher, they would not notice by themselves. Following Vigotsky's (1978) ideas, children can do and understand much more with the help of adults than by themselves. Therefore, in a classroom setting, the teacher is responsible for structuring interaction and developing instruction in small steps based on tasks. The instructor is also charged with providing support until the learner can move through all tasks independently (Cameron, 2001).

Although achievement results of High and Low achievers in listening were not as expected, they progressed significantly throughout the study. Interestingly, Low achievers in the Science group progressed significantly during two of the periods tested (T0-T1 and T0-T3), whereas Science High achievers only progressed significantly during the second year of the study. No significant progress was found during the second year for High and Low achievers in the Arts & Crafts group. The variability in

the results, especially in progress, may also be explained by the differences in the maturational constraints of the learners. Progressing through a foreign language is a complex and non-linear process, with setbacks and multiple regression points (Mitchell, 2003). Young learners, who are still in the process of developing emotionally and cognitively, need time to adapt to new situations. CLIL was for them, and for the teachers, an innovative experimental situation which required a certain adaptation period. As Lyster (2007) points out, processing language through content is a difficult process that requires a certain degree of maturity, even though teachers are helping students to do what they are not yet able to do by themselves.

6.2 Achievement and improvement in Reading skills

Reading was another of the skills tested. Hypothesis 1 and sub hypothesis 1.1 predicted better results for CLIL students in reading. Contrary to what had been hypothesized, the key relevant findings in Reading reveal that, in terms of achievement, neither the CLIL Science group nor the CLIL Arts & Crafts group significantly outperformed their Control Groups at any of the times tested during the study. In fact, in the case of Arts & Crafts, the Control Group attained significantly higher results than the ones obtained by the CLIL group at two of the times tested (T1 and T2). Even though the results were not as expected in terms of achievement, the CLIL groups progressed significantly throughout the study. Nevertheless, the progress was higher for the Control Groups.

There were no differences in the interaction Group/Proficiency level for the CLIL Science students: CLIL High and Low achievers attained lower mean scores than their counterparts in their Control Groups. In the case of Arts & Crafts students, there were statistically significant differences in favour of the High achievers in the Control Group at three of the times tested (T1, T2 and T3). Low achievers in the Control Group

outperformed the CLIL students at T2, while at T3 the mean percentage in reading of the Low achievers in CLIL was higher than that of the Control Group.

The comparison Science vs. Arts & Crafts revealed that CLIL Arts & Crafts students obtained better mean scores and that High achievers seemed to benefit more from reading (significantly better results at T2 and better mean scores at T3). Low achievers obtained higher scores at the end of the study at T3.

Several reasons may be behind the lack of positive results of the CLIL groups as far as reading is concerned. In terms of receptive skills, less time was devoted to reading in the CLIL class. As has already been pointed out, class observations in the Science group revealed that most of the lessons were very much teacher-centred, in the sense that classes revolved around the teachers' oral explanations, and the interaction teacher-student, rather than being based on reading texts. This may have benefited the students' listening skills but not their reading ones. Reading was introduced throughout the study in the form of simplified texts prepared by the teacher herself, which were later used by the students as a resource for revising the contents of the lesson and answering the questions posed in the activities related to each topic. Some of the texts may have been oversimplified in order to make otherwise challenging scientific content comprehensible for the limited proficiency of the primary students. As the teacher explained, she found it very difficult to maintain the balance between simplification and comprehensibility.

In the case of Arts & Crafts, there was a difference in the approach to reading between the first and the second years of the study. During the first year, as in the case of the Science teacher, the lessons revolved very much around the oral explanations provided by the teachers themselves before they engaged the students in practical activities, which did not involve very much reading. During the second year, however, the teachers' decisions to make their CLIL sessions more content-based may have

affected the outcome at the end of the study: several descriptive art projects were carried out by the students themselves. The projects involved reading and finding information about painters or artistic movements. The element of reading became much more present in the classes as the students brought real texts that they had found themselves or with the help of the teacher. This change in approach may have had an influence on the results obtained by the groups and may explain why the first year results in Arts & Crafts had been in favour of the Control Group, whereas the reading results at T3 had been in favour of the CLIL group, although not significantly. It may also provide an explanation as to why the CLIL Arts & Crafts students obtained better mean scores in reading, although not significantly, when they were compared with the CLIL Science students.

As previously mentioned in the discussion about the listening results, comprehensible input (oral and written) is a necessary condition in language acquisition in order to enhance receptive skills. Teachers in content-based situations rely on transforming subject matter into comprehensible input (Snow, 1987; Met, 1994; Tardiff, 1994). The conventional wisdom applied to many reading texts is that simplification is desirable for comprehension. Several studies have proved the efficiency of text simplification as a form of enhancing comprehension (Kim, 1985; Lee, 1986). However, other studies have pointed out that simplification may not necessarily contribute to language acquisition, as the use of simplified syntax and reduced vocabulary is likely to result in 'unnatural' English (Blau, 1982). It may also prevent the students from developing adequate reading strategies that would not be suitable for normal, unsimplified materials (Yano et al., 1994). The process of simplification may also leave the relationship between essential bits of information in the text unclear,

especially in tasks that require understanding and relating the different parts (Honeyfield, 1977) to each other.

The teachers in the study created their own materials and texts, pushed to do so by the lack of materials that would adequately suit the needs in their classes (Halbach, 2011). As has already been said in this dissertation, the teachers chose their own topics and prepared their own texts and activities for the topics. Finding texts did not prove to be difficult for them, but adapting the texts to making them comprehensible, while maintaining some of the original features in order not to lose the essence of the scientific discourse, proved to be really challenging, as they explained in a questionnaire given out at the end of each year (Pladevall-Ballester, forthcoming). As explained in Chapter 3, there seems to be a general consensus on the role of the teacher as one of the influential aspects to be taken into account when analysing the results obtained by the students, because this factor affects not only the model of language provided, but also the type of methodology adopted. As several studies have pointed out, teachers' beliefs, course materials, language practices and their training influence their daily classroom practices (Rixon, 2000; Coonan, 2007; Sohamy and Inbar, 2007; Lundberg, 2007; Matteoudakis et al., 2007; Moon, 2009;). Creating materials and particularly adapting texts are not easy tasks, requiring not only linguistic knowledge but also textual, content and pedagogical knowledge in order to ensure that the original main features of the texts are kept in place.

The results of each of the individual reading questions (see Chapter 4 section 4.3.2.1.2 for a detailed description of each reading test) may be connected to the reasons mentioned above. None of the four reading questions showed significant results in favour of the CLIL Science group and, in fact, the Control Group, exposed to much more systematic reading in the EFL lessons, following the guidelines of a course book,

obtained better mean percentages than the CLIL group in terms of reading. The Arts & Crafts group displayed very similar results: at the end of the first and the second years.

The results in Question 1, where students were asked to match words with their definitions, were significantly different in favour of the Control Group. However, the results for questions 2 and 4 (Comprehension), whose main aim was reading for specific information, were significantly different at the end of the first year in favour of the Control Group, but not at the end of the second year. Even though the mean percentages reflected a certain advantage for the Control Group, at this time the CLIL Arts & Crafts students seemed to catch up with the Control Group. The results in these two questions may reflect the change in methodology explained above, as well as the importance that reading received through the second year projects which may have enhanced the reading ability of the students in Arts & Crafts.

Question 3 (Cloze test) and question 4 (Accuracy), which aimed at assessing the students' use of English and involved an element of grammatical accuracy, did not show significant differences between the CLIL Science group and the Control Group. In the case of Arts & Crafts, the Control Group outperformed the CLIL group in the first year, but at the end of the study, at T3, the CLIL group seemed to catch up with the Control Group. The results of these two questions are in line with some of the claims from the field of SLA, which state that learning grammar is not a lineal process, but rather a process with multiple regression points, and young learners do not learn patterns straightaway as they are not cognitively mature enough to explicitly understand the grammatical structure of the language (Peltzer-Karf and Zangl, 1997; Mitchell, 2003). The different developmental stages that young learners go through may be attributed to individual factors (age, maturational constraints, cognitive development, their L1) as well as to contextual and methodological aspects (Pinter, 2006). As Ranta

(2002) and Skehan (1998) pointed out, young learners in content-based classes benefit from age-appropriate noticing and awareness activities that help them to drive their interlanguage development forward. In the words of Pinter (2006: 84), 'It is better if grammar is noticed and learnt from meaning-focused input'. Learning grammar on the part of young learners is a messy process that requires lots of practice, recycling and guidance in order to make learners attend to language form. As the time devoted to CLIL was limited to one hour a week, there was very limited opportunity to focus on grammatical aspects, even though the children were provided, through scaffolding, with the necessary chunks of language that would allow them to participate in the interaction. Children may have drawn and transferred grammatical knowledge and skill from their EFL classes, and this would explain why, in spite of the poor results in questions 3 and 4 (Accuracy) on the part of the CLIL groups in the early stages of the study, especially the Arts & Crafts students, learners still managed to catch up with the Control group and get similar percentages at the end of the study. The Control Groups in this study were exposed to much more specific language-oriented classes for a minimum of three hours a week.

In the light of the findings presented earlier in this section and in section 6.1, the idea in the first hypothesis in the study, concerning greater results in receptive skills, listening and reading for the students in the EFL+CLIL Groups, is only partially confirmed. Significant results were attained by the CLIL Science students in listening but no significant results were found for any of the groups in reading.

6.3 Achievement and improvement in Writing: Complexity, Accuracy and Fluency

Three different hypotheses on the results in writing were raised at the beginning of this project. In the first place, Hypothesis 1.2 in this study stated that keeping the

number of hours of exposure constant, when contrasting the differential effects of two different programmes, an EFL programme and an EFL+CLIL programme on the learners' linguistic achievement and progress over two academic years, the students exposed to an EFL+CLIL programme would obtain better results in written fluency and complexity than the students exposed to an EFL programme. Secondly, Hypothesis 1.3 stated that students exposed to an EFL+CLIL programme would not obtain better results in accuracy than the students exposed to an EFL programme. Finally, Hypothesis 2 posed that students in an EFL+ CLIL programme exposed to Science in English would obtain better results in writing than students in an EFL+CLIL programme exposed to Arts & Crafts in English.

In the light of these hypotheses, the questions posed were concerned with the achievement and progress of the students' writing, in English in Science and in Arts & Crafts at different times and different time periods when compared to their Control Groups. They were also concerned with the achievement results of the comparison between CLIL Science students versus CLIL Arts & Crafts students. The questions also focused on the influence of the initial proficiency level of the students on the results. As in the two previous sections, each sub section begins with a summary of the results for each of the groups, followed by a discussion of the interpretation of the results.

6.3.1 Achievement and improvement results in Fluency

In terms of Fluency, Hypothesis 1.2 was only partially confirmed in this study as not all the students exposed to an EFL+CLIL programme obtained better results in Fluency than those obtained by the students in the EFL programme. The quantitative results obtained for Fluency by the students in the comparison EFL vs. EFL+CLIL Science indicated that the Control Group significantly outperformed CLIL Science students at the end of the first year. However, the results at the end of the second year

were not statistically significant for any of the groups. Descriptively, though, the CLIL group obtained better mean percentages in TNWE and TNU. The results in these two measures, when the interaction Group/Proficiency level was taken into account showed an advantage for High and Low achievers in the Control Group at T1. However, at the end of the study (T3), the mean percentages were always higher for the students in the CLIL group, although the results were not significantly different. In relation to the progress of the groups, even though both groups, CLIL and Control, improved significantly throughout the study, the CLIL group only progressed significantly during the second year. Both High and Low achievers in CLIL significantly outperformed their Control groups in the final time period (T2-T3).

The achievement results of the CLIL Arts & Crafts group, compared to the Control group in terms of Fluency, showed no significant differences at T1. Nevertheless, the results at T3 indicated that CLIL Arts & Crafts students significantly outperformed the Control Group in two of the measures: TNWE and TNU. The interaction Group/Proficiency level revealed significant differences for High and Low achievers in CLIL during the second year, T2 and T3, as for TNU. In relation to progress, as in the case of the CLIL Science group, the CLIL Arts & Crafts group started to progress significantly during the second year, but, in the case of students in Arts & Crafts, the overall improvement throughout the study was significant. The analysis of the improvement of High and Low achievers revealed that both of them progressed significantly during the second year of the study.

Hypothesis 2 was not confirmed in this study, as the Science students did not significantly outperform the students in Arts & Crafts in terms of Fluency. The results of the comparison between CLIL Science vs. CLIL Arts & Crafts learners showed a slight advantage in all the measures for the students in Arts & Crafts at T1. At T2, the

differences were significant for the students in Arts & Crafts in TNW, TNWE and TNU. At T3, however, the CLIL Science group seemed to catch up with the CLIL Arts & Craft group in TNWE but not in TNU, in which the Arts students still obtained slightly better mean percentages. No differences were found between High and Low achievers in Science and Arts & Crafts during the first year, although descriptive statistics showed a slight advantage for the Arts group. Low achievers in Arts & Crafts significantly outperformed their counterparts in the Control Group at T2. However, at T3, the Low achievers in Science obtained similar or slightly better results in the Fluency measures.

Young learners' fluency may benefit not only from input-based instruction but also from the range of instructional activities designed to negotiate and promote the use of the target language in the classes. Even though producing language constitutes a fundamental part of language acquisition, as stated by Swain's Output Hypothesis (1995) (see Chapter 3), it may also be necessary to link this theory with the Skills Acquisition theory (Anderson, 1982), which views practice as a means of developing fluency. The variability in the results in writing obtained by the learners in the Science and the Arts & Crafts groups may be explained by the way in which the skills in general, and writing in particular, were dealt with in the sessions.

In order to get young learners to write, it is essential for instructors to 'push' the students to produce language. Pushing the students helps them notice gaps in their interlanguage and try out new forms. Being pushed to produce language and thereby forced to automate as well as incorporate new forms into their interlanguage may contribute to the development of fluency. As Curtain (2000) stated, the amount of time spent in language practice in a young learners' language programme is normally a good indicator of performance.

As in the case of reading, writing was not the most practised skill in the CLIL Science classes. The lack of continuous practice may be considered one of the reasons for the slightly poorer results obtained by the students in the Science group in comparison to the Arts & Craft group. Through class observations, it became clear that writing in the Science class was, most of the time, performed either by copying texts from the teachers' powerpoint³¹ presentation or writing answers to the questions posed by the instructor, which had already been answered orally in class and written on the blackboard. Most of the writing activities in the Science class were not very challenging controlled activities and they did not involve the creation of free texts. The lack of free creative writings was also true for the Arts & Crafts students. Nevertheless, most Arts & Crafts topics finished with a written worksheet where the pupils were asked to write the answers to questions related to the contents studied. The questions ranged from providing short explanations on important concepts of the lesson to expressing the personal opinion of the learners on the topic and the work done. The worksheets were systematically collected and corrected by the teacher, who provided feedback to the students on their work. This consistent use of writing may have also had an influence on the results in writing of the students in Arts classes, as it may have provided more opportunities for them to practise, thus allowing them to develop certain aspects of fluency. The writing skills practised in the EFL sessions of the students in Arts & Crafts mainly in the form of short texts may have also contributed to the final outcome in Fluency. The students were at all times encouraged to write short compositions in order to practise the different language aspects dealt with in the classes. The pupils may have

³¹ The teacher in the Science class created her own materials and adapted the texts herself, but the students were not provided with handouts. They were encouraged to copy texts from the powerpoint presentation.

also transferred the skills from the EFL sessions in order to write the texts in the CLIL Science class.

As has been pointed out in the reading discussion, there was a methodological shift in the second year of CLIL implementation in the Arts & Crafts groups towards a more content-based approach to the detriment of their practical approach to the subject. The students in Arts & Crafts carried out small projects on artists and artistic movements that may have pushed them into reading more than they had done in their first year. They were also pushed into doing more writing, in order to present their written projects to the teacher. In these projects, the teachers provided general guidelines in the form of general guidelines so that the students looked for the right type of information. Then, the pupils themselves had to write their own short assignments which were then presented orally to the class. This type of practice may have helped students recognize the communicative value of the language, and motivate them to acquire it, therefore contributing to the development of their written fluency. The gains in Fluency on the part of Arts & Crafts students during the second year may have also been helped by a combination of practice in two different skills, reading and writing. Several authors have already mentioned the beneficial effects of reading on the development of writing, especially in young language learners' programmes (Elly, 1991; Drew, 2010). Time devoted to this skill, which has also been considered a requisite for language development in SLA, together with skills practice, may have also played an influential role in the development of Fluency. Although it has been suggested that practice does not make perfect (Lightbrown, 2000), in the words of Muñoz, 'it does make better' (Muñoz, 2007: 229). Learners need plenty of time and practice to develop not only oral skills but also their reading and writing skills (Lightbrown and Spada, 2006). The amount of time devoted to the practice of the two

skills, combined with the learners' capacity to transfer skills and knowledge from their EFL sessions may have all contributed to the positive results in Fluency obtained by the students in the CLIL Arts & Crafts group. As Ellis suggested (2003), production may be considered the mechanism that connects the learners' memory-based system to the rule-based system, which helps them in the construction of their interlanguage. The fact that students had to find, read and write short projects to be presented to the rest of the students may have pushed them to use and structure the language in meaningful units to try to make themselves understood by their peers in the class. The significant results attained by High and, especially Low achievers in Arts & Crafts in TNU may confirm this finding. Low achievers, who find it difficult to cope with the form focused language requirements in the EFL class, may have been encouraged by the need to communicate meaningful content to their peers. The need to communicate may have prompted them to write more, therefore improving the way that they structured their written language for communicative purposes.

Several authors have supported the assumption that communicative language ability is acquired through purposeful communication (Genessee, 1987; Lyster, 2007), and that language learning may proceed in response to the communication demands of academic work, given certain motivational conditions (Cummins, 1981). The fact that students in Arts & Crafts significantly wrote more words in English than their Control Group may have also been influenced by the continuous use of the dictionary to check or to find out the meaning of unknown words. Dictionary work was a consistent practice in the Arts & Crafts sessions, especially during the second year of CLIL implementation, and may have pushed the students into discovering the need to value the communicative importance of writing as much as possible in English, rather than using words in their own language, which was a characteristic of the writings in the first

stages of the study, almost certainly influenced by the age of the students. It is worth mentioning here that Arts & Crafts students significantly outperformed their Control group at T3 in TNWE.

Although not significantly different, CLIL Science students also outperformed their Control Group at the end of the study: they wrote more words and used more units than their CLIL counterparts. As has already been pointed out above, writing was also worked into the CLIL Science class, although at a smaller scale and in a different way. The amount and type of written practice in the form of very short answers to questions previously worked on class, together with a much more challenging type of language that the Science discourse entails and the fact that the content was more difficult to understand may have contributed to the slightly poorer results of the Science group. Scientific writing may have proved to have been a much more challenging exercise for the limited language proficiency of young learners in the study and may have been a disadvantage in the short term for the students. However, even though the students' results were not significantly different in the two years of the study, their descriptive results were better than those of the Control Group at T3. Therefore, the practice of writing carried out in the CLIL class may have also contributed to the development of the writing skills in the area of Fluency at a smaller scale. However, the students may have been constrained by the challenging nature of the scientific discourse and content, together with the type of written practice which was implemented in the class and which, in a certain way, was not very challenging for the students. The lack of cooperation between the EFL and CLIL teachers may have also had an effect on the final outcome in Fluency. There was very little cooperation between the CLIL and EFL teachers in the case of the programme EFL+CLIL Science and this may have contributed to the lack of skills transfer from EFL to CLIL.

The teacher factor may, therefore, have had an effect on the outcome of the project. As has already been explained in Chapter 4 (Methodology), in the case of Arts & Crafts, two different schools and therefore, two different teachers, were in charge of implementing CLIL. The CLIL and EFL teachers in one of the schools worked closely together in planning the CLIL Arts lessons, and the EFL teacher provided explicit linguistic support to the CLIL teacher. In the case of the other school, the CLIL teacher was able to provide the necessary language support in preparation for the CLIL sessions, as she herself was the EFL teacher for the CLIL and Control Groups. The fact that the teachers worked together (or the fact that there was only one teacher) may have also contributed to the development and outcome of the written language in this study. The CLIL teacher in Science mentioned in one of the questionnaires provided at the end of the second year that she felt that the EFL teachers in her school did not provide enough language support for the CLIL classes. She also said that although from the language perspective, she worked hard on the scaffolding process during the Science lessons, the amount of time she could use in the CLIL class to provide the necessary language support to the students, was limited and probably not sufficient. This lack of joint effort between the EFL and CLIL teachers may have also had an effect on the final outcome of the study. Cooperation and support between CLIL and language teachers has been considered to be, among many others, one of the key factors in the successful implementation of CLIL programmes (see Chapter 2 De Graaff, 2007; Navés, 2011).

6.3.2 Achievement and improvement results in Accuracy

In terms of accuracy, in line with what had been hypothesized (hypothesis 1.3), the students in the CLIL Science group did not significantly outperform their counterparts in the Control group. Even though at T3 descriptive statistics showed a slight advantage for the CLIL group in terms of the TNEFU, the ratio TNEFU/TNU was

slightly higher for the Control Group. A close look at the results of the differences between High and Low achievers showed no significant differences for any of the groups. Despite the lack of significant differences, High and Low achievers in CLIL obtained better mean scores at the end of the study.

The Science CLIL group progressed significantly in one of the measures (TNEFU) during the second year. No significant differences were found in the progress of the Control Group. The analysis of the differences between High and Low achievers indicated significant differences for Low achievers in CLIL during the second year of the study.

Contrary to what had been hypothesized (Hypothesis 1.3), CLIL students in Arts & Crafts obtained better results in Accuracy than their peers in the Control Group. At the end of the study, the achievement results were significant in all the measures used to describe accuracy. The differences between High and Low achievers were also statistically significant in favour of the High and Low achievers in CLIL Arts and Crafts, especially during the second year.

Still contrary to what had been hypothesized (Hypothesis 2), the CLIL learners in Arts & Crafts also significantly outperformed their peers in the CLIL Science Group in all the measures used to assess accuracy. High and Low achievers in Arts and Crafts attained significantly better results than their counterparts in the Science Group.

The unexpected achievement results in accuracy from the CLIL Arts & Crafts students are not in line with the results of Canadian programmes (Lambert and Tucker, 1972; Canale and Swain, 1980; Cummins and Swain, 1986; Genesee, 1987, Harley et al., 1990) although the contexts are not easily comparable because of the difference in the number of hours of exposure. However, very few European CLIL studies, which are, in any case, closer in context to the situation in our study as opposed to the situation

in Canada, have found significant results in this aspect either. They are in line with Dalton Puffers' prediction (2007) on the lack of positive gains in accuracy on the part of CLIL students.

The surprising accuracy results attained by the students in Arts & Crafts in this project may be justified by the particular instructional practices carried out by the CLIL teachers in combination with the outcomes of form-focused instruction which was mainly practised in the EFL classes. Science students did not significantly outperform their Control Group in the first year, but their results at the end of the study also showed an advantageous trend for the CLIL Group. Through class observations, it became obvious that the CLIL teachers in general, but especially those in Arts & Crafts, frequently provided explicit corrective feedback to the students during the lessons. This type of feedback may have also helped the students to focus on aspects of form and to notice errors in their interlanguage which, otherwise, would not have been noticed by young language learners in a content-based class. CLIL Science students also showed an advantage at T3. However, in this particular case, the results in accuracy may have been affected slightly negatively due to the difficulty of the scientific contents, which made the students concentrate more on the material than the language it was presented in.

As has been pointed out in Chapter 3, several authors (Lyster, 2007; Muñoz 2006) have already suggested that a counterbalanced approach of content-based activities and form-focused instructional practices provide the learners with opportunities to process language and to progress in their language development. Learners who do not receive form-focused instruction fail to develop accuracy. The impact of instructional intervention is greater when it is explicit in focusing attention on particular language forms (Lightbrown and Spada, 1999; Doughty, 2001; Muñoz,

2006). The grammatical skills proceduralized in the FL classroom may have been transferred and automatised in the CLIL settings and, combined with content-focused lessons together with the language support given by the teachers in CLIL; they may have had an effect on the results obtained in accuracy. The significant results attained by High, but especially Low achievers in Arts & Crafts may prove the effectiveness of this combination of content and linguistic form.

The teacher and the instructional practices carried out in the EFL classes may have also contributed to the results in Accuracy. As has already been explained above, in the case of Arts & Crafts, EFL and CLIL were represented by the same teacher (or by two teachers in constant cooperation). Therefore, young learners may not have seen any difference in a practical sense between EFL and CLIL from the language point of view and therefore they may have applied the criteria for accuracy in the CLIL class that the same teachers required in the EFL class. On the other hand, the Science students had two different teachers, one for EFL and one for CLIL and these teachers did not work together. Therefore, it may have been more difficult for young learners to satisfy the demands of the teacher in CLIL in terms of accuracy because they did not easily see the connection in terms of language between the English in CLIL and the English in EFL. The interaction between student and teacher is a tool for learning as it contributes to building up joint understanding between them (Alexander, 2003).

6.3.3 Achievement and improvement results in Complexity

As explained in Chapter 4, Complexity was considered from two different perspectives: Lexical and Syntactic Complexity. Lexical Complexity was analysed using four different measures: TNLV, %TNLV/TNWE, TNAdj and % TNAdj/TNWE. Syntactic Complexity was assessed using two measures: ICU and ISU.

Hypothesis 1.2, which predicted better results in Complexity for the students in the EFL+CLIL programme, was only partially confirmed. In terms of Achievement, the CLIL Science group hardly achieved significant results in the first year in Lexical Complexity (only in % TNLV/TNWE) and no significant results were found in the second year of the study either, except in TNAdj. In this measure, the CLIL group, in general, and the High achievers in the group, in particular, significantly outperformed the Control group. The CLIL Science group progressed significantly throughout the study: although no significant improvement was displayed during the first year, they significantly improved during the second year. As for Syntactic Complexity, no differences were found between the CLIL and the Control Groups at the end of the study. However, the scores were slightly better for the CLIL students at T3.

Nevertheless, the results of CLIL Arts & Crafts students were much more in line with Hypotheses 1.2. The students in general, and High achievers in particular, significantly outperformed their peers in the Control Group in the first and second years in Lexical Complexity (TNAdj, %TNAdj/TNWE). Students in Arts & Crafts progressed significantly only during the first year. However, the progress throughout the study was significantly different. No differences were found in Syntactic Complexity measured as ICU and ISU. However, descriptive percentages for ICU indicated a slight advantage for the CLIL group after one and two years of CLIL implementation. In terms of subordination, at T1 and T3, the advantage was for the Control Group. When the interaction High and Low achievers was analysed, it was clear that only High achievers progressed in terms of coordination and subordination. CLIL Low achievers did not significantly progress in this area.

Although Hypothesis 2 predicted an advantage for CLIL Science students over CLIL Arts students, the comparison Science vs. Arts and Crafts revealed, contrary to

what had been hypothesized, that in Lexical Complexity the results in TNAdj were significantly different and in favour of the Arts & Crafts students. In the case of Lexical verbs, the mean scores were very similar, even though the results were significantly different at T2. Low achievers in Arts & Crafts seemed to benefit more from the CLIL approach: they obtained better scores in TNAdj and TNLV, despite the scores at T3 not being significantly different.

No significant differences were found in Syntactic Complexity as for the ICU. CLIL High and Low achievers in Science seemed to obtain higher mean scores at the end of the project. As for subordination, CLIL Arts & Crafts students performed significantly better at T2. This did not replicate at T3, however, where the scores of the Science and Arts & Crafts students were almost the same. Low achievers in Arts & Crafts outperformed Low achievers in Science at T2; however, the results were slightly favourable for the Control Group at T3.

Several studies carried out in the European context have shed light on the gains in lexical complexity among CLIL students (Jiménez-Catalán et al., 2006; Seregély, 2008; Dalton-Puffer, 2010; Navés, 2011). It is natural for young language learners to acquire more nouns than verbs and adjectives, especially in the early stages of language learning. Acquiring lexical verbs and using them in adequate contexts is a challenging process for young students. As shown in Chapter 6 (Results), no major differences were observed in the use of lexical verbs between the CLIL and Control groups and between High and Low achievers. Arts & Crafts students performed better than their CLIL counterparts at T1 and T2, but not at the end of the study, at T3. A close look at the types of verbs used revealed that students in the Control and the CLIL groups used very similar types of verb categories. The lack of differences and the similarity of the verbs used by the CLIL and the Control groups may also be explained by the nature of the

task (a composition titled *My Life*) which did not directly elicit the inclusion of a variety of verb types and which may also have pushed the students into use high-coverage verbs at the expense of other types of verbs which might have been a reflection of those acquired in the CLIL lessons by the different groups.

The positive results in the number of adjectives obtained by the students in Science, and especially by the students in Arts & Crafts, may have been directly related to the descriptive nature of most of the activities carried out during the CLIL lessons, which encouraged and engaged students in the use of adjectives. A close look at the mean scores in terms of adjectives used by CLIL students revealed that the students in Science trebled the mean score attained by their counterparts at T3. In the case of students in Arts & Crafts, their mean score was almost four times higher than the score of the Control Group.

As for Syntactic Complexity, despite the fact that the groups progressed significantly throughout the study, syntactic complexity measured as ICU and ISU displayed no significant differences between the CLIL and the Control Groups. No instances of subordination were found during the first year of the study. It is worth considering that the informants were very young learners, whose first language was still developing and their language proficiency in the target language was very limited. However, in spite of the lack of significantly different results, the writings at T3 were sophisticated and syntactically more complex than the writings at T1, which indicated the students' progress throughout the study. Examples of coordinated units were found at all times during the study, whereas subordinate units only appeared from T2. Coordination was only introduced mainly through *and* or *but* throughout the study. At T3, units containing subordinate nominal units of the type *I think that*, and adverbial subordinate units of reason using 'because' were mainly found in both, Science and

Arts & Crafts students' texts which probably mirrored some of the thinking skills developed in the CLIL lessons. As could be appreciated in the class observations carried out, reasoning and forming opinions were two of the most frequently used thinking skills in the CLIL classes. These findings are in line with the findings of studies carried out with young secondary learners: Tjerandsen (1995) found that the writings of young Norwegian learners' writing included a higher occurrence of nominal clauses and adverbial clauses of reason and purpose. Drew (2003) pointed out the majority of the students in his study wrote understandably and coherently in English but their use of lexical and syntactic complexity was much inferior to their level in Norwegian. The results in Syntactic complexity in our study may just be a reflection of their students' general linguistic competence and writing ability rather than the direct influence of CLIL or even the EFL lessons.

This chapter has provided a discussion of the results obtained by the CLIL and the EFL groups in the three different skills tested in this study. The next chapter will present the final conclusions of the study as well as its pedagogical implications. It will also acknowledge its limitations and present directives for future research.

CHAPTER 7 Conclusions

This chapter draws conclusions from the multiple findings in this research project and acknowledges the limitations of the study. Some implications for learning and teaching in a CLIL setting are highlighted and suggestions are made for future research directions within this field of study.

7.1 Main findings and concluding remarks

The main aim of this piece of research was to find out the effects of two different programmes, an EFL and an EFL+CLIL programme, on learners' linguistic competence in three different skills, those of listening, reading and writing when the number of hours of exposure to the target language was kept constant. The study also aimed at determining the achievement and the progress of the students in the EFL+CLIL Science group compared to the EFL+CLIL students in the Arts & Crafts group. In order to attain reliable results, the number of hours of previous exposure to the language in the school context, as well as the initial proficiency of the learners were taken into account for statistical analysis. Quantitative as well as qualitative data obtained through class observations as well as through interviews with the relevant CLIL and EFL teachers, were used to reach the final conclusions.

The findings reported in this study could be considered to be unique in some aspects. Although many studies have investigated the effect of CLIL on the linguistic competence of learners, few studies such as this one have kept the number of hours of exposure to the language the same for the CLIL and EFL groups at data collection times, and very few have tried to take into account variables such as the initial proficiency level of the students as well as the amount of previous exposure to the language. Studies dealing with the linguistic results of students in general and young

learners in particular exposed to different CLIL subjects are also rare. Hence, in order to measure the different linguistic gains of the students involved in this project, three main research questions and several research subquestions were used to guide this study and will be answered below.

The main questions and subquestions were the following:

RQ1 Keeping the number of hours of exposure to English the same for both groups, CLIL and EFL, do the CLIL students' listening and reading skills benefit from their exposure to the CLIL experience?

RQ1.1 Are there any differences in achievement between CLIL and EFL learners statistically significantly in favour of CLIL learners at different times (T0, T1, T2, T3)?

RQ1.2 Are there any differences in progress between CLIL and EFL learners significantly in favour of CLIL learners after one year (T0-T1) and two years (T2-T3) of CLIL implementation? Are there differences in favour of CLIL students when we consider their progress from T0-T3?

Results obtained in Listening partially confirm the effectiveness of CLIL programmes of the type implemented in the schools in this project since, according to the results obtained, all the CLIL groups moderately benefit from the exposure to content in English when compared to their control groups. CLIL Science students progressed significantly throughout the two-year study, and achieved significantly different results at T3 with respect to the Control Group whereas, although Arts & Crafts students also improved significantly throughout the study, their achievement at T3 was not statistically significant. However, on the positive side, as has already been pointed out, the groups progressed significantly and the CLIL Science group improved significantly during the final time period (T2-T3).

As for Reading, results seem to confirm that CLIL groups made significant progress during the study. However, despite this improvement, no significant differences in achievement were found in favour of either CLIL Science students or CLIL Arts & Crafts learners. The students in Arts & Crafts were significantly outperformed by the Control Group at times T1 and T2 during the study and in several of the individual reading questions used to assess this skill, also at T1 and T2. The reading results at T3 showed no differences between the CLIL and the Control Groups.

RQ 1.3 How does the initial level of English proficiency affect the students' performance in the CLIL and the EFL group?

Taking into account that all groups had the same number of hours, no significant differences in the listening skill were found between CLIL High achievers in Science and CLIL Low achievers in Arts & Crafts when compared to their Control groups. However, the Low achievers in the CLIL Science group obtained significantly different results in Listening in relation to the Control group. High achievers in Arts & Crafts were significantly outperformed by their Control group. Despite the achievement results, High and Low achievers progressed throughout the study.

As for Reading, which once again involved the same number of hours of exposure for both groups, no differences in achievement were found between High and Low achievers in Science when compared to their Control Groups, nor were any found between Low achievers in Arts & Crafts and its Control Group. However, CLIL High achievers in Arts & Crafts were outperformed by their Control Group. Despite the achievement results, all groups progressed significantly at the end of the study.

RQ 2 Keeping the number of hours of exposure the same for both groups, CLIL and non-CLIL, do the CLIL students' writing skills measured in terms of complexity, accuracy and fluency benefit from their exposure to the CLIL experience?

RQ 2.1 In terms of fluency, are there differences in achievement between CLIL and EFL learners significantly in favour of CLIL learners at different times (T0, T1, T2, T3)?

In terms of achievement in Fluency, significantly different results were obtained by CLIL Arts & Crafts students at T3 in TNWE and TNU. The students in Arts & Crafts wrote more words in English and more units than their counterparts in the Control Groups. CLIL Science students attained significantly different results at T1 as for the %TNWE/TNW. No differences were found between the CLIL Science group and its Control group at T3.

RQ 2.2 In terms of fluency, are there differences in progress between CLIL and EFL learners significantly in favour of CLIL learners after one year (T0-T1) and two years (T2-T3) of CLIL implementation? Are there differences in favour of CLIL students when we consider their progress from T0-T3?

CLIL students in Science improved significantly in several of the measures used to assess fluency: TNW, TNWE and the TNU during two time periods: T1-T2 and T2-T3. Arts & Crafts students also improved in TNW and TNU during T1-T2 and in TNWE during two time periods: T1-T2 and T2-T3. Both CLIL Science students and those in Arts & Crafts progressed significantly during the study.

RQ 2.3 In terms of accuracy, are there differences in achievement between CLIL and EFL learners significantly in favour of CLIL learners at different times (T0, T1, T2, T3)?

Students in Arts & Crafts significantly outperformed their peers in the Control Group in all the measures used to assess accuracy (TNEFU and % TNEFU/TNU). No differences were found between CLIL Science and CLIL Arts & Crafts students compared to their Control Groups.

RQ 2.4 In terms of accuracy, are there differences in their progress between CLIL and EFL learners significantly in favour of CLIL learners after one year (T0-T1) and two years (T2-T3) of CLIL implementation? Are there differences in favour of CLIL students when we consider their progress from T0-T3?

CLIL students in Science and in Arts & Crafts improved significantly throughout the study in terms of TNEFU. The Science students outperformed their Control Group during the second year (T2-T3) and the students in Arts & Crafts also outperformed their Control Group counterparts from T1-T2.

RQ 2.5 In terms of complexity (lexical and syntactic), are there differences in achievement between CLIL and EFL learners significantly in favour of CLIL learners at different times (T0, T1, T2, T3)?

Arts & Crafts students significantly outperformed their Control group in TNAdj and the % TNAdj/TNWE at all times tested and in TNLV at T1 and T2. As for CLIL Science students, even though the Control Group outperformed the CLIL group at T1 in TNAdj and the ratio TNAdj/TNWE, CLIL Science students significantly outperformed their Control counterparts in TNAdj at T3. No differences were found between the CLIL Science and the Control Group during the study in terms of TNLV and the ratio TNLV/TNWE, except at T1 in the ratio TNLV/TNWE. In terms of Syntactic Complexity, no differences were found for any of the groups. In fact, no instances of subordination were found during the first year of the study.

RQ 2.6 In terms of complexity (lexical and syntactic), are there differences in progress between CLIL and EFL learners significantly in favour of CLIL learners after one year (T0-T1), two years (T2-T3) of CLIL implementation? Are there differences in favour of CLIL students when we consider their progress from T0-T3?

CLIL Science and Arts& Crafts made significant improvements throughout the study in Lexical Complexity. CLIL Science students progressed significantly from the second year of the study in the TNLV and the TNAdj, and CLIL Arts & Crafts students improved significantly mainly during the first year. As for Syntactic Complexity, the results showed significant improvement during the second year of the study for all students.

RQ 2.7 How does the initial level of English proficiency affect the students' performance in writing in the CLIL and the EFL group?

No significant differences were found between High and Low achievers in the EFL and the EFL+CLIL Science programme in Fluency, Accuracy and Syntactic Complexity. Only the High achievers in the CLIL Science group obtained significantly higher results in Lexical Complexity (TNAdj) than their control counterparts.

However, the results of the students in Arts & Crafts did reveal significant advantages for High and Low achievers in Fluency (TNU), Accuracy (TNEFU) and Lexical Complexity (TNAdj, %TNADJ/TNWE) with respect to the corresponding Control groups, but no differences were observed in Syntactic Complexity.

RQ 3 Do the CLIL Science students' listening, reading and writing skills benefit more than the Arts & Crafts students from their exposure to the CLIL experience?

RQ 3.1 Are there differences in achievement in listening and reading between CLIL Science students and CLIL Arts & Crafts students at different times (T0, T1, T2, T3)?

The comparison allowed us to see a distinct advantage in achievement in Listening for the students exposed to Science in English at T3. As for Reading, no differences were found between students in Science and Students in Arts & Crafts.

RQ 3.2 In terms of writing measured as Complexity, Accuracy and Fluency, are there differences in achievement between CLIL Science and CLIL Arts & Crafts students at different times (T0, T1, T2, T3)?

The students exposed to Arts & Crafts in English outperformed their CLIL counterparts in Science in several aspects of Fluency at T1 and T2 in TNWE, TNW and TNU. They also outperformed their counterparts in Science in all measures in Accuracy and in TNAdj and %TNADJ/TNWE in Lexical Complexity at all times tested.

RQ 3.3 How does the initial level of English proficiency affect the students' performance in writing in the CLIL Science and the CLIL Arts & Crafts group?

No significant differences were found between CLIL Science High achievers and CLIL Arts & Crafts High achievers in Listening. However, High achievers in Science outperformed those in Arts & Crafts in Reading at T2.

Low achievers in Arts & Crafts, however, did manage to attain some significant results during the second year of the study in Listening, Fluency (TNWE and TNU), Accuracy (TNEFU) Lexical Complexity (TNADj, %TNAdj/TNWE) and Syntactic Complexity (ISU).

The present study therefore provides partial statistical evidence to support both the questions examined as well as the effectiveness of the CLIL programmes applied in primary schools in our context. Some of the CLIL students' listening ability improved more than the students in the Control Group, as did their writing capacity, which mainly improved in aspects such as Fluency, Accuracy and Lexical Complexity. The groups exposed to the combination EFL+CLIL significantly progressed throughout the study, although most of the significant results in progress appeared during the second year. The results also provide partial evidence to support the effectiveness of CLIL in improving the language competence of Low achievers, as they managed to significantly

outperform the Low achievers in the Control Group in some of the aspects analysed. Although not all results have been statistically significant in favour of CLIL, only a few of them have come out in favour of the Control Groups. This, together with a tendency evident in the descriptive results of the analysis, which show better mean scores in general from the CLIL groups, pushes us towards the first conclusion: the CLIL approach may benefit the interlanguage development of young learners. Although this finding has already been reported in other studies, as cited in previous sections, what makes it relevant here is that the results in this study were obtained keeping the number of hours of exposure the same between the CLIL and the Control Groups at data collection time, and that only one hour a week was devoted to CLIL. In other words, exposing the students to just one hour of CLIL every week, in comparison with those students who just received EFL hours, has produced the same or better results in some of the abilities tested in this study. This conclusion is drawn specifically from some of the positive results obtained by the CLIL groups in different skills, as well as from the fact that, although not significant at all times, CLIL students obtained better mean scores than their counterparts in the Control Groups and, most importantly of all, from the fact that, on very few occasions during the study, CLIL students were significantly outperformed by the Control Groups. Taking all of the above into account, an increased exposure to CLIL could produce even better results.

However, the variability in the findings of the different groups involved in the project drives us towards the second conclusion in this study: second language success in a CLIL programme is not only a matter of the input received in the CLIL classes, but also a matter of the *type* of input received in combination with the *practice* in the different skills and scaffolding strategies provided by the instructors. Therefore, the subject chosen and the methodological approach have an important role in the

implementation of any CLIL programme. This conclusion was drawn from the positive results obtained by the groups exposed to different subjects (Science or Arts & Crafts) in the three different skills analysed in this project: Listening, Reading and Writing. The Listening skills of the students exposed to Science in general, and those of the Low achievers in particular, benefitted in the short term in this project, whereas the Listening skills of the Arts & Crafts students did not. The different results between the Science and Arts & Crafts groups were explained by the effect that exposure to challenging comprehensible scientific discourse, together with the challenging scaffolding strategies provided by the CLIL Science teacher, had on the students' listening ability. Being exposed to challenging input seems to benefit the listening skills of the students, as the extra cognitive effort made to understand difficult concepts, without relying constantly on visual help, seems to trigger the mechanisms necessary to increase the listening capacities of young learners. However, the Reading results which were at all times in favour of the Control Groups can be explained by the poor exposure and practice in this skill, as well as by the poor ability on the part of the teachers in adapting and simplifying the challenging reading texts used in the class. Fluency and Complexity in Writing, especially in Arts & Crafts, also benefitted from continuous classroom practice as it provided opportunities for output through guided writings, thus helping young learners to understand the communicative value of the written texts.

Another conclusion reached from findings on Accuracy is that the success in CLIL is not only determined by exposure to cognitively challenging input but also by the focus on form approach used by the instructor, which helps learners to process language through content. This conclusion was drawn from the fact that CLIL students in Arts & Crafts obtained better results in Accuracy than students in the Science class. It was suggested that due to the challenging nature of the content in Science, the cognitive

effort required on the part of the students was done at the expense of the progress in language, in spite of the scaffolding strategies used by the teacher. Therefore, the students' linguistic progress, as was seen in the results at the end of the study, was not as good as that of the students in Arts & Crafts, who managed to improve in terms of Accuracy. This was probably due to the fact that the less challenging nature of the topics in the Arts classes, in combination with continuous correcting feedback on the part of the teacher and the transfer of grammatical skills from the EFL classes led the students to concentrate much more on the language, noticing certain linguistic aspects and thus producing more accurate written forms. This situation was also assisted by the cooperation and joint work carried out by the EFL and CLIL teachers, which helped students to understand the role of language in the CLIL classes and make the necessary connections to transfer the skills learnt in the EFL classes.

Furthermore, it can also be concluded that language gains in CLIL may be a long-term process which may be influenced, as explained above, by the cognitive demands of content. The idea of a long-term project, especially with young language learners, can be supported from findings in the progress of the students. Most of the significant results attained by Science and Arts & Crafts students in improvement in the different skills analysed were obtained during the second year of the implementation of the CLIL projects in the schools. This idea can also be supported by the fact that, although not statistically significant, the students in CLIL obtained better scores at the end of the project than the students in the Control Groups. The nature of the input received in CLIL represented a challenge for young learners, who needed time to adapt to new forms of learning, and CLIL was, for them, a challenging new way of learning content and language at the same time. We can attribute this to maturational constraints and to the fact that the students are in constant cognitive and personal development. As

they grow older, they become more efficient learners, and this fact may have been reflected in the outcome of CLIL. Incidental language learning requires large amounts of exposure to the language and this was not the case of the students involved in this project, which was implemented in a minimal input situation.

Another conclusion reached from findings on the influence of the initial proficiency level of the students is that CLIL Low achievers seem to benefit from the CLIL approach. This conclusion was drawn from the fact that Low achievers in Science outperformed the Low achievers in the Control group in Listening, and Low achievers in Arts & Crafts outperformed the Low achievers in the Control group in Writing (Fluency, Accuracy and Lexical Complexity). As was suggested in the discussion of the results, Low achievers not only benefit from immersion in CLIL, but also from the scaffolding strategies used by the teachers. Studying content through a foreign language may have helped Low achievers to make greater cognitive efforts, understand the communicative value of language, and therefore motivate them to learn this new language.

A final conclusion can be reached from the findings in this study. Students in an EFL+ CLIL programme not only benefit from the CLIL approach. They also benefit from the transfer of skills that they carry forward from their EFL lessons. In order to get good results, instructors need to support each other so that students can transfer skills from one type of approach to the other, thus contributing to the general improvement in the language level of the students at school.

The success in CLIL can be seen as a combination of many overlapping and intersecting contextual factors, which include elements of the extra-curricular context, the educational or curricular context and the learners' individual context (Housen et al., 2011). The findings presented in this dissertation are the results of an experimental

project carried out in a context of minimal input situation, in which the main source of input for the students was the input they received in the EFL and CLIL classes. From an educational perspective, the project was an experimental exercise carried out in three primary schools which decided to teach a curricular subject in English for the first time, with the objective of improving the pupils' language competence in English. The project was carried out by teachers with many years of teaching experience as primary teachers, but with relatively little experience and training in CLIL, who were keen and enthusiastic about the project. From the learners' perspective, the participants were very young learners in their final two years of primary schooling, who had been learning English as a foreign language at school since their very first years. Therefore, the introduction of CLIL required, on their part, the necessary time to adapt to a new way of learning the language. Factoring in all these elements, it can be said that the results can be used to support the theory that, given the right and sufficient exposure and teaching conditions (linguistically and methodologically qualified teachers, good pedagogical methods and classroom practices, adequate teacher training and suitable resources, among many other conditions), the CLIL approach can improve the language competence of pupils in primary education.

However, the results of the investigation have also highlighted a number of challenges schools may face if the proper conditions for the correct use of CLIL are not implemented: the importance of the teacher's linguistic competence, as well as the need to receive appropriate CLIL training, both of which are necessary in order to allow teachers to foster genuine, good communication in the classroom as well as offering cognitively-demanding tasks to young learners and preparing adequate materials for the CLIL lessons. CLIL has always been considered an umbrella term under which different approaches to content-based teaching can be found. However, only under the right

conditions may learners fully benefit from high-quality teaching and from input that is both meaningful and understandable.

7.2 Limitations and further research

Some limitations of the present study will be acknowledged in this section. They are concerned with the generalization of the findings to the whole CLIL approach, the need to isolate the influence of some aspects that may have affected the outcome of the study and the nature of the tasks chosen. Directions for future research will also be addressed.

In the first place, one of the most obvious limitations is the generalization of the findings to a broader context. There are two questions here. The first question concerns the type of programme applied in the schools. The results mirror the situation of CLIL in Catalonia, where many schools devote just one hour a week to CLIL. There are several reasons for this situation, including the fact that the school syllabus is already overcrowded, and that there has been to date very little research done in our context trying to prove the effectiveness of CLIL in terms of language learning. As for the second question, the participants were young learners in CLIL, and therefore the findings can only be generalized in the context of primary education. Research-based longitudinal empirical studies, such as this one, should be carried out in similar school contexts at secondary level to allow comparisons to be made between the different educational stages and to allow researchers to see the development of CLIL in the long run.

A second limitation concerns the large amount of results and their variability in this study, which has made it difficult to draw very specific conclusions. As has already been said, the results in this study may have been influenced by the particular contextual factors related to the schools involved in the project, such as the teachers, the linguistic

project of the centres and the methodological aspects in the implementation of the CLIL approach.

Finally, there is a limitation concerning the tests administered in this study. As participants were students in their final two years of primary education, the tests chosen were those specifically designed to determine the language level of young students mainly exposed to EFL courses. As research in the context of CLIL and young learners is still at an emerging stage, there is a lack of specific instruments to measure the language gains in a CLIL setting. More finely-grained tests should be a priority for further research.

Data has been analysed from just three perspectives: listening, reading and writing. Other areas such as speaking and motivation have not been considered in this study due to the lack of time and space. The analysis of these areas would undoubtedly provide interesting findings on the implementation of CLIL.

Finally, CLIL is all about learning content through a foreign language. This study has looked at the linguistic gains achieved from the implementation of the CLIL approach in primary schools. It is very important, however, to make sure that the implementation of CLIL is not carried out at the expense of content learning. Further and more extensive research is necessary in order to determine whether a longer period of instructional time, greater exposure to the language, the use of informed methodology and strategies and properly trained teachers will provide even better results. CLIL definitely seems to be the best way forward in terms of enhancing foreign language teaching in primary education.

REFERENCES

- Ackerl, C. (2007). Lexico-Grammar in the essays of CLIL and non-CLIL students: error analysis of written production. *Vienna English Working Papers*, 16 (3): 6-11.
- Adler, J. (1998). A Language of Teaching Dilemma: Unlocking the Complex Multilingual Secondary Mathematics Classroom. *The Learning of Mathematics*, 18 (1): 24-33.
- Admiraal, W., Westhoff, G., and de Bot, K. (2006). Evaluation of Bilingual secondary education in The Nederland's: Students' language proficiency. *Educational Research and Evaluation* 12 (1):75-93.
- Aguilar, M. and Muñoz, C. (2013). The effect of proficiency on CLIL benefits in Engineering Students in Spain. *International Journal of Applied Linguistics*, 15(2): 183-197.
- Agustín Llach, M.P. (2009). The role of Spanish L1 in the vocabulary of CLIL and non-CLIL EFL learners. In Y. Ruiz de Zarobe and R.M. Jiménez-Catalán (eds.), *Content and Language Integrated Learning. Evidence from Research in Europe*. Bristol: Multilingual Matters. 112-129.
- Airey, J. (2004). Can you teach it in English? Aspects of the language choice debate in Swedish Higher Education. In R Wilkinson (ed), *Integrating Content and Language.Meeting the Challenge of a Multilingual Higher Education*. Maastricht university. 97-108.
- Airey, J. and Linder, C. (2006). Language and the experience of learning university physics in Sweden. *European Journal of Physics*, 27 (3): 553-60.
- Alba, J.O. (2009). Themes and vocabulary in CLIL and non-CLIL instruction. In Y.Ruiz de Zarobe, and R.M.J Catalán (eds.), *Content and Language Integrated Learning: Evidence from research in Europe*. Bristol: Multilingual Matters. 130-156.
- Aline, D. and Hosoda, Y. (2006). Team teaching participation patterns of homeroom teachers in English activities classes in Japanese public elementary schools. *JALT Journal*, 28(1): 5-22.
- Anderson, J.R. (1985). *Cognitive psychology and its implications* (2nd ed) New York: Freeman.
- Atkinson, R. C. and Shiffrin, R. M. (1971). *The control processes of short-term memory*. Institute for Mathematical Studies in the Social Sciences, Stanford University.
- Baddeley, A. (2003). Working memory and language: an overview. *Journal of communication disorders*, 36 (3): 189-208.
- Baetens-Beardsmore, H. (eds.), (1993). *European Models of Bilingual Education*. Clevedon: Multilingual Matters.
- Bailey, A.L., Butler, F. A., and Sato, E. (2007). Standards to standards linkage under title III: Exploring common language demands in ELD and science standards. *Applied Measurement of Education*, 20 (1): 53-78.
- Baker, C. R. and Jones, S. P. (eds.). (1998). *Encyclopedia of bilingualism and bilingual education*. Multilingual Matters.
- Ballester, C. F. (2010). Llicència, A. Pràctica Docent Efectiva AICLE. Available at <http://www.xtec.cat/sgfp/llicencies/200910/memories/2052m.pdf>

- Bange, P. (2005). *L'apprentissage d'une langue étrangère: cognition et interaction*. Paris: L'Harmattan.
- Bernaus, M., Masgoret, A., Gardner, R.C., and Reyes, E. (2004). Motivation and attitudes towards learning languages in multicultural classrooms. *The International Journal of Multiculturalism* 1(2): 75-89.
- Berry, C. F. and Mindes, G. (1993). *Planning a Theme-Based Curriculum: Goals, Themes, Activities, and Planning Guides for 4's and 5's*. Good Year Books.
- Bialystok, E. (1991). *Language Processing in Bilingual Children*. Cambridge: Cambridge University Press.
- Bialystok, E. (1994). Analysis and control in the development of second language proficiency. *Studies in Second Language Acquisition*, 16, 157-168.
- Bialystok, E. (2001). *Bilingualism in Development: Language, Literacy and Cognition*. Cambridge: Cambridge University Press.
- Blau, E. K. (1982). The effect of syntax on readability for ESL students in Puerto Rico. *TESOL Quarterly*, 16, 517-526.
- Bley Vroman, R. (1989). What is the logical problem of foreign language learning? In S.Gass and J.Schachter (eds.), *Linguistic Perspectives on Second Language Acquisition*, 41-68. Cambridge: Cambridge University Press.
- Blondin, C., Candelier, M., Edelenbos, P., Johnstone, R., Kubanek-German, A., and Traute, T. (1998). Foreign Languages in Primary and Pre-School Education: Context and Outcomes. A review of recent research within the European Union. London: CILT.
- Bret Blasco, A. (2011). *Implementing CLIL in a primary school in Spain: the effects of CLIL on L2 English learners' oral production skills*. MA thesis. Unpublished manuscript. Universitat Autònoma de Barcelona
- Brinton, D.M., Snow, M.A., and Wesche, M.B. (1989). *Content-based Second Language Instruction*. New York: Newbury House.
- Brumfit, C. (1998). Language teaching and information technology. CILT Research Forum *Information technology: The pedagogical implications for language teaching and learning*. Available at <http://www.cilt.org.uk/research/resfor2/brumfit.htm>.
- Bruner, J. (1985). Child's talk: Learning to use language. *Child Language Teaching and Therapy*, 1(1): 111-114.
- Bruton, A. (2011). Is CLIL so beneficial, or just selective? Re-evaluating some of the research. *System*, 39 (4): 523-532.
- Bruton, A. (2013). CLIL: Some of the reasons why... and why not. *System*, 41 (3): 587-597.
- Burstall, C. (1975). Primary French in the balance. *Foreign Language Annals*, 10 (3): 245-252.
- Butler, Y. G. (2004). What level of English proficiency do elementary school teachers need to attain to teach EFL? Case studies from Korea, Taiwan, and Japan. *Tesol Quarterly*, 38(2): 245-278.
- Butler, Y. G. (2005). Comparative perspectives towards communicative activities among elementary school teachers in South Korea, Japan, and Taiwan. *Language Teaching Research*, 9, 423-446.

- Cameron, L. (2001). *Teaching Languages to Young Learners*. Cambridge Language Teaching Library. Cambridge: Cambridge University Press.
- Canale, M. and Swain, M. (1980). Theoretical Bases of Communicative Approaches to Second Language Teaching and Learning. *Applied Linguistics* 1 (1): 1-47.
- Carroll, J. B. (1991). No demonstration that < i> g</i> is not unitary, but there's more to the story: Comment on Kranzler and Jensen. *Intelligence*, 15 (4): 423-436.
- Celaya, M.L., Pérez-Vidal, C., and Torras, M.R. (2001). Matriz de criterios de medición para la determinación del perfil de competencia lingüística escrita en inglés (LE). *Revista Española de Lingüística Aplicada*, 14, 87-98.
- Cenoz, J (2002) Age and the acquisition of English as a third language in the Basque Autonomous Community: Research Perspectives. In F. Etxeberria & U. Ruiz Bikandi (eds.), *¿Trilingües a los 4 años?. Donostia: Ibaeta Pedagogia*. 65-83.
- Cenoz, J. (2003). The influence of age on the acquisition of English: general proficiency, attitudes and code mixing. In M.P. García Mayo and M.L. García Lecumberri (eds.), *Age and the Acquisition of English as a Foreign Language: Theoretical Issues and Field Work*. Clevedon: Multilingual Matters. 77-93.
- Cenoz, J. (2004). Teaching English as a third language: The effect of attitudes and motivation. *Trilingualism in Family, School and Community*, 202-218.
- Ceuleers, E. (2008). Variable Identities in Brussels. The Relationship between Language Learning, Motivation and Identity in a Multilingual Context. *Journal of Multilingual and Multicultural development*, 29 (4): 291-309.
- Chamot, A.U. and O'Malley, J.M. (1994). *The CALLA handbook: Implementing the Cognitive Academic Language Learning Approach*. White Plains, New York: Addison Wesley Longman
- Charboneau, R. (2012). Approaches and practices relating to the teaching of EFL reading at the Norwegian primary level. *The Young Language Learner: Research-based insights into teaching and learning*, 51-70.
- Clewell, B. and Campbell, P. (2004). *Highly Effective USL Schools: An Outlier Study*. Washington, The Urban Institute/ Campbell-Kibler.
- Cohelo, E. (2005). *La transversalidad del aprendizaje lingüístico en el Currículo: el proceso de andamiaje en la educación. Multilingüismo, competencia lingüística y nuevas tecnologías. Cuadernos de Educación*. ICE Universitat de Barcelona, 5.
- Collentine, J. and Freed, B. (2004). Learning context and its effects on second language acquisition. *Studies in Second Language Acquisition*, 26 (2): 153-171.
- Collier, V. (1987). Age and rate of Acquisition of Second Language for Academic Purposes. *TESOL Quarterly*, 21, 617-641.
- Collier, V. (1989). How long? A synthesis of Research on Academic Achievement in Second Language. *TESOL Quarterly*, 23, 509-531.
- Collier, V. (1992). A Synthesis of Studies Examining Long-Term Language Minority Student Data on Academic Achievement. *Bilingual Research Journal*, 16 (1-2): 187-212.

- Collier, V. P. (1995). *Acquiring a second language for school* (p. 19). Washington, DC: National Clearinghouse for Bilingual Education.
- Cook, V. (2001). Using the first language in the classroom. *Canadian Modern Language Review/La Revue canadienne des langues vivantes*, 57 (3): 402-423.
- Coonan, M. (2007). Insider views of the CLIL class through teacher self observation introspection. *International Journal of Bilingual Education and Bilingualism* 10, (5): 625-46.
- Coyle, D. (1999). Theory and planning for Effective Classrooms: Supporting Students in Content and Language Integrated Learning Contexts. In J. Masih (ed.), *Learning Through a Foreign Language*. London: CILT (Centre for Information on Language Teaching Research)
- Coyle, D. (2000). Meeting the Challenge-The 3Cs curriculum. In S. Green (ed) *New perspectives on Teaching and Learning Modern Languages*. Clevedon: Multilingual Matters.
- Coyle, D. (2002). Against All Odds: Lessons from Content & Language Integrated Learning in English Secondary Schools. In W.C. Daniel and G.M. Jones (eds.), *Education and Society in Plurilingual Contexts* Brussels: VUB Brussels University.
- Coyle, D. (2005). The Teaching Observatory: exploring zones of interactivity. In Holmberg, G., Shelley, M. and White, C., (eds.), *Languages and Distance Education: Evolution and Change*, Clevedon: Multilingual Matters.
- Coyle, D. (2006). Content and Language Integrated Learning: Motivating learners and teachers. *Scottish Languages Review*, 13, 1-18.
- Coyle, D. (2006). Developing CLIL: Towards a theory of practice. *CLIL in Catalonia, from theory to practice*. APAC Monographs, 6.
- Coyle, D. (2007). Content and Language Integrated Learning: Towards a Connected Research Agenda for CLIL Pedagogies. *The International Journal of Bilingual Education and Bilingualism*. 10 (5): 543-562.
- Coyle, D., Hood, P., and Marsh, D. (2010). *CLIL: Content and language integrated learning*. Cambridge: Cambridge University Press.
- Crandall, J. and Tucker, G.R. (1990). Content-based Instruction in Second and Foreign Languages.. In A. Sanivan (ed.), *Language Teaching Methodology for the Nineties*. Singapore: SEAMEO Regional Language Centre. 83-96.
- Crawford, J., & Krashen, S. (2007). *English learners in American classrooms: 101 questions, 101 answers*. Teaching Resources. New York: Scholastic Inc.
- Cummins, J. (1984). *Bilingualism and Social Education: Issues in Assessment and Pedagogy*. Clevedon: Multilingual Matters.
- Cummins, J. (1991). *Interdependence of First and Second Language Proficiency*. In E. Bialistock (ed.), *Language processing in Bilingual Children*. Cambridge: Cambridge University Press. 79-88.
- Cummins, J. (2003). Bilingual Children's Mother Tongue: Why is it Important for Education? *Sprogforum*, 19, 15-20

- Cummins, J. and Swain, M. (1983). Analysis-by Rhetoric: Reading the Text or the Reader's own projections?. *Applied Linguistics*, 4, 22-41.
- Cummins, J. and Swain, M. (1986). *Bilingualism and Education: Aspects of Theory, Research and Practice*. London: Longman.
- Curtain, H. (2009). Assessment of early learning of foreign languages in the USA. In M. Nikolov (ed.), *The age factor and early language learning*. Berlin: Mouton de Gruyter. 59-82.
- Czura, A., Papaja, K., and Urbaniak, M. (2009). Bilingual Education and the emergence of CLIL in Poland. In D. Marsh, P. Mehisto, D. Wolff, R. Aliaga, T. Asikainen, M.J. Frigols, S. Hugues, and G. Langé (eds.), *CLIL Practice. Perspectives on the field*. Finland: University of Jyväskylä. 172-188.
- Dafouz Milne, E. (2006). Solidarity strategies in CLIL university lectures: teacher's use of pronouns and modal verbs. *Vienna English Working Papers* 15 (3): 9-14.
- Dafouz Milne, E. (2011). English as the medium of instruction in Spanish Contexts: A look at teacher discourses. In Y. Ruiz de Zarobe, J.M. Sierra and F. Gallardo del Puerto (eds.), *Content and Foreign Language integrated learning. Contributions to Multilingualism in European contexts*. Frankfurt.am-Main: Peter lang. 189-209.
- Dafouz, E. and Guerrini, M. C. (2009). *CLIL across educational levels :Experiences from primary, secondary and tertiary contexts*. Madrid: Richmond.
- Dalton-Puffer, C. (2006). Questions in CLIL Classrooms: Strategic questioning to encourage speaking. In Martinez Flor, A. and Usó, E. (eds.), *Current trends in the development and teaching of the four skills*. Berlin: Mouton de Gruyter, 187-213.
- Dalton-Puffer, C. (2007). *Discourse in CLIL classrooms*, Amsterdam: John Benjamin.
- Dalton-Puffer, C. (2008). Outcomes and Processes in Content and Language Integrated Learning. (CLIL): Current Research from Europe. In D. Werner and V. Laurenz (eds.) *Future Perspectives in English Language Teaching*. Heidelberg: Carl Winter. 139-157.
- Dalton-Puffer, C., and Nikula, T. (2006). Pragmatics of content-based instruction: teacher and student directives in Finnish and Austrian classrooms. *Applied Linguistics*, 27 (2): 241-267.
- Dalton-Puffer, C., Nikula, T., and Smit, U. (2010). *Language use and language learning in CLIL classrooms*. Amsterdam: John Benjamins
- Dalton-Puffer, C. and Smit, U (eds.) (2007). *Empirical Perspectives in CLIL Classroom Discourse*. Frankfurt: Peter lang.
- Darn, S. (2009). Teaching Other Subjects through English, Cross-Curricular Resources for Young Learners, Uncovering CLIL. *ELT journal*, 63(3): 275-277.
- De Bot, K. (1996). The psycholinguistics of the output hypothesis. *Language Learning*, 46, 529-555.
- De Corte, E. (2000). Marrying Theory Building and the Improvement of School Practice: A Permanent Challenge for Instructional Psychology. *Learning and Instruction*, 10(3), 249-266.

- De Graaff, R., Koopman, J., and Westhoff, G. (2007). Identifying Effective L2 Pedagogy in Content and Language Integrated Learning (CLIL). *Vienna English Working Papers*, 16 (3):12-20.
- De Graaff, R., Koopman, G. J., Anikina, Y. and Westhoff G. (2007). An observation tool for effective L2 pedagogy in content and language integrated learning (CLIL). *International Journal of Bilingual Education and Bilingualism*, 10 (5): 603-624.
- De Jabrun, P. (1997). Academic Achievement in Late Partial Immersion French. *Babel: Australia*, 32 (2): 20-23
- DeKeyser, R. (1998). Beyond focus on form: Cognitive perspectives on learning and practicing second language grammar. In C. Doughty, and J. Williams (eds.), *Focus on form in classroom second language acquisition*. Cambridge, UK: Cambridge University Press. 42-63.
- DeKeyser, R. (2000). The robustness of critical period effects in second language acquisition. *Studies in Second Language Acquisition* 22, (4): 499-533.
- DeKeyser, R. (2003). Implicit and Explicit Learning. In C.J. Doughty and M.H. Long (eds.), *Handbook of Second Language Acquisition*. Oxford: Blackwell. 313-348.
- DeKeyser, R. (Ed.). (2007). *Practice in a second language: Perspectives from applied linguistics and cognitive psychology*. Cambridge University Press.
- DeKeyser, R. and Larson-Hall, J. (2005). What does the critical period really mean? In J.F.Kroll and A.M.B. de Groot (eds.), *Handbook of Bilingualism: Psycholinguistic approaches*. Oxford: Oxford University Press. 88-108.
- Djigunović, J.M. (1998). Uloga afektivnih faktora u učenju stranoga jezika [Role of affective factors in FL learning]. Zagreb: Filozofski fakultet Sveučilišta u Zagrebu.
- Djigunovic, J.M. (2009). Individual differences in early language programmes. In M. Nikolov (Ed.), *The age factor and early language learning*. Berlin: Mouton de Gruyter. 199-225
- Djigunović, J.M. (2010). Starting age and L1 and L2 interaction. *International Journal of Bilingualism*, 14 (3): 303-314.
- Djigunovic, J.M. and Krevelj, S. (2010). Instructed early SLA: Development of attitudes. *SRAZ*, 54, 127-146.
- Donaldson, M. (1978). *Children's minds*. Harper Collins UK
- Donato, R., Tucker, G.R., and Igarashi, K. (2000). Converging Evidence, Attitudes, Achievements and Instruction in the Later Years of FLES. *Foreign Language Annuals*, 33, 377-393.
- Dörnyei, Z. (1990). Conceptualising Motivation in Foreign Language Learning. *Language Learning* 40 ,(1): 45-78.
- Dörnyei, Z. (1994). Motivation and Motivating in the foreign language classroom. *The Modern Language Journal* 78, (3): 273-284.
- Dornyei, Z. and Skehan, P. (2003). Individual Differences in Second Language Learning. In C. J., Doughty, and M. H. Long, (eds.), *The Handbook of Second Language Acquisition*, 589-630.

- Dörnyei, Z. (2009). *The Psychology of Second Language Acquisition*. Oxford: Oxford University Press
- Doughty, C. and Williams, J. (1998). Pedagogical choices in focus on form. In C. Doughty and J. Williams (eds.), *Focus on form in classroom second language acquisition*. Cambridge: Cambridge University Press. 197-261.
- Drew, I. (2010). A longitudinal study of young language learners' development in written English. *Norsk Lingvistisk Tidsskrift. Argang*, 28, 39-72.
- Drew, I. and Hasselgreen, A. (2008). Young language learner (YLL) research: An overview of some international and national approaches. *Acta Didactica Norge*, 2 (1): 5.
- Drew, I. and Pedersen, R. (2012). Readers Theatre: a group reading approach to texts in mainstream EFL classes. In A. Hasselgreen, I. Drew and B. Sorheim (eds.), *The Young Language Learner. Research based Insights into Teaching and Learning*. Bergen: Fagbokforlaget. 71-85.
- Drew, I. and Sorheim, B. (2009). *English Teaching Strategies*. Oslo: Det Norske Samlaget.
- Echevarria, J. and Graves, A. (1998). *Sheltered content instruction*. Boston: Pearson Allyn and Bacon.
- Echevarria, J., Vogt, M. E. and Short, D. J. (2007) *Making Content Comprehensible for English Learners: The SIOP Model*. Boston: Pearson Allyn and Bacon.
- Edelenbos, P., Johnstone, R., and Kubanek, A. (2006). The main pedagogical principles underlying the teaching of foreign languages to very young learners. Languages for the Children of Europe. Published research, good practice and main principles, European Commission Final report 89/04. (Available at <http://ec.europa.eu/education/languages>)
- Edelenbos, P. and Johnstone, R. (2009). Early foreign language learning: Published research, good practice and main principles. *The age factor and early language learning*, 40, 39.
- Ellis, R. (1985). *Understanding Second Language*. Oxford University Press.
- Ellis, R. (1994). *The Study of Second Language Acquisition*. Oxford University Press.
- Ellis, R. (1997). *Second Language Acquisition*. Oxford: Oxford University Press.
- Ellis, R. (2001). *Form Focus Instruction and Second Language Learning*. Oxford: Blackwell.
- Ellis, R. (2004). Modelling Learning Difficulty and Second Language Proficiency: The Differential Contributions of Implicit and Explicit Knowledge. *Applied Linguistics*, 27, 431-460.
- Ellis, R. (2008). *The Study of Second Language Acquisition* (2nd ed). Oxford: Oxford University Press.
- Ellis, R. and Barkhuizen, G. P. (2005). *Analysing Learner Language*. Oxford: Oxford University Press.
- Enever, J. (Ed.) (2011). *ELLiE: Early language learning in Europe*. London: The British Council.
- Enever, J., Moon, J., and Raman, U. (eds.) (2009). *Young learner English language policy and implementation: international perspectives*. Reading, UK: Garnet Education Publishing.

- Erlam, R. (2005). Language aptitude and its relationship to instructional effectiveness in second language acquisition. *Language Teaching Research*, 9(2): 147-171.
- Escobar, C. and Unamuno, V. (2008) Languages and language learning in Catalan Schools: From the bilingual to the multilingual challenge. In A. Hélot and A.M. De Mejía (eds.), *Forging Multilingual Spaces. Integrated Perspectives Towards Bilingual Education Bridging: the Gap between Prestigious Bilingualism and the Bilingualism of Minorities*. Multilingual Matters.
- Escobar, C. (2010). Pre-service CLIL Teacher-Education in Catalonia: expert and novice practitioners teaching and reflecting together. In D.Lasagabaster and Y. Ruiz de Zarobe, *CLIL in Spain: Implementation, Results and Teacher Training*. Newcastle, UK: Cambridge Scholars Publishing. 189-218.
- Escobar Urmeneta, C. (2012). Content-Rich Language Learning in Context-Rich Classrooms. *APAC*, 74, 39-47
- Escobar, C. and Nussbaum, L. (2008). Tareas de intercambio de información y procesos de aprendizaje en el aula AICLE. In A. Camps and M. Milian (ed.), *Miradas y voces. Investigaciones sobre la educación lingüística y literaria en entornos plurilingües*. Barcelona: Graó.
- Escobar, C. and Pérez-Vidal, C. (2004). Teacher education for the implementation of a Content and Language Integrating Learning approach (CLIL) in the school system. In R. Wilkinson (ed.), *Integrating Content and Language. Meeting the challenge of multilingual education*. Maastricht: Universitaire Pers Maastricht. 402-415.
- Escobar, C. and Sánchez Sola, A. (2009). Language Learning through Tasks in a CLIL Science Classroom. *Porta Linguarum* 11, 65-83.
- European Commission (1995). *White Paper on Education and Training. Teaching and Learning Towards the Learning Society*. Brussels: European Commission.
- European Commission, Directorate-General for Education and Culture, Vocational Training and Language Policy (2005). *Europeans and Languages: A Eurobarometer Special Survey*. (Available at <http://eacea.ec.europa.eu/page/portal/Eurydice>).
- European Commission (2007). *Commission Working Document: Report on the Implementation of the Action Plan "Promoting Language Learning and Linguistic Diversity"*. Brussels.
- Eurydice European Unit (2006). Content and Language Integrated Learning (CLIL) at school in Europe. European Commission. Eurydice, Belgium.
- Evnitskaya, N. and Aceros, J.C. (2008). "We are a good team": El contrato didáctico en parejas de aprendices de lengua extranjera. *Revista Española de Lingüística Aplicada* 21, 45-70.
- Felix, S. (1985). More evidence on competing cognitive systems. *Second Language Research* 1, 47-72.
- Fernandez Fontecha, A. (2009). Spanish CLIL: Research and official actions. In Y. Ruiz de Zarobe and R.M. Jiménez-Catalán (eds.), *Content and Language Integrated Learning. Evidence from research in Europe*. Bristol: Multilingual Matters. 3-21.
- Fernández, R. and Halbach, A. (2011). Analysing the Situation of Teachers in the CAM Bilingual Project after Five Years of Implementation. In Y. Ruiz de Zarobe, J.M.

- Sierra and F. Gallardo del Puerto (eds.), *Content and Foreign Language Integrated Learning: Contributions to Multilingualism in European Contexts*. Berlin: Peter Lan. 241-270
- Fernández, R., Pena, C., García, A., and Halbach, A. (2005). La implantación de proyectos educativos bilingües en la Comunidad de Madrid: las expectativas del profesorado antes de iniciar el proyecto. *Porta Linguarum* 3, 161-173.
- Flynt, E. S. and Brozo, W. (2010). Visual literacy and the content classroom: A question of now, not when. *The Reading Teacher*, 63 (6): 526-528.
- Fortanet Gómez, I. and Ruiz Garrido, M.F. (2009). Sharing CLIL in Europe. In M.L. Carrió-Pastor (ed.), *Content and Language Integrated Learning: Cultural Diversity*. Frankfurt-am-main: Peter Lang. 47-75.
- Frigols, M.J., Marsh, D., and Naysmith, J. (2007). Competence Building for Teachers of CLIL. In D. Marsh and D. Wolff (eds.), *Diverse Contexts. Converging Goals. CLIL in Europe*. Peter Lang. Frankfurt.
- García, O. (2009). *Bilingual education in the 21st century: A global perspective*. Chichester: Wiley-Blackwell.
- García Mayo, M. P. (2003). Age, length of exposure and grammaticality judgements in the acquisition of English as a foreign language. In M. P. García Mayo and M. L. García Lecumberri (eds.), *Age and the Acquisition of English as a Foreign Language*. Clevedon, MultilingualMatters. 94-114.
- Gardner, H. (1983). *Frames of Mind*. New York: Basic Books Inc.
- Gass, S. (1997). *Input, interaction and the second language learner*. Mahwah, N.J.: Lawrence Erlbaum.
- Gassner, D. and Maillat, D. (2006). Spoken Competence in CLIL. A pragmatic taken on recent Swiss data. In *Vienna English Working Papers* 15 (3): 15-22.
- Gatbonton, E. and Segalowitz, N. (2005). Rethinking communicative language teaching: A focus on access to fluency. *Canadian Modern Language Review*, 61 (3):325-353.
- Generalitat de Catalunya. Departament d'Ensenyament.(1992). *Curriculum. Educació Primària*. Barcelona: Departament d'Ensenyament.
- Genesee, F. (1987). *Learning through Two Languages: Studies of Immersion and Bilingual Education*. Cambridge: Newbury House.
- Genesee, F. (1992). Second/foreign language immersion and at-risk English-speaking children. *Foreign Language Annals*, 25, 199-213.
- Grabe, W. and Stoller, F. L. (1997). Content-based instruction: Research foundations. In M.A. Snow and D.M. Brinton (eds.), *The Content-based Classroom: Perspectives on Integrating Language and Content*. New York: Longman.
- Graddol, D. (2006). *English Next* (vol 62). London: The British Council.
- Grenfell, M. (eds.) (2002). *Modern Languages across the Curriculum*. Routledge Palmer. London.

- Haenni Hoti, A., Heinsmann, S., and Müller, M. (2009). I can you help? Assessing Speaking Skills and Interaction Strategies of Young Learners. In M. Nikolov (ed), *The age factor and early language learning*. Berlin: Mouton de Gruyter. 119-140.
- Halbach, A. (2009). The primary school teacher and the challenges of bilingual education. In E.Dafouz (ed.), *CLIL across educational levels. Experiences from primary, secondary and tertiary levels*. Madrid: Santillana. 19-26.
- Halbach, A. (2010). From the Classroom to University and Back: Teacher Training for CLIL in Spain at the University de Alcalá. In D. Lasagabaster and Y. Ruiz de Zarobe (eds.), *CLIL in Spain: Implementation, Results and Teacher Training*. Newcastle upon Tyne: Cambridge Scholars Publishing, 243-256.
- Halliwell, S. (1992). *Teaching English in the Primary Classroom*. London: Longman.
- Hardi, J. (2004). The attitude and motivation of primary school children for learning English as a foreign language. *Magyar Pedagogia*, 104 (2): 220-242.
- Harley, B. and Swain, M. (1984). The Interlanguage of immersion students and its implication for Second Language Teaching. In A. Davies, C. Cripser and A.P.R. Howatt (eds.), *Interlanguage*. Edinburgh: Edinburgh University Press.
- Harley, B. and Wang, W. (1997). The critical period hypothesis: Where are we now? In A. de Groot and J.Kroll (eds.), *Tutorials in bilingualism: Psycholinguistic Perspectives*. London: Longman. 19-51.
- Hasselgreen A. (2003). The Bergen ‘Can do’ project, ECML/Council of Europe, Graz.
- Hasselgreen A. (2005). Assessing the language of young learners. *Language Testing*, 22 (3): 337-354.
- Hasselgreen, A., Kalédaité, V., Martín, N., and Pižorn, K. (2011). *Assessment of young learner literacy linked to the Common European Framework of Reference for Languages*. Council of Europe Publications.
- Heining-Boyton, A. and Haytema, T. (2007). A ten-year chronicle of student attitudes toward foreign language in the elementary school. *The Modern Language Journal* 91 (2): 149-168.
- Hellekjaer, G. (2004). Unprepared for English Medium Instruction: A critical look at beginner students. In R Wilkinson (ed), *Integrating Content and Language. Meeting the Challenge of a Multilingual Higher Education*. Maastricht University. 97-108.
- Hellekjaer, G. (2010). Language matters: Assessing Lecture Comprehension in Norwegian English-medium Higher education. In C. Dalton-Puffer, T. Nikula and U.Smit (eds.), *Language use and language learning in CLIL classrooms*. Amsterdam: John Benjamins. 233-258.
- Honeyfield, J. (1977). Simplification. *TESOL Quarterly*, 11, 431-440.
- Housen, A. and Kuiken, F. (2009). Complexity, accuracy, and fluency in second language acquisition. *Applied Linguistics*, 30 (4): 461-473.
- Housen, A. and Pierrard, M (eds.) (2005). *Investigations in Instructed Second Language Acquisition*. New York: Walter de Gruyter.

- Housen, A., Schoonjans, E., Janssens, S., Welcomme, A., Schoonheere, E., and Pierrard, M. (2011). Conceptualizing and measuring the impact of contextual factors in instructed SLA: the role of language prominence. *IRAL* 49 (2): 83-112.
- Hughes, A. (2011). Teaching reading in English as a foreign language to young learners: A global reflection. *What research has to say about reading instruction*. In S.J. Samuels and A.E. Farstrup (eds.), *What research has to say about reading instruction*. Newark, DE: International Reading Association. 315-358.
- Hulstijn, J. (1990). A comparison between the information-processing and the analysis/control approaches to language learning. *Applied Linguistics*, 11: 30-45.
- Hunt, K. (1965). *Gramatical Structures Written at Three Grade Levels*. Urbana, IL: The National Council of Teachers of English.
- Hutchinson, T. and Waters, A. (1987). *English for Specific Purposes: A Learning centred approach*. Cambridge: Cambridge University Press.
- Hüttner, J. and Rieder-Bünemann, A. (2007). The effect of CLIL instruction on children's narrative competence. *Vienna English Working Papers*, 16,(3): 20-27.
- Hymes, D. (1972). *Reinventing anthropology*. New York: Random House.
- Inbar-Lourie, O. (2008). Assessment culture. In: E. Shohamy (ed.), *Language Testing and Assessment*. In Hornberger, N. *Encyclopedia of language and education* (2nd ed). New York: Springer Science and Business Media, Inc. 285-300
- Inbar-Lourie, O. (2010). English only? The linguistic choices of teachers of young EFL learners. *International Journal of Bilingualism*, 14 (3): 351-367.
- Inbar-Lourie, O. and Shohamy, E. (2009). Assessing young language learners: What is the construct? In M. Nikolov (ed.), *The age factor and early language learning*. Berlin: Mouton de Gruyter. 83-96.
- Infante, D., Benvenuto, G., and Lastrucci, E. (2008). Integrating content and language at primary school in Italy: Ongoing experimental research. *International CLIL Reserach Journal* 1(1):74-82.
- Infante, D., Benvenuto, G., and Lastrucci E. (2009). The effects of CLIL from the perspective of experienced teachers. In D. Marsh, P. Mehisto, D. Wolff, R. Aliaga, T. Asikainen, and M.J. Frigols-Martin (eds.), *CLIL practice: Perspectives from the field*. S. Hugues and G. Langé, Finland: University of Jyväskylä. 156-163.
- Ioannou-Georgiou, S. and Pavlou, P. (2003). *Assessing young learners*. Oxford: Oxford University Press.
- Jäppinen, A. K. (2005). Thinking and content learning of mathematics and science as cognitional development in content and language integrated learning (CLIL):Teaching Through a Foreign Language in Finland. *Language and Education*, 19 (2):147-168.
- Jäppinen, A.K. (2006). CLIL and future learning. In S. Bjorklun, and K. Mard-Miettinen (eds.), *Exploring dual-focused education. Integrating language and content for individual and societal needs*. Vaasan Yliopiston Julkaisuja.
- Jarvinen, H.M. (1999). Second Language acquisition through CLIL at primary school level. In J. Masih (ed.), *Learning through a foreign language: models, methods and outcomes*. London, CILT. 72-80.

- Jarvinen, H.M. (2005). Language learning in Content-based instruction. In A. Housen and M. Pierrard (eds.), *Investigations in second language acquisition*. Berlin: Mouton de Gruyter. 433-56.
- Jexenflcker, S. and Dalton-Puffer, C. (2010). The CLIL differential: Comparing the writing of CLIL and non-CLIL students in higher colleges of technology. In C. Dalton-Puffer, T. Nikula and U. Smit (eds.), *Language Use and Language Learning in CLIL Classrooms*. Amsterdam: John Benjamins. 169-89.
- Jiménez-Catalán, R.M. and Ojeda, J. (2010). Girls' and boys' lexical availability in English as a Foreign Language. *International Journal of Applied Linguistics*, 158: 57-76.
- Jiménez-Catalán, R.M. and Ruiz de Zarobe, Y. (2009). The receptive vocabulary of EFL learners in two instructional contexts: CLIL versus non-CLIL instruction. In Y. Ruiz de Zarobe and R.M. Jiménez-Catalán (eds.), *Content and Language Integrated Learning. Evidence from Research in Europe*. Bristol: Multilingual Matters. 81-92.
- Jiménez Catalan, R.M., Ruiz de Zarobe, Y., and Cenoz, J. (2006). Vocabulary profiles of English Foreign Language Learners in English as a subject and as a Vehicular Language. *Vienna English Language Papers*, 15 (3): 23-27.
- Johnson, R. (1996). *Language teaching and skill learning*. Oxford: Blackwell.
- Johnson, J. and Newport, E. (1989). Critical Period Effects in Second Language Learning: the influence of maturational state on the acquisition of English as a second language. *Cognitive Psychology*, 21, 60-99.
- Johnson, K. and Swain, M. (1997). *Immersion Education: International perspectives*. New York. Cambridge University Press.
- Johnstone, R. (1994). *Teaching MFL at Primary School: approaches and implications*, Fort William, SCRE.
- Johnstone, R. (2009). An early start: What are the key conditions for generalized success? In J. Enever, J. Moon and U. Raman (eds.), *Young Learner English Language Policy and implementations: international perspective*. Reading, UK: Garnet Education Publishing. 31-42.
- Juan-Garau, M. (2010). Oral Fluency Development in Secondary Education CLIL Learners. *Vienna English Working Papers* 19 (3): 42-48.
- Juffs, A. and Harrington, M. (1995). Parsing effects in second language sentence processing. *Studies in Second Language Acquisition*, 17 (4): 483-516.
- Kim, Y.-J. (1985). *Variables in reading comprehension*. MA thesis. Unpublished manuscript. Seoul National University, Korea.
- Klesmer, H. (1994). Assessment and teacher perceptions of ESL student achievement. *English Quarterly*, 26 (3): 5-7.
- Knight, M. (1990). Third and final report on English at Ebersteinska skölan. In Näslund, L (ed) Engelskpraktiska tekniska studievägar vid Ebersteinskaskölan i Norköping. Unvärdering av en Forsöksverksamhet. Stockholm. Skolöverstyrelsen.
- Kovács, J. (2005) CLIL in Hungary. In The CLIL quality matrix. Central workshop report, coordinator D. Marsh. (Available at <http://www.ecml.at/mtp2/CLILmatrix.pdf>.)

- Krashen, S. (1997). The comprehension hypothesis: Recent evidence. *English Teachers' Journal (Israel)*, 51, 17-29.
- Krahnke, K. (1987). *Approaches to Syllabus Design for Foreign Language Teaching* New York: Prentice Hall.
- Krashen, S.D. (1982). *Principles and Practice in Second Language Learning and Acquisition*. New York: Pergamon.
- Krashen, S.D. (1985). *The Input Hypothesis: Issues and Implications*. London: Longman..
- Krashen, S., Scarcella, R., and Long, M. (1982). Age, rate and eventual attainment in second language acquisition. *TESOL Quarterly* 9 (13): 573-582.
- Kroll, B. (1998). Assessing Writing Abilities. *Annual Review of Applied Linguistics*, 18 (1): 219-240.
- Lambert, W. and Tucker, R. (1972). *Bilingual Education of Children: the St Lambert experiment*. MA: Newbury House Publishers.
- Langé, G. (2007). Postscript to CLIL 2006 and Future Action. In D. Marsh and D. Wolff (eds.), *Diverse Contexts, Converging Goals: CLIL in Europe*. Peter Lang: Frankfurt am Main.
- Langé, G. (2007). *Windows on CLIL. Content and Language Integrated Learning in the European Spotlight*. In A. Maljers, D. Marsh and D. Wolff (eds.). European Platform for Dutch Education, Alkmaar.
- Lapkin, S., Swain, M., and Shapson, S. (1990). French Immersion Research Agenda for the 90's. *Canadian Modern Language Review*, 46 (4): 638-74.
- Larsen-Freeman, D. (2001). Individual cognitive/affective learner contributions and differential success in second language acquisition. In M. Breen *Learner contributions to language learning*. Harlow: Longman. 12-24.
- Larson-Hall, J. (2008). Weighing the benefits of studying a foreign language at a younger starting age in a minimal input situation. *Second Language Research* 24 (1): 35-63.
- Lasagabaster, D. (2003). Attitudes towards English in the Basque Autonomous Community. *World Englishes* 22 (4): 585-597.
- Lasagabaster, D. (2008). Foreign language competence in content and language integrated courses. *The Open Applied Linguistics Journal* 1, 31-42.
- Lasagabaster, D. and Ruiz de Zarobe, Y. (2010). Ways forward in CLIL: Provision issues and future planning. In D. Lasagabaster and Y. Ruiz de Zarobe (eds.), *CLIL in Spain: Implementation, results and Teacher training*. Newcastle upon Tyne: Cambridge Scholars Publishing, 278-95.
- Lasagabaster, D. and Sierra, J.M. (2009). Language attitudes in CLIL and traditional EFL classes, *International CLIL Research Journal*, 1(2): 4-17.
- Lasagabaster, D. and Sierra, J.M. (2010). Immersion and CLIL in English: more differences than similarities. *ELT Journal*, 64 (4): 367-375.
- Lazaruk, W. 2007, Linguistic , academic and cognitive benefits of French Immersion. *The Canadian Modern Language Review*, 63 (5): 605-628.

- Lee, J. H. (1986). *A study on the influence of vocabulary, syntax, and discourse structure upon reading comprehension*. MA thesis. Unpublished manuscript., Seoul National University, Korea.
- Lenneberg, E.H. (1967). *Biological Foundations of Language*. New York: Wiley.
- Lennon, P. (1990). Investigating Fluency in EFL: A quantitative approach. *Language Learning*, 40, 387-417.
- Lightbown, P. and Spada, N. (2006). *How Languages are Learned*. Oxford: Oxford University Press.
- Llach, M. P. A. and Gallego, M. T. (2009). Examining the Relationship between Receptive Vocabulary Size and Written Skills of Primary School Learners/Examen de la relación entre el conocimiento de vocabulario receptivo y las destrezas escritas de los alumnos de primaria. *Atlantis*, 31, 129-147.
- Llinares, A. and Dafouz, E. (2010). Content and Language Integrated Programs in the Madrid Region: Overview and research findings. In D. Lasagabaster and Y. Ruíz de Zarobe (eds.), *CLIL in Spain: Implementation, Results and Teacher Training*. Newcastle: Cambridge Scholars, 95-113.
- Llinares, A. and Whittaker, R. (2006). Linguistic analysis of secondary school students' oral and written production in CLIL contexts: Studying social science in English. *Vienna English Working Papers* 15, 28-32.
- Llinares, A. and Whittaker, R. (2010). Writing and speaking in the history class: A comparative analysis of CLIL and first language contexts. In C. Dalton-Puffer, T. Nikula and U. Smit (eds.), *Language Use and language learning in CLIL classrooms*. Amsterdam: John Benjamins Publishing Company, 125-143.
- Long, M. (1983). Native speaker/non-native speaker conversation and the negotiation of comprehensible input. *Applied linguistics*, 4 (2): 126-141.
- Long, M. (1990) Maturation constraints on language development. *Studies in Second Language Acquisition* 12(3): 251-285.
- Long, M. (1991). Focus on Form: A Design Feature in Language Teaching Methodology. In K de Bot, R. Ginsberg & C Kramsch (eds.) *Foreign Language Research in Cross-cultural Perspective*. Amsterdam: John Benjamin Publishing Company, 39-52
- Long, M. (1996). The Role of Linguistic Environment in Second Language Acquisition. In Ritchie W.C. & Bathia, T.K.(eds.). *The Handbook of Acquisition*. San Diego: Academic Press, 413-468.
- Long, M. (1997). Construct validity in SLA: A response to Furth and Wagner. In *Modern Language Journal*, 81, (3): 318-323.
- Long, M. (2005). Problems with supposed counter-evidence to the Critical Period Hypothesis. *IRAL* 43, (4): 287-318.
- Lopriore, L. (2009). *Development of young learners perception of foreign language learning and teaching*. Paper presented at the American Association of Applied Linguistics Conference, Denver, Colorado.
- Loranc-Paszylk, B. (2009). Integrating reading and writing into the context of CLIL classroom: Some practical solutions. *International CLIL Research Journal*, 1(2): 47-53.

- Lorenzo, F. 2007. The sociolinguistics of CLIL: Language planning and language change in 21st century Europe. *Revista Española de Linguística Aplicada*, 1: 27_38.
- Lorenzo, F. (2008). Instructional discourse in bilingual settings: an empirical study of linguistic adjustments in CLIL. *Language Learning Journal* 36,(1): 21-23.
- Lorenzo, F., Casal, S., and Moore, P. (2010). The effects of content and language integrated learning in European education: Key findings from the Andalusian Bilingual Sections Evaluation Project. *Applied Linguistics* 31, (3): 418-42.
- Luczywek, I. (2009). Three models of integrating school subjects in Poland. In . D. Marsh, P. Mehisto, D. Wolff, R. Aliaga, T. Asikainen, M.J. Frigols, S. Hugues, and G. Langé (eds.) *CLIL practice. Perspectives from the field*. Finland: University of Jyväskylä. 172-178.
- Lundberg, G. (2007). Developing teachers of young learners: In-service for educational change and improvement. In M. Nikolov, J. Mihaljevic Djigunovic, M. Mattheoudakis, G. Lundberg, and T. Flanagan (eds.), *Teaching Modern Languages to Young Learners: Teachers, Curricula and Materials*. Strasbourg: Council of Europe. 21-34
- Lyster, R. (1987).Speaking Immersion. *The Canadian Modern Language Review*, 43 (4) 701-717.
- Lyster, R. (2004). Research on form-focused instruction in immersion classrooms: Implications for theory and practice. *Journal of French Language Studies*, 14 (3): 321-341.
- Lyster, R. (2007). *Learning and Teaching Languages through Content: a Counterbalanced Approach*. Amsterdam: John Benjamins Publishing Company.
- Lyster, R. and Ranta, L. (1997) Corrective feedback and learner uptake. *Studies in Second Language Acquisition*, 19, 37-66.
- Mackey, A. (1999). Input, interaction and second language development. *Studies in Second Language Acquisition*, 21, 557-581.
- MacWhinney, B. (2005). A unified model of language development. In J.F.Krol and A.M.B De Groot (eds.), *Handbook of Bilingualism: Psycholinguistic Approaches*. Oxford: Oxford University Press.49-67
- Maljers, A., Marsh, D., and Wolff, D. (eds.) (2007). *Windows on CLIL. Content and Language Integrated Learning in the European Spotlight*. European Platform for Dutch Education.
- Marinova-Todd, S.H., Marshall, B., and Snow, C.E. (2000). Three misconceptions about age and L2 learning. *TESOL Quarterly*, 34(1): 9-34.
- Marschollek, A. (2002). *Cognitive and affective flexibility in foreign languages*. Münster: Lit.
- Marsh, D. (1994). Bilingual Education and Content and Language Integrated Learning *International Association for Cross-Cultural Communication, Language Teaching in the Member States of the European Union*. University of Sorbonne, Paris.
- Marsh, D. (2000). *Using Languages to Learn and Learning to Use languages: An Introduction to CLIL for Parents and Young People*. Jyväskylä, Finland: UniCOM, University of Jyväskylä on behalf of TIE-CLIL. (Available at <http://www.clilcompendium.com/luk.pdf>).

- Marsh, D. (2012). *Quality interfaces: Examining evidence and exploring solutions in CLIL*. Eichstätt: Eichstaett Academic Press.
- Marsh, D. and Langé, G. (eds.). (1999). *Implementing Content and Language Integrated Learning*. A Research-driven TIE_CLIL Foundation Course Reader. Jyväskylä, Finland: Continuing Education Centre, University of Jyväskylä on behalf of TIE-CLIL (European Lingua Project).
- Marsh, D., Maljers, A., and Hartiala, A.(eds.) (2001). *Profiling European CLIL Classrooms*. UniCOM. University of Jyväskylä.
- Marsh, D. and Wolff, D. (eds.) (2007). *Diverse Contexts. Converging Goals. CLIL in Europe*. Peter Lang. Frankfurt.
- Martin, C. (2000). Modern foreign languages at primary school: A three-pronged approach? *Language Learning Journal*, 22 (1): 5-10.
- Matteoudakis, M., Dvorakova, K., and Lang, K. (2007). Story-based language teaching: An experimental study on the implementation of a module in three European countries. In M. Nikolov et al. (eds.), *The TeMoLaYoLe book: Teaching modern languages to young learners: teachers, curricula and materials*. Strasbourg/Graz: Council of Europe/European Centre for Modern Languages. 59-76
- Mckay, P. (2006). *Assessing Young Language Learners*. Cambridge: Cambridge University Press.
- Mehisto, P. (2007). What a School Needs to Consider before Launching a CLIL. In D. Mars and D. Wolff (eds.), *Diverse Contexts- Converging Goals. CLIL in Europe*. Frankfurt: Peter Lang. 61-78
- Mehisto, P., and Asser, H. (2007). Stakeholder perspectives: CLIL programme management in Estonia. *International Journal of Bilingual Education and Bilingualism*, 10, (5): 683-701.
- Mehisto, P., Frigols, M. J., and Marsh. D.(2008). *Uncovering CLIL: Content and Language Integrated Learning and Multilingual Education*. Oxford:Macmillan.
- Merisuo-Storm,T. (2007). Pupils' attitudes towards foreign-language learning and the development of literacy skills in bilingual education. *Teaching and Teacher Education*,23, (2): 226-235.
- Met, M. (1991). Learning Language through Content: Learning Content through Language. *Foreign Language Annals*, 24, (4): 281-95.
- Met, M. (1998). Curriculum decision-making in content-based language teaching. In J. Cenoz & Genesee (eds.), *Beyond bilingualism: Multilingualism and multilingual education*. Clevedon, UK: Multilingual Matters.35-63.
- Met, M. (1999). Content based instruction: *Defining Terms and Making Decisions*. College Park, M.D.: National Foreign Language Center.
- Met, M., and Lorenz, E. (1997). Lessons from U.S. immersion programs: Two decades of experience. In R. Johnson & M. Swain (eds.), *Immersion education: International perspectives*. Cambridge: Cambridge University Press. 243-264.
- Mewald, C. (2007). A Comparison of Oral Language Performance of Learners in CLIL and in Mainstream Classes at Lower Secondary Level in Austria .In C. Dalton-Puffer, D.

- Kastovsky, N. Ritt and Schendl, H. (eds.), *Syntax, Style and Grammatical Norms*. Frankfurt: Peter Lang. 139-173.
- Miret, M. (2009). CLIL effects on writing performance and overall proficiency. MA thesis. Unpublished manuscript. Barcelona. University of Barcelona
- Mitchell, R. (2003). Rethinking the concept of progression in the national curriculum for modern foreign languages: a research perspective. *Language Learning Journal*, 27(1): 15-23.
- Mohan, B. (1986). *Language and Content*. Reading, Massachusetts: Addison-Wesley Publishing Company.
- Mohan, B., Leung, C., and Davison, C. (Eds.). (2001). *English as a Second Language in the Mainstream. Teaching, Learning and Identity*. Essex, England: Longman.
- Mohan, B. and Van Naersenn, M. (1997). Understanding Cause-Effect: Learning through Language. *Forum*, 35 (4): 22-40.
- Moon, J. (2000). *Children Learning English*. London: Macmillan Heinemann.
- Moon, J. (2009). The teacher factor in early foreign language learning programmes: The case of Vietnam. In M. Nikolov (ed). *The age factor and early language learning*. Berlin: Mouton de Gruyter. 311-336
- Moon, J. and Nikolov, M (Eds.) (2000). *Research into teaching English to young learners*. Pécs: University Press Pécs.
- Muñoz, C. (ed) (2006). The Effects of Age on Foreign Language Learning. In C. Muñoz (ed), *Age and the Rate of Foreign Language Learning*. Clevedon: Multilingual Matters. 1-41
- Muñoz, C. (2007). Age-related differences and second language learning practice. In R. Dekeyser (ed), *Practising for second language use: Perspectives from applied linguistics and cognitive psychology*. Cambridge, UK: Cambridge University Press. 229-255.
- Muñoz, C. and Nussbaum, L. (1997). Les enjeux linguistiques dans l'éducation en Espagne. In *AILE: Acquisition and Interaction en Langue Étrangère* N° 10, *Appropriation the Langues en Situation de Contact*.
- Muñoz, C. and Tragant, E. (2001). Attitudes and Motivation towards foreign language learning. In S.H. Foster-Cohen and A. Nizgorodsew (eds.), *Eurosla Yearbook 2000*. Amsterdam: John Benjamins Publishing Company. 211-224
- Nagy, K. (2009). English language teaching in Hungarian primary schools with special reference to the teacher's mother tongue use. MA thesis. Unpublished manuscript University of Stirling, Scotland.
- Navés, T. (2002). Successful CLIL Programmes. In Navés, T. Muñoz, and C. Pavesi, M. *Second Language Acquisition for CLIL. The CLIL Professional Development Course*. 93-102.
- Navés, T. (2009). Effective Content and Language Integrated Programmes. In Y. Ruiz de Zarobe and R. Jimenez Catalán (eds.), *Content and Language Integrated Learning Evidence from Research in Europe*. Clevedon: Multilingual Matters. 22-40.

- Navés, T. (2011). The promising benefits of integrating content and language for EFL writing and overall EFL proficiency. In Y. Ruiz de Zarobe, J. M. Sierra and F. Gallardo del Puerto (eds.). *Content and Foreign Language Integrated Learning. Contributions to Multilingualism in European Contexts*. Bern: Peter Lang.
- Navés, T. and Muñoz, C. (1999). CLIL experiences in Spain. In Marsh, D and Langé, G (eds.), *Implementing Content and Language Integrated Learning. A Research-driven TIE_CLIL Foundation Course Reader*. Jyväskylä, Finland: Continuing Education Centre, University of Jyväskylä on behalf of TIE-CLIL (European Lingua Project). 131–144.
- Navés, T. and Muñoz Lahoz, C. (1999). Implementation of CLIL in Spain. In D. Marsh & G. Langé (Eds.) *Implementing Content and Language Integrated Learning. A Research-driven TIE_CLIL Foundation Course Reader*. Jyväskylä, Finland: Continuing Education Centre, University of Jyväskylä on behalf of TIE-CLIL (European Lingua Project). 145-158.
- Navés, T. and Victori, M. (2010). CLIL in Catalonia: An Overview of Research studies. In Y. Ruiz de Zarobe & D. Lasagabaster (eds.), *CLIL in Spain: Implementation, Results and Teacher Training*. Cambridge Scholars Publishing. 30-54
- Nemser, W. (1971). Approximate systems of foreign language learners. *International Review of Applied Linguistics*, 9 (2): 115-123.
- Nikolov, M. (1999). Why do we learn English? Because the teacher is short. A study of Hungarian children's foreign language learning motivation. *Language Teaching Research*, 3, (1): 33-56.
- Nikolov, M. (2008). Primary school means methodology! An empirical study of lower primary EFL classes. *Modern Nyelvoktatás*, 10, (1): 3-19.
- Nikolov, M. (ed.) (2009). *Early learning of modern foreign languages: processes and outcomes*. Bristol: Multilingual Matters.
- Nikolov, M. (2011). The age factor in context. In M. Nikolov (ed), *The Age Factor and Early Language Learning*. Berlin: Mouton de Gruyter.
- Nikolov, M. and Djigunovic, J.M. (2006). Recent research on age, second language acquisition and early foreign language learning. *Annual Review of AppliedLinguistics*, 26 (1): 234-260.
- Nikolov, M. and Djigunović, J.M. (2011). All shades of every color: An overview of early teaching and learning of foreign languages. *Annual Review of Applied Linguistics*, 31(1): 95-119.
- Nikula, T. (2005). English as an object and tool of study: interactional effects and pragmatic implications. *Linguistics and Education*, 16 (1): 27-58.
- Nikula, T. and Marsh, D. (1997). *Language and Content Integrated Learning in the Primary and Secondary sector*. Helsinki: National Board of Education.
- Novotná, J., Hadj-Mousová, Z., and Hofmannová, M. (2001). Teacher training for CLIL. Competences of a CLIL teacher. In *Proceedings SEMT 01*. J. Novotná, M. Hejný. (eds.).Praha: Univerzita Karlova, Pedagogická fakulta. 122-126.

- Novotná, J. and Hofmannová, M. (2007). Czech Republic. In A. Maljiers, D. Marsh and D. Wolff (eds.), *Windows on CLIL*. GRAZ: ECML. 39-51.
- Nunan, D. (1995). *Language Teaching Methodology. A textbook for Teachers*. New York: Phoenix. ELT.
- Ojeda Alba, J. (2009). Themes and vocabulary in CLIL and non-CLIL instruction. In Y. Ruiz de Zarobe and R.M. Jiménez-Catalán (eds.), *Content and Language Integrated Learning. Evidence from Research in Europe*. Bristol: Multilingual Matters. 81-92.
- Olaizola, I.V. and Mayo, M. D. P. G. (2009). Tense and agreement morphology in the interlanguage of Basque/Spanish bilinguals: CLIL versus non-CLIL. In *Content and language integrated learning: evidence from research in Europe*. Bristol: Multilingual Matters. 41-57
- Oller, D. K. and Pearson, B. Z. (2002). Assessing the effects of bilingualism: A background. In D. K. Oller and R. E. Eilers (eds.), *Language and literacy in bilingual children*. Tonawanda, NY: Multilingual Matters Ltd. 3-21.
- Pallotti, G. (2009). CAF: Defining, Refining and Differentiating Constructs. *Applied Linguistics*, 30 (4): 590-601.
- Patkowski, M. (1980). The sensitive period for the acquisition of syntax in a second language. *Language Learning* 30, 449-472.
- Pavesi, M., Bertochi, D., Hofmannová, M., Kazianka, M., and Langé, D. (2001). Teaching Through a Foreign Language. A guide for teachers and schools to using Foreign Language in Content Teaching. In D. Langé (Ed.), *Insegnare in una Lingua Straniera. Unterrichten durch eine Fremdsprache. Teaching through a Foreign Language. Enseñar en una Lengua Extranjera. Enseigner dans une Langue Vivante* (pp. 32). Milano, Italy: M.I.U.R Direzione Generale della Lombardia on behalf of TIECLIL.
- Peltzer-Karf, A. and Zangl, R. (1997). Vier Jahre Vienna Bilingual Schooling: Eine Langzeitstudie. Vienna; Bundesministerium für Unterricht und kulturelle Angelegenheiten., Abteilung 1. (1).
- Pena Díaz, C. and Porto Requejo, M.D. (2008). Teachers' beliefs in a CLIL education project. *Porta Linguarum* 10, 151-161.
- Perani, D. and Abulatevi, J. (2005). The neural basis of first and second language process. *Current Opinion in Neurobiology*, 15, 202-206.
- Pérez-Cañado, M.L. (2012). CLIL research in Europe: Past, present and future. *International Journal of Bilingual Education and Bilingualism*, 15 (3): 315-341.
- Pérez-Vidal, C. (1999). Teacher Education: the Path towards CLIL. In D Marsh and Marsland (eds.), *CLIL Initiatives for the Millenium*, Report on the CEILINK Think-Tank16-18 Jyväskylä: University of Jyväskylä.
- Pérez-Vidal, C. (2000). Multilingual education: Problems and Solutions in Catalonia. In Van de Craen, P.& Pérez-Vidal, C. (eds.), *The Multilingual Challenge/Le défi Multilingue*. Barcelona: Printulibro.
- Pérez-Vidal, C. (2001). Dimensiones del aprendizaje de contenidos curriculares y lenguas extranjeras. In D: Marsh, A. Maljiers and A. Hartiala (eds.), *The CLIL compendium. Profiling European CLIL classrooms*. Jyväskylä: University of Jyväskylä. 215-253.

- Pérez-Vidal, C. (2007). The Need for Focus on Form in Content and Language Integrated Approaches: An Exploratory Study. *Revista Española de Linguística Aplicada*, 1, 39-54.
- Pérez-Vidal, C. (2008). La enseñanza bilingüe: El enfoque integrado de Contenidos i Lenguas en Europa. *Aula de Innovación Educativa* Editorial Graó.
- Pérez-Vidal, C. and Juan-Garau, M. (2010). To CLIL or not to CLIL? From bilingualism to multilingualism in Catalan/Spanish communities in Spain. In Y. Ruiz de Zarobe and D. Lasagabaster (eds.), *CLIL in Spain: Implementation, Results and Teacher Training*. CUP: Cambridge Scholars. 115-138
- Phillips, S. (1993). *Young Learners*. Oxford: Oxford University Press.
- Piaget, J. (1965). The stages of the intellectual development of the child. *Educational psychology in context: Readings for future teachers*, Thousand Oaks, CA: Sage Publications. 98-106.
- Pienemann, M. (1998). *Language processing and second language development: Processability theory* (Vol. 15). John Benjamins Publishing.
- Pinter, A. (2006). *Teaching young language learners*. Oxford: Oxford University Press.
- PISA Reports available from [http:// www.pisa.oecd.org](http://www.pisa.oecd.org)*
- Pladevall Ballester, E. (forthcoming) CLIL subject selection and young learners' listening comprehension skills. *International Journal of Applied Linguistics*.
- Pladevall Ballester, E. (forthcoming) Exploring primary school CLIL perceptions in Catalonia: students', teachers', head teachers' and parents' opinions and expectations. *International Journal of Bilingual Education and Bilingualism*.
- Pulvermüller, F. and Schumann, J.H. (1994). Neurobiological mechanisms of language acquisition. *Language Learning* 44, 681-734.
- Quezada, M.S., Wiley, T.G., and Ramirez, J.D. (2000). How the reform agenda short changes English learners. *Educational Leadership*, 57 (4): 57-61.
- Ramirez, J. D., Yuen, S., Ramey, D., and Pasta, D. (1991). Longitudinal study of Structured English Immersion Strategy, Early-exit and late-exit Bilingual Education Programs for Language-minority children (Final Report, Vols. 1 & 2). San Mateo, CA: Aguirre International.
- Richards, J. and Rodgers, S. (2001). *Approaches and Methods in Language Teaching*. Cambridge: Cambridge University Press.
- Rixon, S. (1999). *Young learners of English: Some research perspectives*. London: Longman.
- Rixon, S. (2000). Young Learners of English: background and issues. *Modern English Teacher*, 9 (4): 5-10.
- Rixon, S. (2007). Cambridge ESOL YLE tests and children's first steps in reading and writing in English. *Cambridge ESOL Research Notes*, 28: 7-14.
- Robinson, P. (1995). Attention, memory and the noticing hypothesis. *Language Learning*, 46 (2): 283-331.

- Roeper, T. (1996). The role of merger theory and formal features in acquisition. *Language acquisition and language disorders*, 14, 415-450.
- Romu, A., and Sjöberg-Heino, L. (1999). A practitioner's perspective on bilingual teaching in Alhaisten Primary school. In *Learning through a Foreign Language. Models. Methods and outcomes*. J. Masih (ed). London: Centre for Information on Language Teaching and Research. 81-88.
- Ruiz de Zarobe, Y. (2004). Aprendizaje Integrado de Contenidos Curriculares en Inglés como Lengua Extranjera: Diferencias con el Aprendizaje del Inglés como Asignatura. XXV Congreso Anual de AESLA Universidad de Murcia, 19-21 de abril, 2004.
- Ruiz de Zarobe, Y. (2004). Edad y competencia oral en la Adquisición del inglés como tercera lengua. *Review of Applied Linguistics*, 143-144: 121-132. Belgium: University of Leuven.
- Ruiz de Zarobe, Y. (2005). Age and Third Language Production: A longitudinal study. *International Journal of Multilingualism*, 2, (2): 105-113.
- Ruiz de Zarobe, Y. (2007). CLIL in a Bilingual Community: Similarities and Differences with the Learning of English as a Foreign Language *Vienna English Working Papers*. 16, 3.
- Ruiz de Zarobe, Y. (2008). CLIL and Foreign Language Learning: a longitudinal study in the Basque Country. *International CLIL Research Journal*, 1(1): 60-73.
- Ruiz de Zarobe, Y. and Catalán, R. M. J. (eds.). (2009). *Content and language integrated learning: Evidence from research in Europe* (Vol. 41). Multilingual Matters.
- Ruiz de Zarobe, Y. and Lasagabaster, D. (2010). *CLIL in Spain: Implementation, Results and Teacher Training*. Newcastle, UK: Cambridge Scholars Publishing.
- Saville-Troike, M. (2006) *Introducing Second Language Acquisition*. Cambridge: Cambridge University Press.
- Schmidt, R. (1990). The role of consciousness in second language learning. *Applied Linguistics*, 11(2): 129-158.
- Scovel, T. (2000). A critical review of the critical period research. *Annual Review of Applied Linguistics*, 20 (1): 213-223.
- Seikkula-Leino, J. (2007). CLIL learning: achievement levels and affective factors. *Language and Education*, 21(4): 328-341.
- Selinker, L. (1972). Interlanguage. *IRAL*, 10 (1-4), 209-232.
- Senesac, B. V. K. (2002). Two-way bilingual immersion: A portrait of quality schooling. *Bilingual Research Journal*, 26 (1), 85-101.
- Seregély, E. M. (2008). *A comparison of lexical learning in CLIL and traditional EFL classrooms* (Unpublished doctoral dissertation, University of Vienna).
- Serra, C. (2007). Assessing CLIL at Primary School: A longitudinal Study. *International Journal of Bilingual Education and Bilingualism*, 10 (5): 582-602.
- Serra, J.M. (1997). Immersió lingüística, rendiment acadèmic i classe social. *Quaderns per a l'anàlisi*, 9. Barcelona: Horsori.
- Setati, M. (1998). Code-switching in a Senior Primary Class of Second-language Mathematics Learners. *The Learning of Mathematics*, 18 (1): 34-40.

- Sheen, R. (2002). Focus on form and Focus on Forms. *ELT Journal*, 56 (3): 303-305.
- Shohamy, E. and Inbar, O. (2007). Effective teaching of EFL to young learners: Homeroom or FL teachers? Paper presented at TeMoLaYoLe conference, University of Pécs, Hungary.
- Sierra, J. M. and Lasagabaster, D. (eds). (2005). *Multilingüismo, competencia lingüística y nuevas tecnologías. Cuadernos de Educación*. ICE . Universitat de Barcelona.
- Simó, N. and Telford, J. (2012). Vic: The challenges facing schools in a small, newly diverse, Catalan city. *Improving Schools*, 15(3): 211-227.
- Singleton, D. (2005). The Critical Period Hypothesis: A coat of many colours. *International Review of Applied Linguistics in Language Teaching*, 43(4): 269-285.
- Skehan, P. (1998). *A Cognitive Approach to Language Learning*. Oxford: Oxford University Press.
- Skehan, P. and Foster, P. (1999). The influence of task structure and processing conditions on narrative retellings. *Language learning*, 49 (1): 93-120.
- Slattery, M. and Willis, J. (2001). *English for Primary Teachers*. Oxford: Oxford University Press.
- Slavoff, G.R. and Johnson, J. S. (1995). The effects of age on the rate of learning a second language. *Studies in Second Language Acquisition* 17, (1):, 1-16.
- Smith, J. and Patterson, F. (1998). *Positively Bilingual: Classroom Strategies to Promote the Achievement of Bilingual Learners*. Nottingham: Nottingham Education Authority.
- Snow, M. A. (1998). Trends and Issues in Content-Based Instruction. *Annual Review of Applied Linguistics*, 18, 243-267.
- Snow, M., Met, M., and Genessee, F. (1989). A Conceptual Framework for the Integration of Language and Content in Second/Foreign Language Instruction. *TESOL Quarterly*, 23 (2): 201-217.
- Södergard, M. (2006). From Kindergarten to Grade 6. The immersion experience from the pupils' point of view. In *Exploring dual- focussed education. Integrating language and content for individual and societal needs*. S. Björklund and K. Mard-Miettinen (ed). Berstrom and Södergard.
- Spanos, G. (1989). On the Integration of Language and Content Instruction. *Annual Review of Applied Linguistics* 10 (1): 227-240.
- Stern, H.H. (1983). *Fundamental Concepts of Language Teaching*. Oxford: Oxford University Press.
- Stotz, D. and Meuter, T (2003). Embedded English: Integrating content and language in a Swiss primary school project. In *Le Bulletin Suisse de Linguistique Appliquée* 77, 83-101.
- Swain, M. (1984). A Review of Immersion Education in Canada: Research and Evaluation Studies. In *Studies on immersion education. ELT Documents*, 119: 35-51.
- Swain, M. (1985). Communicative competence: Some roles of comprehensible input and comprehensible output in its development. *Input in second language acquisition*, 15, 165-179.

- Swain, M. (1988.) Manipulating and complementing Content teaching to maximize Second language Learning, *TESL Canada Journal*, 6 (1): 68-83.
- Swain, M. (1995). Three functions of output in second language learning. In G. Cook and B. Seidlhofer (eds.), *Principal and practice in applied linguistics: studies in honour of H.G.Widdowson*. Oxford: Oxford University Press. 125-144.
- Swain, M. (1998). Focus on form through conscious reflection. In C. Doughty and J. Williams (eds.), *Focus on form in classroom second language acquisition* Cambridge, UK: Cambridge University Press. 64-81.
- Swain, M. (2000). French Immersion Research in Canada: Recent Contributions to SLA and Applied Linguistics. *Annual Review of Applied Linguistics*, 20, 199-212.
- Swain, M. and Carroll, S. (1997). The immersion observation study. In B. Harley, P. Allen, J. Cummins, and M. Swain (eds.), *Development of bilingual proficiency. Final report. Volume II: Classroom treatment*. Toronto, ON: Modern Language Centre OISE/UT. 190-316.
- Swain, M. and Lapkin, S. (1995). Problems in output and cognitive processes they generate: a step towards second language learning. *Applied Linguistics*, 16, 370-391.
- Swain, M, and Lapkin, S. (1998). Interaction and Second Language Learning: two adolescent French immersion students working together, *The Modern Language Journal*, 82 (3): 320-337.
- Swain, M. and Lapkin, S. (2005). The evolving socio-political context of immersion education in Canada: some implications for programme development. *International Journal of Applied Linguistics*, 15 (2): 169-186.
- Szpotowicz, M., Mihaljevic Djigunovic, J. and Enever, J. (2009). Early language learning in Europe: A multinational, longitudinal study. In J. Enever, J. Moon and U. Raman (eds.) *Young Learner English Language Policy and Implementation: International perspectives*. Reading, UK: Garnet Publishing. 141-147.
- Takala, S. (2002) Positioning CLIL in the Wider Context. In Marsh, D. (ed), CLIL/EMILE. The European Dimension. UniCOM Continuing Education Centre, University of Jyväskylä, Finland. 40-42.
- Taylor, L. and Saville, N (2002).Developing English Language Tests for young learners. In University of Cambridge Local Examinations Syndicate. *Research Notes 7* (pp 2-5). Cambridge University Press.
- Thomas, W. P. and Collier, V. (1997). School Effectiveness for Language Minority Students'. *In NCBE Resource Collection Series*, 9. Washington, DC: National Clearinghouse for Bilingual Education.
- Thomas, W. P. and Collier, V. P. (2002). A national study of school effectiveness for language minority students' long-term academic achievement. (Available at www.crede.ucsc.edu/research/llaa/l1_final.html).
- Ting, T. (2007). Insights from Italian CLIL-Science Classrooms: Refining Objectives,Constructing Knowledge and Transforming FL-Leraners into FL-users. In *Vienna English Working Papers*, 16, (3): 60.

- Tjierandsen, E. (1995). Stories you make think is boring. Can she make funny. Sentence complexity in the written work of Norwegian 8th graders. MA thesis, University of Bergen.
- Torras, M.R., Navés, T., Celaya, M.L. and Pérez-Vidal, C. (2006). *Age and IL Development in Writing*. In C. Muñoz (Ed), *Age and the Rate of Foreign Language Learning*. Clevedon: Multilingual Matters. 156-183.
- Tragant, E. (2006). *Language Learning, Motivation and Age*. In Muñoz, C. (ed) *Age and the rate of Foreign Language Learning*. Clevedon: Multilingual Matters. 237-267
- Ullmann, M. (1999). History and Geography through French: CLIL in a UK secondary school
In J Masih (ed.), *Learning through a foreign language. Models, methods and outcomes*, London: Centre for Information on Language Teaching and Research. 96-105.
- Ullman, M. (2005). A cognitive neuroscience perspective on second language acquisition: The declarative/procedural model. In C. Sanz (ed), *Mind and context in adult second language acquisition*. Methods, theory and Practice. Washington: Georgetown University Press. 141-178.
- Vallbona, A. (2009). Implementing CLIL methodology in two courses of primary education in a state-funded school: A case study. MA thesis. Unpublished manuscript. Bellaterra, UAB.
- Vallbona, A. (2011). Implementing CLIL in the Primary Classroom. Research and Future Challenges .In Urmeneta, C. E., Evnitskaya, N., Moore, E., and Patiño, A. Implementing CLIL in the primary classroom: Results and future challenges. *AICLE–CLIL–EMILE: Educació plurilingüe: Experiències, research & polítiques 2*. (151-157).
- Van de Craen, P. Ceuleers, E., & Mondt, K. (2007). Cognitive development and bilingualism in primary schools: teaching maths in a CLIL environment. *Diverse Contexts–Converging Goals. CLIL in Europe. Frankfurt am Main: Peter Lang*, 185-200.
- Van de Craen, P. and Pérez-Vidal, C. (2001). *The Multilingual Challenge/ Le Défi Multilingue. Final Report 2000*. Barcelona: European Commission. Pompeu Fabra University.
- Van de Craen, P., Mondt, K., Allain, L and Gao, Y. (2007). Why and How CLIL works. An Outline for a CLIL Theory. *Vienna English Working Papers*, 16 (3): 70-78.
- Van de Craen, P., Mondt, K., Allain, L., and Gao, Y. (2007). Cognitive Development Bilingualism in Primary Schools: Teaching Maths in a CLIL Environment. In D. Marsh and D. Wolff (eds.), *Diverse Contexts-Converging Goals: CLIL in Europe*. Frankfurt am Main: Peter Lang. 185-200.
- VanPatten, B. (1990). Attending to form and content in the input. In *Studies in Second Language Acquisition*, 12, 287-301.
- VanPatten, B. (1996). *Input processing and grammar instruction in second language acquisition*. Greenwood Publishing Group.
- VanPatten, B. (2003). *From Input to Output. A teacher's Guide to Second Language Acquisition*. Boston: McGraw Hill.
- Varkuti, A (2010), Linguistic Benefits of the CLIL Approach: Measuring Linguistic Competences, *International CLIL Research Journal*, 1(3) 201-221. University of Jyväskylä.

- Vázquez, G. (2007) Models of CLIL: An Evaluation of its status drawing on the German experience. A critical report on the limits of reality and perspectives. *RESLA* 1, 95-111.
- Victori, M. and Vallbona, A. (2008) A case study on the implementation of CLIL in a primary education school: results, benefits and challenges. Paper presented at CLIL TBL Seminar, UB, Barcelona, December, 12.
- Vygotsky, L.S. (1978). *Mind in society: The development of higher psychological processes*. Cambridge: MA: Harvard University Press..
- Washburn, Linda. (1997). English immersion in Sweden: A case study of Röllingby high school 1987-1989. Unpublished PhD thesis. Stockholm University
- Wesche, M.B. (2001). French Immersion and Content-Based Language Teaching in Canada. *The Canadian Modern Language Review*, 58 (1):84-102.
- Wesche, M.B. (2002). Early French immersion: How has the original Canadian model stood the test of time? In P. Burmeister, T. Piske and A. Tohde (eds.), *An integrative view of language development: Papers in honour of Hemming Wode*,. Trier, Germany: Wissenschaftlicher Verlag Trier. 357-375.
- Wesche, M. and Skehan, P. (2002). Communicative teaching, content-based instruction, and task-based learning. *Handbook of applied linguistics*. Ney York: Oxford University Press. 207-228.
- Whittaker, R. and Llinares, A. (2009). CLIL in Social Science classrooms: Analysis of spoken and written productions. In Y. Ruiz de Zarobe and R.M. Jiménez-Catalán *Content and Language Integrated learning. Evidence from Research in Europe* (eds.). Bristol: Multilingual Matters. 215-34.
- Whittaker, R., Llinares, A., and McCabe, A. (2011). Written discourse development in CLIL at secondary school. *Language Teaching Research*, 15 (3): 343-362.
- Wode, H. (1999). Language learning in European Immersion classes. In *Learning through a Foreign Language. Models, methods and outcomes*. J Masih (ed). London: Centre for Information on Language Teaching and Research. 16-25.
- Wolfe-Quintero, K., Inagaki, S., and Kim, H.Y. (1998). *Second Language Development in Writing: Measures of Fluency, Accuracy and Complexity*. Honolulu: University of Hawai'i Press.
- Wolff, D. (2001). Towards an advanced Level Programme in Content-Based Language Education. In P. Van de Craen and C. Pérez-Vidal (eds.), *The Multilingual Challenge/Le défi multilingue*. Barcelona: Printolubro, 27-39.
- Wolff, D. (2007). CLIL: Bridging the Gap between School and Working Life. In D. Marsh and D. Wolff (eds.), *Diverse Contexts- Converging Goals: CLIL in Europe*. Frankfurt: Peter Lang.
- Wray, A. (2002). *Formulaic Language and the Lexicon*. Cambridge: CUP.
- Vollmer, H. J., Lena, H., Randi, T., Debbie, C., and Verena. K. (2006). Subject-specific competence and language use of CLIL learners: The case of geography in grade 10 of secondary schools in Germany. Paper presented at the ESSE8 Conference in London, 29 August 2006
- Wood, D. (1998). *How Children Think and Learn*. Oxford: Blackwell.

- Wood, D., Bruner, J. S., and Ross, G. (1976). The role of tutoring in problem solving. *Journal of child psychology and psychiatry*, 17 (2): 89-100.
- Yano, Y., Long, M. H., and Ross, S. (1994). The effects of simplified and elaborated texts on foreign language reading comprehension. *Language Learning*, 44, 189–219.
- Zydati, W. (2007). *Deutsch.Englische Zge in Berlin: Eine evaluation des bilingualen sachfachunterrichts an gymnasien. Kontext, kompetenzen, konsequezen.* Frankfurt-am-Main: Peter-lang.

APPENDICES

Appendix A. Listening Test A1

est 2

Part 4 - 5 questions -

Listen and tick (✓) the box. There is one example.

Where did Katy go on her birthday?



A



B



C

1 How did Katy and her parents get there?



A



B



C

2 What time did the family arrive at the theatre?



A



B



C

3 Which was Katy's favourite actor?



A



B



C

4 Where did Katy sit in the theatre?



A



B



C

5 What did Katy buy at the theatre?



A



B



C

Appendix B. Listening Test B1

est 2

Part 3
- 5 questions -

Which thing must Alex take to each lesson tomorrow?

Listen and write a letter in each box. There is one example.



History

C



Sports



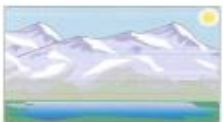
Music



Science



Art



Geography

Listening



A



B



C



D



E



F



G



H

Appendix C. Reading Test A1

Test 2

Reading and Writing

Part 1

- 10 questions -

Look and read. Choose the correct words and write them on the lines.
There is one example.

swans	queens	dinosaurs	a journalist
-------	--------	-----------	--------------

	These people work in circuses and people laugh at the things they do.	clowns
1	These animals can live in caves, in trees and sometimes in our houses! In some places, you see them flying and looking for food in the evening.	clowns
2	This is a kind of big, round fruit which is dark green outside and red or pink inside.	clowns
3	These are very rich women who are very important. Some of them live in castles	clowns
4	These animals lived a long time ago. They don't live now.	clowns
5	These people draw and paint pictures to sell.	clowns
6	You eat this from a bowl with a spoon or drink it from a big cup. You can make it with vegetables and meat.	clowns
7	This is made with fruit and sugar.	clowns
8	These animals live in water. People enjoy watching them swim, jump and do clever things.	clowns
9	This person looks for news stories and then writes them for us to read.	clowns
10	These animals can carry a lot of heavy things and can walk for a long time in hot weather without any water.	clowns

a watermelon	dolphins	a farmer
--------------	----------	----------

Appendix D. Reading Test B1

1004 4

Part 2
- 7 questions -

Look and read. Write **yes** or **no**.



Examples

There is a red butterfly which is flying above the wall. yes

The boy is swinging from the tree. no

Questions

- 1 The brown birds are bigger than the birds that are in the tree.
- 2 The rabbit which is near the girl has eaten all the yellow flowers.
- 3 Someone has put the knives and forks on the plates.
- 4 The woman is wearing a T-shirt which is yellow with red stripes.
- 5 The man is cooking some eggs on the fire.
- 6 The woman is sitting under an umbrella and talking on the phone.
- 7 The pockets on the girl's shorts are different colours.

Appendix E. Reading Test C

Part 6
- 5 questions -

Read the text. Choose the right words and write them on the lines.

Farms



Example

There's a lot of work to do on
 farm. In cold weather, people must give the sheep and cows
 1 food. Then there are a lot of things
 and places which people must clean. After that, they
 2 to work in the fields. At some farms,
 3 you can vegetables, plants, cheese and
 beautiful big brown eggs in a shop on the farm.
 4 The work on farms stops. There are
 always things to do in sunny weather, in the rain
 5 in the snow.

Example

- | | | | |
|---|-------|-------|---------|
| | any | some | a |
| 1 | their | her | its |
| 2 | do | can | have |
| 3 | buys | buy | buying |
| 4 | not | never | nothing |
| 5 | but | than | or |

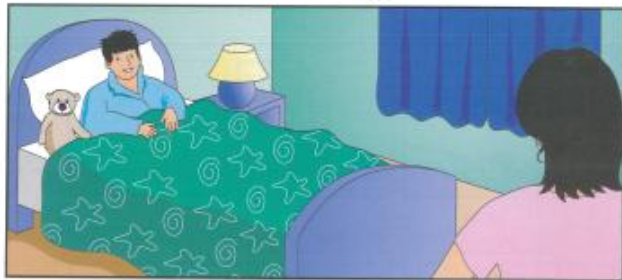
Appendix F. Reading Test D1

Part 5

- 10 questions -

Look at the pictures and read the story. Write words to complete the sentences about the story. You can use 1, 2 or 3 words.

Paul's birthday



It was Thursday morning. Paul opened his eyes and sat up in bed. "It's my birthday. I'm ten today," he said to his favourite bear. He looked at his old toys in the cupboard opposite his bed. "I don't need those toys now," he said. His mother came into his room with some presents. "Happy Birthday, Paul!" she said. "What do you want to do today? Go to the zoo? The cinema?" "I don't want to go out," Paul said. "I'm ten now. I must put my old toys in the basement." "OK," said his mother quietly.

Examples

Paul's birthday was on Thursday.

Paul talked to his favourite bear first.

Questions

- 1 Paul's old toys were in a cupboard which was
- 2 Paul's said "Happy Birthday!"
- 3 Paul wanted to take his toys downstairs to the

Reading and Writing



After breakfast, they found three boxes which they carried upstairs. Paul's mum picked up his toy train and put it in the first box. "No," said Paul. "I like playing with that sometimes." Then she put Paul's kangaroo very carefully in the second box. "No," said Paul. "I like his face. He's always smiling."

- 4 Paul and his mother carried upstairs.
- 5 The first toy that Mum picked up was Paul's
- 6 The kangaroo went in the box.
- 7 Paul liked his kangaroo's happy

Test 2



Paul's mum put his kite in the third box.
"No," said Paul. "Dad and I play with that in the park sometimes."
"Dad's toy cars are under our bed, Paul, and he's a lot older than you. You don't need these boxes," said Mum. "How about going to the zoo?"
Paul laughed. "Oh, yes!" he said. "Let's go!"

- 8 Paul and his father play in with Paul's kite sometimes.
- 9 Paul's father's were under the bed.
- 10 In the end, Paul wanted to go to with his mum.

Appendix G. Reading Test A2

Test 3

Reading and Writing

Part 1

- 10 questions -

Look and read. Choose the correct words and write them on the lines.
There is one example.

a mechanic a station pilots bridges

	These are often made of wool and if you wear them in cold weather, your hands don't get cold.gloves.....	
nurses	1 You enjoy going in these if you like flying. You can sometimes see them in the sky or at airports.	lorries
	2 You often walk over these when you cross rivers or roads. Traffic can go over bigger ones.	
shorts	3 This person plays a sport for his job. He must run fast and kick a ball.	ambulances
	4 If you want to go somewhere in your town but you don't have a car and there's no bus, you can go in this.	
a taxi	5 Some people have to wear these at work or at school so everyone looks the same.	gloves
	6 If you have a problem with your car, this person can help you.	
	7 When people need to catch trains, they must go to this place.	
uniforms	8 People use these to take big or heavy things by road from factories to shops, for example.	a street
	9 Some people like to wear these if the weather is hot and sunny.	
	10 These people work in hospitals and they help people who are ill.	

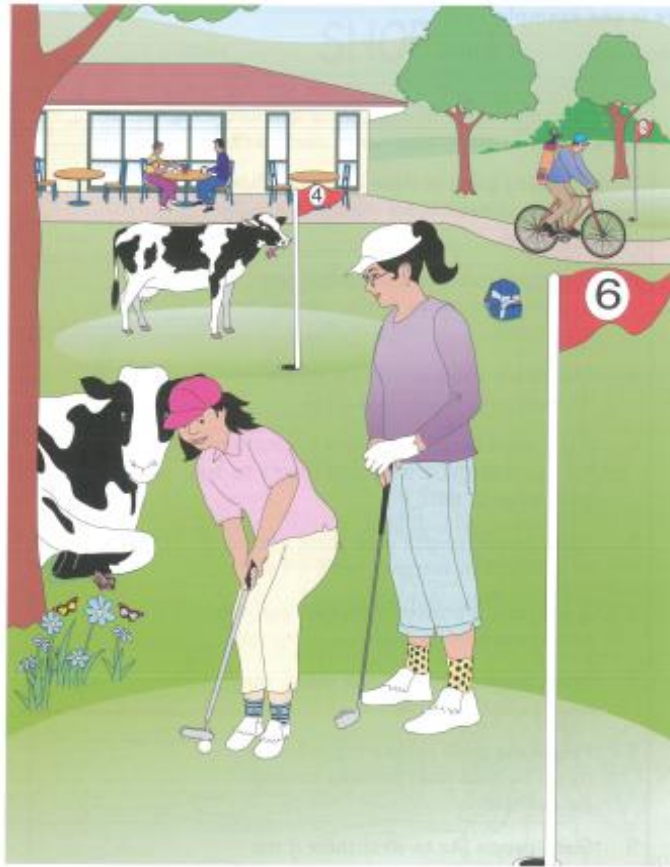
belts planes a footballer

Appendix H. Reading Test B2

Part 2
- 7 questions -

Reading and Writing

Look and read. Write **yes** or **no**.



Questions

- 1 There's a different number on each of the three flags.
- 2 You can see some butterflies in the air just above the flowers.
- 3 Both the people who are at the front of the picture are wearing socks with stripes on them.
- 4 The woman is carrying some glasses in her left hand.
- 5 One of the cows is lying down and the other cow is eating some grass.
- 6 There are a few little white balls by the girl's feet.
- 7 Someone has put a rucksack on the ground.

Examples

There are three round tables outside the café. yes

Two people are riding bicycles in the picture. no

Appendix I. Reading Test D2

Test 1

Part 5

- 10 questions -

Look at the pictures and read the story. Write words to complete the sentences about the story. You can use 1, 2 or 3 words.

Grandpa Bill



Last week, Sam and Anna's Grandpa Bill came to their house. They loved days with their grandfather because he did lots of exciting things with them. On Saturday, the children went to the zoo in their grandfather's car. Sam loved the monkeys and Anna loved the lions. At the end of the morning, they wanted to go home again, but Grandpa Bill couldn't start the car.

Examples

Grandpa Bill went to the children's house *last week*

Sam and Anna *loved* being with their grandfather.

Questions

1 Grandpa Bill took the children to the zoo on

2 At the zoo, Sam liked the a lot.

3 When the children wanted to go home, their grandfather couldn't the car.

Reading and Writing



They ran quickly to the bus station to catch a bus. On the bus, they started talking to two older children, Ben and Alex, who were with their parents. Grandpa Bill told the four children some very funny stories and they laughed a lot.

Then the mother said, "Would you all like to come to our house for lunch? We've got a swimming pool in the garden. You can swim there before lunch." Grandpa Bill said, "Thank you. We'd like that!"

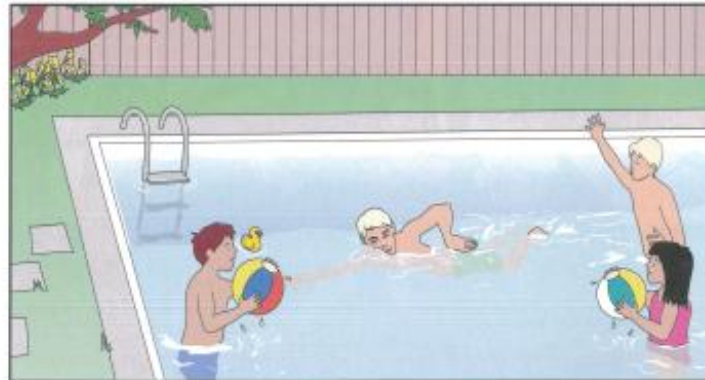
4 Grandpa Bill, Anna and Sam caught a bus at the

5 The two older children on the bus were called

6 Grandpa Bill's were very funny.

7 The older children's family had a in their garden.

Test 1



The family's house was very old. It was big and all the walls were blue. The four children swam then played badminton. Grandpa Bill watched them play. Then for lunch they all had chicken and vegetables. Alex took pictures of the family's three new friends and then Grandpa Bill, Sam and Anna said goodbye and went home.

- 8 The colour of the family's house was
- 9 The children played after their swim.
- 10 The family and their new friends ate for lunch.

Appendix J. Students' Background Questionnaire

CURS:

GRUP:

ESCOLA:

DATA :

PROFESSOR(A) D'ANGLÈS :

1. NOM:..... 2. COGNOMS:

3. Edat:..... anys

4. Data de naixement: ... de..... de

5. Lloc de naixement:.....

6. Sexe: Dona Home

7. Llengües que parles habitualment a casa:

amb la mare amb el pare

amb germans amb altres (especificar)

.....

En quina llengua et sents més còmode?

Català

Castellà

Les dues indistintament

8. Professió de la mare:.....

Professió del pare:

9. A quin curs vas començar a fer anglès a l'escola?

PARVULARI: 3anys 4 anys 5 anys

PRIMÀRIA: 1er 2on 3er 4art 5è 6è

Quantes hores setmanals feies d'anglès a l'escola quan feies Primària

4h 3h 2h 1h

A quina escola anaves quan vas començar a fer anglès?.....

10. Has repetit algun curs? Quin/s?.....

11. Has fet o fas anglès extraescolar?.....

A quin curs vas començar a fer-les? anys

Durant quant temps n'has fet?

Quantes hores per setmana fas o feies aproximadament?:

4h setmanals 3h setmanals 2h setmanals 1h setmanal

12. Has fet alguna estada en un país de parla anglesa en la que utilitzessis l'anglès de manera habitual?

SI NO

Si és que sí,

Quantes vegades?.....

A on?.....

Quants anys tenies?..... Durant quant de temps?.....

Amb qui?.....

Vas fer un curs d'anglès?.....

13. Parles anglès fora de classe habitualment?

Mai De vegades Molt sovint

Si és que sí, amb qui parles anglès?

Amb amics Mai De vegades Molt sovint

Amb persones conegudes Mai De vegades Molt sovint

Amb estrangers a l'estiu Mai De vegades Molt sovint

14. Utilitzes l'anglès fòra de l'escola ?

Mai De vegades Molt sovint

Si és que sí, en quines activitats i freqüència? (sols marca les que utilitzis)

Per cercar informació a Internet 2 o 3 cops setmanals 2 o 3 cops mensuals

Per xatejar o fer jocs per Internet 2 o 3 cops setmanals 2 o 3 cops mensuals

Per llegir texts o revistes en anglès 2 o 3 cops setmanals 2 o 3 cops mensuals

Per escoltar cançons 2 o 3 cops setmanals 2 o 3 cops mensuals

Per veure la TV en anglès 2 o 3 cops setmanals 2 o 3 cops mensuals

Appendix K. CLIL Teachers' Background Questionnaire

Part 1. Biodata & Professional Background

- 1- Name and surname:
- 2- Age:
- 3- Education centre name:
- 4- I work as a:
 - a. Primary teacher b. Secondary teacher c. University teacher d. Other (please specify)
- 5- I have been teaching for years.
- 6- Which is/are your major university degrees?
 - a. Diplomatura b. Llicenciatura c. Master d. Doctorat
- 7- English language qualifications. Please, write the highest one
- 8- Do you use English in your private life (emails, chat, movies, books, magazines)?
 - a. I don't use English in my private life
 - b. Yes, I use English some few times (once/twice a month)
 - c. Yes, I use English sometimes (three/four times a month)
 - d. Yes, I use English often (twice a week)
 - e. Yes, I use English frequently (more than twice a week)
 - f. Yes, I use English every day
- 9- Foreign languages that you speak and your level of fluency:
 - a. ----- very fluent, fluent, sufficient, non fluent
 - b. ----- very fluent, fluent, sufficient, non fluent
 - c. ----- very fluent, fluent, sufficient, non fluent
- 10- Do you attend some refresher/updating courses in the English language/EFL teaching?
 - a. Four or more times a year b. Two/Three times a year c. Once a year d. No
- 11- Have you participated in any European mobility programmes like Comenius/Grundvig/Erasmus?
 - a. Yes, please specify b. No

Part 2. CLIL Professional Background

- 12- Level at which you teach CLIL:
 - a. Primary: 1 First Cycle 2 Second Cycle 3 Third Cycle
 - b. Secondary Education (ESO): 1 First grade 2 Second grade 3 Third grade 4 Fourth grade

c. High School Degree (Batxillerat): 1 First 2 Second

13- Before starting your experience with CLIL you: (some possible answers)

- a. You attended a CLIL course in your country
- b. You attended a CLIL course in a foreign country
- c. You learnt about CLIL by yourself
- d. You learnt about CLIL working in collaboration with the EFL teacher
- e. You learnt about CLIL being a CLIL student and you decided to implement it in your teaching
- f. Other

14- Your CLIL training is (in total):

- a. Up to 20 h
- b. 20-40 h
- c. 40-60 h
- d. 60-80 h
- e. 80-100 h
- f. More.....

15- Experience in CLIL teaching before this academic year:

- a. Occasional (short)
- b. About two years
- c. About three or more years (number of years.....)

16- How is your CLIL teaching developed?

- a. Single CLIL units (please specify duration)
- b. Single CLIL workshops (please specify duration)
- c. Mainstream CLIL course (please specify duration)
- d. Optional CLIL course (please specify duration)

17- Content-subject in which you are implementing CLIL/AICLE:
.....

18 – Did the school make any special arrangements for you to prepare the CLIL subject? Which ones? (e.g. fewer teaching hours)

19 – Did you volunteer to teach CLIL or were you chosen by the school?

20- State what you consider to be the main benefits of teaching in CLIL
.....

21 - Write the main challenges/difficulties you face when you teach in CLIL
.....

Part 3. CLIL Implementation

22- Are you an EFL teacher or a content teacher?

23 - If you are a content teacher, in what percentage did you plan the CLIL implementation with the EFL teacher?

- a. 100%
- b. 80-60%
- c. 40-20%
- d. I plan the CLIL implementation alone, on my own

24- If you are an EFL teacher, in what percentage did you plan the CLIL implementation with the subject/content teacher?

Language Competence of Young Learners Exposed to EFL and CLIL

a. 100% b. 80-60% c. 40-20% d. I plan the CLIL implementation alone, on my own
25- Do you use a textbook?

a. Yes Which one?..... b. No

26- How often do you use ...

1- Materials adapted from the Internet

a. Always 80-100% b. Frequently 80-40% c. Sometimes 40-20% d. Few times 20%-0%

2- Materials created for English Foreign Language Learners

a. Always 80-100% b. Frequently 80-40% c. Sometimes 40-20% d. Few times 20%-0%

3- My own created materials

a. Always 80-100% b. Frequently 80-40% c. Sometimes 40-20% d. Few times 20%-0%

27- How often do you use...

1 Audio materials

a. Always 80-100% b. Frequently 80-40% c. Sometimes 40-20% d. Few times 20%-0%

2 Audio-visual materials (e.g. videos)

a. Always 80-100% b. Frequently 80-40% c. Sometimes 40-20% d. Few times 20%-0%

3 Visual materials

a. Always 80-100% b. Frequently 80-40% c. Sometimes 40-20% d. Few times 20%-0%

4 Written materials

a. Always 80-100% b. Frequently 80-40% c. Sometimes 40-20% d. Few times 20%-0%

28- How often do your CLIL students:

1 Work in pairs or small groups

a. Always 80-100% b. Frequently 80-40% c. Sometimes 40-20% d. Few times 20%-0%

2 Participate in cooperative activities in the classroom

a. Always 80-100% b. Frequently 80-40% c. Sometimes 40-20% d. Few times 20%-0%

3 Participate in on-line learning environments

a. Always 80-100% b. Frequently 80-40% c. Sometimes 40-20% d. Few times 20%-0%

4 Participate in hands-on activities (experiential activities)

a. Always 80-100% b. Frequently 80-40% c. Sometimes 40-20% d. Few times 20%-0%

5 Do some Internet searching

a. Always 80-100% b. Frequently 80-40% c. Sometimes 40-20% d. Few times 20%-0%

6 Use ICT for lesson development

a. Always 80-100% b. Frequently 80-40% c. Sometimes 40-20% d. Few times 20%-0%

7 Make a product

- a. Always 80-100% b. Frequently 80-40% c. Sometimes 40-20% d. Few times 20%-0%

8 Design and do a project

- a. Always 80-100% b. Frequently 80-40% c. Sometimes 40-20% d. Few times 20%-0%

29- Which language do your CLIL students usually use in the CLIL class?

- a. when they interact with the teacher
b. when they interact among themselves

30- Is there spontaneous use of English in the CLIL class?

- a. Always 80-100% b. Frequently 80-40% c. Sometimes 40-20% d. Few times 20%-0%

If yes, what kind?

31 - How often do you use translation into Spanish/Catalan.

- a. Always 80-100% b. Frequently 80-40% c. Sometimes 40-20% d. Few times 20%-0%

32 - Do you think your CLIL students understand everything you say in English?

- a. Always 80-100% b. Frequently 80-40% c. Sometimes 40-20% d. Few times 20%-0%

33- What do you do when your CLIL students do not understand you? Write three strategies you use.

1.
2.
3.

34- Which listening activities do you do in the CLIL class? Write some examples.

How do your students respond to them?

35- Which reading activities do you do in the CLIL class? Write some examples.

How do your students respond to them?

36- Which speaking activities do you do in the CLIL class? Write some examples.

How do your students respond to them?

37- Which writing activities do you do in the CLIL class? Write some examples.

How do your students respond to them?

38- Is there one of the four skills which dominates in your CLIL class? Which one? Why?

39 - Do you work with the four skills in the CLIL class in the same way as you would do in the English class?

- a. Yes Why? b. No Why?

40 - When you assess your CLIL students (you can select more than one option)

- a. I do it the same way as if I was teaching in the first language
b. I use different assessment strategies (please specify which strategies)
c. I design the assessment with the content teacher
d. I design the assessment with the EFL teacher
e. I design self-assessment tasks for my students

41 - You feel confident enough when you are teaching CLIL because

1 I master the content subject register/discourse in English

- a. I strongly agree b. I somewhat agree c. I somewhat disagree d. I strongly disagree

2 I master the specific subject vocabulary and terminology in English

- a. I strongly agree b. I somewhat agree c. I somewhat disagree d. I strongly disagree

3 I master the classroom management language in English

- a. I strongly agree b. I somewhat agree c. I somewhat disagree d. I strongly disagree

42 - Do you think your students benefit from the CLIL class? In terms of content? In terms of language? Both?

43 - Do you think your students enjoy the CLIL class? Why (not)?

Further comments or final suggestions.

THANK YOU!

Appendix L. CLIL Teachers' Opinion Questionnaire

NAME

Please answer the following questions in Catalan or in English. Thank you !

1. What is your general impression of this first CLIL course ?
2. What would you say have been your strengths and weaknesses of the course ?
3. Have you felt comfortable in class ? Why ?
4. Did the students follow the class easily ?
5. Have the students learnt content ?
6. Have the students learnt language ? Specify how much improvement you perceive and in what areas you perceive it.
7. What language did you speak in class ?
8. What language did students speak in class ?
 - a. When asking questions ?
 - b. When answering questions ?
 - c. Among themselves ?
 - d. When doing groupwork ?
9. How do you think you can improve your classes ?
10. Do you think you would need help from the content teacher ?
11. Do you think you would need institutional support (eg. extra time to prepare lessons) ?
12. How do you prepare activities ? Do you create them or downlaod them from The Internet... ?
13. Do you think you would need more CLIL training ?
14. What are the best and the worst moments during a CLIL class/CLIL course ?
15. Would you like to continue being a CLIL teacher in the future?

Appendix M. CLIL Class Observation Protocol

	Highly evident	3	Somewhat evident	2	1	Not evident	0
Preparation	4	3	2	1	0		
Clearly define content objectives.							
Clearly define language objectives.							
Use supplementary materials to make lessons clear and meaningful.							
Adapt content to all levels of student proficiency.							
Provide meaningful and authentic activities that integrate lesson concepts with language practice opportunities.							
Building Background	4	3	2	1	0		
Make clear links between students' past learning and new concepts							
Explicitly link concepts to students' background experience.							
Emphasize key vocabulary							
Comprehensible Input	4	3	2	1	0		
Speak appropriately to accommodate students' proficiency level.							
Clearly explain academic tasks.							
Use a variety of techniques to make content concepts clear.							
Strategies	4	3	2	1	0		
Consistently use scaffolding techniques through the lesson							
Employ a variety of question and task types							
Interaction	4	3	2	1	0		
Provide frequent opportunities for interaction and discussion.							
Group students to support language and content objectives (e.g. pair-work, group-work, etc.)							
Consistently provide sufficient wait time.							
Give ample opportunities for clarification for concepts in L1 (e.g. peer, L1 texts, etc.)							
Practice-Application	4	3	2	1	0		
Supply lots of hands-on materials							
Provide activities for students to apply content/language knowledge.							
Integrate all language skills (i.e. reading, writings, listening, and speaking) into each lesson.							
Lesson Delivery	4	3	2	1	0		
Clearly support content objectives							
Clearly support language objectives							
Engage students 90-100% of the lesson							
Appropriately pace the lesson to students' ability level.							

*Write comments on the back.

Appendix N. Sample of Writings from High and Low achievers

Low achiever

Time 0.

I Genesis and 10 yeards ol, I'm very happy my favourites colours is purple and blue I like dance and singer. My favourite food is macarronis I go to Eanglissh class.

Time 1.

I have 2 brothers and he's names is Erick and Ismael my parent are very beautifuld the name of my mother is Adriana and name of my father is Rufino, I have a lots of cosins. This hollidays I go to Lloret de Mar I have friends example Carla, Olivia, Ari, Nuria F, Nuria G, Carme, Sara, Lina, Roger, Manel...And my favorites animals are: dogs and dolphins

Time 2.

Yesterday I was in cinema of my best friend Mireia, we soe film Tin Tin and the adventure of (Unicornio). In this momen I travell do the test of English it's a little bit dificult. Next week I go to my best friend to his house to go play and to go to the shopping and of course to go to the cinema. We go to see kika super bruja in Spanish language.

Time 3.

When I was little I live in Mallorca. I flew witch plane to Vic. I was two years old. I 've got a lot of friends. I'm 12 years old. My favourite food is macarronie I study in Vic My best friends are Aina and Mireia. I will be one dancer. I will goen to New York and live in a big house. I love New York because it is big and funny.

High Achiever

Time 0

I'm anna. I'm ten. I love school. My family is very neas. My life is very good. My mom is Maite my dad is Jordi, my sister is Gemma. My favority colors is orange and green. I love cats and dogs. My favorite book is Peter Pan. My life (construeix) un present and futur. I love angles. My freindes is childrens in my classmates.

Time 1

Hello! I'm Anna and I have got a sister she's name is Gemma, my mum name is Maite and the name of my dad is Jordi. I have got two cats and a canary. I live in Guissona, the name of my school is Mare de Déu del Roser I like dancing, playin futbol, playing

basketball...I have got long brown hair and small brown eyes, My sister Gemma has got long brown hair and big brown eyes, My mum Maite has got short brown hair and big brown eyes, and my dad has got short blond hair and small blue eyes. My cat Tina has got black hair and yellow eyes, and she's hair is blond and brown. My cat Mickey has got blue eyes and he's hair is blond. My canary Ricki has got black eyes, and he's yellow.

Time 2

Hello, I'm Anna, my father is Jordi, my mum is Maite and I have got a sister, she is Gemma. I'm from Catalonia and I live in Guissona. When I was young, my favourite food was rice, now my favourite food is pasta, and in my future I don't know.

When I was young, my favourite hobby was play with my sister, the same to now and in my future I don't know.

When I was young, my favourite hobby was play with my sister, the same to now and in my future I don't know.

I have three pets: a canary, he is Coco. A cat, she is Tina and her soon, he is Mickey.

I have got family in: Agramunt, Montclar, Sunyer...

My birthday is on 11th of December and I have lots of presents.

Last year I do my first comunion and I went too nice clothes.

I like my sister friends come in my house to stay sleeping. I'ts funny!

Time 3

Now, I'm 10. My mother's name is Maite and my father is Jordi. They're work in a baker. They work many hours. I have got a sister, her name is Gemma she's 17. I have got two cats: Tina is the mother and Mickey is her son. I have got a canary, too, his name is Ricky. I live in Guissona. In my past I lived in Guissona. And I have got the same pets, and two tortoises, too. They're cold Rebeca and Mario. I have got the first day at dance school when I went to P-5. At the future, I will like work at the clotes shop or design clotes because I like clothes very much. Now my favourite food is pasta. When I six my favourite food was rice. Now I want to eat rice every day. I love it. I go to school every day. I walk to school and in the afternoon I go to English classes and sports.