



UNIVERSITAT DE
BARCELONA

Language Learning from Subtitled TV Series in the Primary EFL Classroom

Montserrat Casulleras Marquès

ADVERTIMENT. La consulta d'aquesta tesi queda condicionada a l'acceptació de les següents condicions d'ús: La difusió d'aquesta tesi per mitjà del servei TDX (www.tdx.cat) i a través del Dipòsit Digital de la UB (diposit.ub.edu) ha estat autoritzada pels titulars dels drets de propietat intel·lectual únicament per a usos privats emmarcats en activitats d'investigació i docència. No s'autoritza la seva reproducció amb finalitats de lucre ni la seva difusió i posada a disposició des d'un lloc aliè al servei TDX ni al Dipòsit Digital de la UB. No s'autoritza la presentació del seu contingut en una finestra o marc aliè a TDX o al Dipòsit Digital de la UB (framing). Aquesta reserva de drets afecta tant al resum de presentació de la tesi com als seus continguts. En la utilització o cita de parts de la tesi és obligat indicar el nom de la persona autora.

ADVERTENCIA. La consulta de esta tesis queda condicionada a la aceptación de las siguientes condiciones de uso: La difusión de esta tesis por medio del servicio TDR (www.tdx.cat) y a través del Repositorio Digital de la UB (diposit.ub.edu) ha sido autorizada por los titulares de los derechos de propiedad intelectual únicamente para usos privados enmarcados en actividades de investigación y docencia. No se autoriza su reproducción con finalidades de lucro ni su difusión y puesta a disposición desde un sitio ajeno al servicio TDR o al Repositorio Digital de la UB. No se autoriza la presentación de su contenido en una ventana o marco ajeno a TDR o al Repositorio Digital de la UB (framing). Esta reserva de derechos afecta tanto al resumen de presentación de la tesis como a sus contenidos. En la utilización o cita de partes de la tesis es obligado indicar el nombre de la persona autora.

WARNING. On having consulted this thesis you're accepting the following use conditions: Spreading this thesis by the TDX (www.tdx.cat) service and by the UB Digital Repository (diposit.ub.edu) has been authorized by the titular of the intellectual property rights only for private uses placed in investigation and teaching activities. Reproduction with lucrative aims is not authorized nor its spreading and availability from a site foreign to the TDX service or to the UB Digital Repository. Introducing its content in a window or frame foreign to the TDX service or to the UB Digital Repository is not authorized (framing). Those rights affect to the presentation summary of the thesis as well as to its contents. In the using or citation of parts of the thesis it's obliged to indicate the name of the author.

LANGUAGE LEARNING FROM SUBTITLED TV SERIES IN THE PRIMARY EFL CLASSROOM

Tesi doctoral presentada per

Montserrat Casulleras Marquès

com a requeriment per a l'obtenció del títol de

Doctora en Filologia Anglesa

Programa de Doctorat: *Estudis Lingüístics, Literaris i Culturals*

Línia de recerca: *Lingüística Aplicada*

Departament de Llengües i Literatures Modernes i d'Estudis Anglesos

Directora i Tutora: **Dra. Immaculada Miralpeix Pujol**

Universitat de Barcelona

2023

ACKNOWLEDGEMENTS

First and foremost, I am extremely grateful to my supervisor Dr. Imma Miralpeix for her invaluable advice, continuous support and patience throughout all these years. Imma, this PhD thesis would not have been possible without you. Since the very beginning you have encouraged me to do better and to keep going. Thank you for your knowledge, feedback, comments and the many hours and effort you put into this dissertation. I will never be able to thank you enough for all your dedication. I have learnt a lot from you and I will always be grateful.

I would like to express my sincere gratitude to Prof. Carmen Muñoz for being tutor of this thesis when it was first registered. I would also like to thank the GRAL group for welcoming me when I could attend any talks or conferences at the University of Barcelona. I would like to extend my thanks to Dr. Joan Tomàs Pujolà, Dr. Natalia Fullana and Dr. M. Luz Celaya for having been part of the follow-up committee during all these years.

I would also like to thank 'Col.legi Sant Josep' from Vilafranca del Penedès for letting me pilot my study and 'Col.legi Igualada' from Igualada for letting me collect data for more than five months. Thank you to my work colleagues for lending me a hand when I needed it and to the students without whom this study would not have been possible.

I am indebted to Dr. Francesca Walls, who helped me in giving this thesis a native accent. Thank you for your feedback, fast and efficient work. I would also like to thank Judit Solà, who provided me with the statistical help I needed at a crucial time.

Per últim, però no menys important, voldria donar les gràcies a la meva família, especialment als meus pares, Josep i Montserrat, a qui dedico aquesta tesi. Moltes gràcies pel vostre suport i recolzament durant tots aquests anys. Sense vosaltres aquesta tesi tampoc seria possible. Gràcies al meu germà Jordi, per recolzar-me i ajudar-me sempre amb totes les dificultats tècniques que han anat sorgint i milions de gràcies a les meves nebodes Laia i Júlia, per donar-me energia i fer-me riure sempre que ho he necessitat. Finalment, moltes gràcies als meus amics i amigues per encoratjar-me i recolzar-me tant en els bons moments com en els dolents. A tots vosaltres, moltes gràcies per tot!

ABSTRACT

Even if audiovisual materials have been shown to offer several advantages for L2 learning (Danan, 2004), in previous literature very little attention has been paid to L2 learning from multimodal input in young low-proficiency learners (d'Ydewalle & Van de Poel, 1999; Matiolo et al., 2015; Alexiou, 2015). Furthermore, most of the studies conducted so far are one-off studies, while there is scarce longitudinal research on extensive viewing with young populations in instructional settings (Gesa, 2019; Avello, 2023).

In this study, two groups of Spanish/Catalan students learning English as FL at primary school, watched one episode a week of the animated TV series *Curious George* with the audio in English but subtitles either in Spanish (L1S, N=47) or English (L2S, N=45) over a period of 5 months. After each episode, students were immediately tested on episode comprehension and L2 written word-form recognition. In order to explore the role of several variables on their scores, before the intervention learners answered a questionnaire enquiring about previous out-of-school exposure (OSE) to audiovisual materials. They were also tested on language aptitude, L1 / L2 Reading Speed (RS) and L2 receptive Vocabulary Size (VS). Finally, in order to examine possible changes in their viewing behaviour when they were repeatedly exposed to multimodal input over a long period of time, they answered the same video watching questionnaire before and after the intervention.

Results showed that both groups were understanding the videos and learning vocabulary from them: although differences between the groups were not always significant, the L1S group consistently scored higher in comprehension, while the L2S group was better at L2 word-form recognition of known and new words appearing in the episodes. Linguistic aptitude was revealed to be closely related to episode comprehension in both groups, and to vocabulary scores especially in the L1S group. Regarding possible proficiency effects, L2 VS had a significant effect on L2 comprehension in both groups, whereas RS was only significantly related to the scores obtained by the L1S group.

It was also observed that the L1S group required less time to read the subtitles than their L2S peers since the beginning of the study, as participants' RS was better in the L1. No significant changes were observed in the viewing behaviour of the two groups, although the L2S group showed a significant shift in their lexical learning perception at the end of the intervention, as they perceived that they had learned more vocabulary. Findings also showed that those participants who had been previously exposed to multimodal input at home did not actually obtain significantly higher results in the tests than their peers who had not been.

The outcomes from this dissertation suggest that exposure to multimodal input promotes L2 comprehension and vocabulary learning in young low-proficiency L2 learners. However, they also indicate that gains are not always evident and do not present a clear ascending order as amount of exposure accumulates. The findings also point towards the important role of language aptitude and L2 VS for L2 learning through audiovisual materials. The results of the study, which are discussed considering the previous research available on the topic, also offer new insights into extensive viewing with young learners for teachers and practitioners.

RESUM

Els materials audiovisuals han demostrat oferir múltiples avantatges per a l'aprenentatge d'una segona llengua (L2) (Danan, 2004). Malgrat això, pocs estudis previs s'han centrat en l'aprenentatge de la L2 a través d'input multimodal en aprenents joves de baix nivell de proficiència (d'Ydewalle & Van de Poel, 1999; Matielo *et al.*, 2015; Alexiou, 2015). A més, hi ha molt poca recerca longitudinal que analitzi l'exposició regular a la televisió en versió original en poblacions joves en contextos d'instrucció formal (Gesa, 2019; Avello, 2023).

En aquest estudi, dos grups d'alumnes bilingües castellà/català que aprenen anglès com a llengua estrangera a Primària, van visionar setmanalment un episodi de la sèrie dels dibuixos animats *Curious George* –amb l'àudio en anglès, però subtitulats en castellà (L1S, N= 47) o en anglès (L2S, N= 45)– durant un període de cinc mesos. Després de cada episodi, l'alumnat va realitzar un test per avaluar-ne la comprensió i reconeixement de vocabulari escrit en anglès. Per tal d'investigar el paper de diverses variables en els resultats obtinguts, abans de començar la intervenció l'alumnat va completar un qüestionari on se li preguntava sobre la seva exposició extraescolar a materials audiovisuals. També van realitzar tests per mesurar la seva aptitud lingüística, la velocitat lectora en L1 (castellà) o L2 (anglès) i la mida del seu vocabulari en L2. Finalment, per tal d'analitzar els possibles canvis dels joves aprenents en la seva manera de visionar l'input multimodal després d'estar-hi exposats de forma repetida en el temps, van respondre un mateix qüestionari sobre el visionat dels vídeos abans i després de la intervenció.

Els resultats mostren que ambdós grups entenien els vídeos i aprenien vocabulari després de visionar cada episodi: tot i que les diferències entre grups no eren sempre significatives, el grup L1S va obtenir puntuacions més altes de forma consistent en la comprensió, mentre que el grup L2S va obtenir millors resultats en el reconeixement de vocabulari, tant de paraules que ja coneixia prèviament com de noves que apareixien als episodis. L'aptitud lingüística va resultar estar relacionada estretament amb els resultats dels tests de comprensió en ambdós grups, i també amb els de vocabulari, especialment en el grup L1S. Pel que fa als possibles efectes de la proficiència, la mida del vocabulari en L2 de l'alumnat va mostrar un efecte significatiu en la comprensió dels episodis en els dos grups, mentre que la velocitat lectora només es relacionava significativament amb les puntuacions obtingudes pel grup L1S.

També s'observà que el grup L1S va necessitar menys temps per llegir els subtítols que els seus companys del grup L2S des de l'inici de l'estudi, ja que la velocitat lectora dels participants era millor en L1. No es va observar cap canvi significatiu en la manera de visionar els vídeos per part dels dos grups, tot i que el grup L2S va mostrar un gir significatiu en la seva percepció d'aprenentatge de vocabulari al final de la intervenció, quan va considerar que aprenia més paraules. També s'evidencia que aquells participants que havien estat prèviament exposats a input multimodal fora de l'escola no obtenia resultats significativament més alts que els seus companys que no ho havien estat.

Els resultats d'aquesta tesi suggereixen que l'exposició a input multimodal promou la comprensió i l'aprenentatge de vocabulari d'una segona llengua en aprenents joves de nivell inicial baix. No obstant això, indiquen que els beneficis no són sempre evidents i no presenten un clar ordre ascendent paral·lel a la quantitat d'exposició a l'input multimodal. També assenyalen l'important rol de l'aptitud lingüística i la mida del vocabulari en L2 dels aprenents per a l'aprenentatge de llengües estrangeres a partir de materials audiovisuals. Els resultats de l'estudi, que s'argumenten en relació a la recerca existent sobre el tema, ofereixen noves perspectives sobre l'exposició regular dels aprenents joves a input multimodal, i són també rellevants per a la pràctica docent.

TABLE OF CONTENTS

CHAPTER ONE – Introduction	1
1.1. Contextualisation	1
1.2. TV series for L2 learning	2
1.3. Motivation for the present dissertation	6
1.4. Organisation of the thesis	7
CHAPTER TWO – Multimodal input in L2 learning	9
2.1. The role of authentic input in L2 learning	9
2.2. Input modalities	11
2.2.1. Unimodal and bimodal input	11
2.2.2. Multimodal input	14
2.3. Theories on learning from verbal and non-verbal input	15
2.3.1. Dual Coding Theory (Paivio, 1986)	15
2.3.2. Cognitive Load Theory (Sweller, 1988)	17
2.3.3. Cognitive Theory of Multimedia Learning (Mayer, 2009)	18
2.4. Media and audiovisual materials for L2 learning	19
2.4.1. TV cartoons and L2 learning	21
2.5. Subtitling for L2 learning	28
2.5.1. L1 Subtitles (L1S)	30
2.5.2. L2 Subtitles (L2S)	30
2.5.3. L1 Subtitles (L1S) vs. L2 Subtitles (L2S)	32
2.5.4. Keyword subtitles	34
2.5.5. Advantages and disadvantages of subtitling	36
2.5.6. Reading behaviour with subtitles	39
CHAPTER THREE – Multimodal input and L2 comprehension	41
3.1. Previous considerations	41
3.1.1. Listening comprehension in the L2	41
3.1.1.1. Bottom-up and top-down processing in listening comprehension	42
3.1.1.2. Measuring listening comprehension	44
3.1.1.2.1. Multiple Choice (MC) tests	45
3.1.1.2.2. Item sequencing	46
3.1.1.2.3. True / False (T / F) questions	47
3.1.2. Bimodal input and L2 comprehension	48
3.2. Multimodal input and L2 comprehension	50
3.2.1. Studies with adult learners	59
3.2.2. Studies with adolescent learners	69
3.2.3. Studies with young learners	72
CHAPTER FOUR – Multimodal input and L2 vocabulary learning	77
4.1. Vocabulary learning in the L2	77
4.1.1. Single words vs. multi-word units	79
4.1.1.1. Collocations	81
4.1.1.2. Compounds	82
4.1.1.3. Phrasal verbs	83

4.1.1.4. Linking adverbials	84
4.1.1.5. Speech act formulae	84
4.1.2. Implicit / Explicit and Intentional / Incidental L2 vocabulary learning	86
4.1.2.1. L2 vocabulary learning and young learners	91
4.1.2.2. Incidental vocabulary learning from reading and listening	92
4.1.2.2.1. Incidental word-form recognition from reading and listening	97
4.1.2.3.2. Incidental L2 vocabulary learning from multimodal input	100
4.1.3. Factors affecting word learning: evidence from studies on reading, listening and TV watching	103
4.2. Effects of L1/L2 subtitling on vocabulary learning from multimodal input	109
4.2.1. Studies with adult learners	115
4.2.2. Studies with young learners	122
CHAPTER FIVE – Individual differences and other factors that may influence L2 learning from multimodal input	127
5.1. Learners’ Individual Differences (IDs): aptitude and age	127
5.2. Aptitude	129
5.2.1. Definition and measurement	129
5.2.1.1. The Modern Language Aptitude test (MLAT)	130
5.2.1.2. The LLAMA language aptitude tests	131
5.2.2. The role of aptitude in L2 learning	133
5.2.3. Aptitude in young and adult learners	135
5.3. Age	140
5.3.1. Young learners and FL learning	141
5.3.1.1. Out-of-school exposure (OSE) to English in young learners	144
5.4. L2 proficiency, vocabulary size and reading speed	150
5.4.1. L2 proficiency level	150
5.4.2. L2 vocabulary size (VS)	155
5.4.3. Reading speed (RS)	160
CHAPTER SIX – Research questions and method	167
6.1. Research Questions	167
6.2. Methodology	169
6.2.1. Participants	169
6.2.2. Instruments	171
6.2.2.1. <i>Curious George</i> TV series	171
6.2.2.2. Vocabulary Size test	176
6.2.2.3. Reading speed tests	177
6.2.2.3.1. Spanish reading speed test	178
6.2.2.3.2. English reading speed test	179
6.2.2.4. Aptitude test: LLAMA_B	181
6.2.2.5. Out-of-school exposure (OSE) questionnaire	183
6.2.2.6. Video-watching questionnaire	184
6.2.2.7. Episode-based comprehension tests	185
6.2.2.7.1. Item analysis	189
6.2.2.8. Episode-based vocabulary tests	190
6.2.2.9. School English mark	194
6.2.3 Procedure	195

6.2.3.1. Pilot study	195
6.3.2.2. Main study	198
6.2.3.2.1. Pre-treatment sessions	199
6.2.3.2.2. Treatment sessions	203
6.2.3.3 The teacher-researcher	206
6.2.4. Data analysis	208
6.2.4.1. Test scoring	208
6.2.4.1.1. Vocabulary size test	208
6.2.4.2.2. Reading speed tests	209
6.2.4.2.3. LLAMA_B test	209
6.2.4.2.4. Episode-based comprehension and vocabulary tests	210
6.2.4.2. Statistical analyses	210
6.2.4.2.1. RQ1: Effects of watching L1/L2 subtitled TV series on L2 comprehension and vocabulary learning	211
6.2.4.2.2. RQ2: Relationship between L2 comprehension and vocabulary scores with aptitude and proficiency variables	212
6.2.4.2.3. RQ3: Relationship between extensive viewing and viewing Behaviour	212
6.2.4.2.4. RQ4: Effects of OSE on L2 comprehension and vocabulary learning from extensive viewing	213
CHAPTER SEVEN – Results	215
7.1. RQ1: Comprehension and vocabulary learning from extensive viewing	215
7.1.1. L2 comprehension	216
7.1.2. Vocabulary learning	218
7.1.3. Summary of findings	223
7.2. RQ2: Aptitude and proficiency effects on learning from extensive viewing	224
7.2.1. Aptitude in relation to comprehension and vocabulary scores: correlations	225
7.2.2. Proficiency variables in relation to comprehension and vocabulary scores: correlations	227
7.2.2.1. L1 / L2 reading speed	227
7.2.2.2. L2 vocabulary size	228
7.2.2.3. English class grade	229
7.2.3. Proficiency variables in relation to comprehension and vocabulary scores: regression analyses	231
7.2.4. Summary of findings	234
7.3. RQ3: Extensive viewing and viewing behaviour	235
7.3.1. Self-reported level of understanding	236
7.3.2. Viewing behaviour	238
7.3.3. Reading and learning from subtitles	240
7.3.3.1. Time taken to read subtitles	241
7.3.3.2. Reasons for reading subtitles	243
7.3.3.3. Subtitles and vocabulary learning	245
7.3.4. Summary of findings	246
7.4. RQ4: Effects of OSE on L2 comprehension and vocabulary learning from extensive viewing	247

CHAPTER EIGHT – Discussion	251
8.1. Introduction	251
8.2. Extensive viewing in the light of theories on learning from multimodal input	252
8.2.1. Cartoons and L2 learning	254
8.3. L2 comprehension and vocabulary learning: L1 vs. L2	256
8.3.1. Effects of subtitled TV viewing on comprehension: L1S vs. L2S	257
8.3.2. Effects of subtitled TV viewing on word-form recognition: L1S vs. L2S	261
8.4. The role of aptitude and proficiency on L2 comprehension and vocabulary learning through multimodal input	266
8.4.1. Aptitude	267
8.4.2. Proficiency	270
8.5. Extensive viewing: participants' perceptions and viewing behaviour at the beginning and end of the treatment	278
8.5.1. Self-reported level of understanding	278
8.5.2. Viewing behaviour	279
8.5.3. Self-perception of learning from the subtitles	281
8.5.3.1. Time spent reading subtitles	281
8.5.3.2. Reasons for reading subtitles	284
8.5.3.3. Subtitles for vocabulary learning	285
8.6. Out-of-school exposure and learning from multimodal input	286
CHAPTER NINE – Conclusion	291
9.1. Review of findings	292
9.2. Limitations of the study	296
9.3. Directions for further research	299
9.4. Pedagogical implications	302
REFERENCES	305
APPENDIX A – Vocabulary Size Test (VST)	337
APPENDIX B – Spanish and English reading speed tests	341
APPENDIX C – Out-of-school exposure questionnaire	347
APPENDIX D – Video watching questionnaire	349
APPENDIX E – Episode-based comprehension and vocabulary tests	351
APPENDIX F – Single words and multiword units categorisation	395
APPENDIX G – Single words and multiword units per episode	399
APPENDIX H – Vocabulary profile of the <i>Curious George</i> episodes / running time of episodes	403
APPENDIX I – Transcripts from <i>Curious George</i> (E9 and 19)	405
APPENDIX J – Item type and idea units of <i>Curious George</i> Episodes	413
APPENDIX K – Difficulty and Discrimination Index	415
APPENDIX L – Idea units and items created for episode 16	427

INDEX OF TABLES

Table 3.1 – Previous research on L1 / L2 subtitling and comprehension (L1S: L1 subtitles, L2S: L2 subtitles)	54
Table 4.1 – Categorization of MWU by Granger and Paquot (2008)	80
Table 4.2 – Previous research on L1 / L2 Subtitling for L2 vocabulary learning (L1S: L1 subtitles, L2S: L2 subtitles)	111
Table 6.1 – Participants in the Study	171
Table 6.2 – Target words (TWs) and distractors in each vocabulary test.	192
Table 6.3 - Word type figures for known and unknown TW and distractors.	193
Table 6.4 – Calendar of the study	199
Table 7.1 – Descriptive statistics and Mann-Whitney results for comprehension scores	217
Table 7.2 – Descriptive statistics and Mann-Whitney results for comprehension scores (E10 and E20)	217
Table 7.3 – Descriptive statistics and Mann-Whitney results for vocabulary scores	219
Table 7.4 – Descriptive statistics and Mann-Whitney results for known words scores	221
Table 7.5 – Descriptive statistics and Mann-Whitney results for new words scores	222
Table 7.6 – Descriptive statistics and Mann-Whitney results for the vocabulary scores (E10 and E20)	222
Table 7.7 – Mann-Whitney results for known and new words scores and the known vs. new words scores	223
Table 7.8 – Descriptive statistics for aptitude, L1/L2 RS, VS and English class grade scores	224
Table 7.9 – Spearman Rho correlations between the comprehension scores and aptitude, L1/L2 RS, VS and English class grade scores	225
Table 7.10 – Spearman Rho correlations between the vocabulary scores and aptitude, L1/L2 RS, VS and English class grade scores	225
Table 7.11 – Standard regression analyses for L1S and L2S. Predictor variables: LLAMA_B, RS L1 / L2 and VS. Dependent variables: comprehension and vocabulary scores	232
Table 7.12 – ANOVA results. Predictor variable: LLAMA_B. Dependent variables: Comprehension and vocabulary scores	233
Table 7.13 – ANOVA results. Predictor variable: RS (L1RS for L1S and L2RS for L2S). Dependent variables: Comprehension and vocabulary scores	233
Table 7.14 – ANOVA results. Predictor variable: VS. Dependent variables: Comprehension and vocabulary scores	234

Table 7.15 – Raw numbers and percentages for ‘self-reported level of understanding’ at T1 / T2	237
Table 7.16 – McNemar’s results for ‘self-reported level of understanding’	237
Table 7.17 – Raw numbers and percentages for ‘viewing behaviour’ (L1S)	239
Table 7.18 – Raw numbers and percentages for ‘viewing behaviour’ (L2S)	239
Table 7.19 – McNemar’s results for ‘viewing behaviour’	240
Table 7.20 – Raw numbers and percentages for ‘time taken to read subtitles’	241
Table 7.21 – McNemar’s results for ‘time taken to read subtitles’	242
Table 7.22 – Raw numbers and percentages for ‘reasons for reading subtitles’ (L1S)	243
Table 7.23 – Raw numbers and percentages for ‘reasons for reading subtitles’ (L2S)	243
Table 7.24 – McNemar’s results on ‘reasons for reading subtitles’	244
Table 7.25 – Raw numbers and percentages for ‘subtitles and vocabulary learning’	245
Table 7.26 – <i>T</i> -tests results comparing students with and without OSE in terms of their comprehension and vocabulary scores	248

INDEX OF FIGURES

Figure 4.1 – What it means to know a word. From Nation (2001, p.27)	78
Figure 6.1 – Example of a VST Item (from Nation & Beglar, 2007)	177
Figure 6.2 – Screenshot from the results of the reading speed Spanish test	179
Figure 6.3 – Screenshot from the results of the English reading speed test	181
Figure 6.4 – LLAMA_B screenshot	182
Figure 6.5 – Item nº 1 of the T / F task in the comprehension test (E16)	188
Figure 6.6 – Item nº 4 of the MC task in the comprehension test (E16)	189
Figure 6.7 – Sequencing task in the comprehension test (E16)	189
Figure 7.1 – Mean comprehension scores for all episodes. Ovals show the significant differences between groups	218
Figure 7.2 – Mean total vocabulary scores across all 20 episodes. Ovals show significant differences between groups	220
Figure 7.3 – Descriptive data for ‘self-reported level of understanding’ at T1 / T2 in both groups	238
Figure 7.4 – Descriptive data for ‘viewing behaviour’ at T1 / T2 in both groups	240
Figure 7.5 – Descriptive data for ‘time taken to read subtitles’ at T1 / T2 in both groups	242
Figure 7.6 – Descriptive data for ‘reasons for reading subtitles’ at T1 / T2 in both groups	244
Figure 7.7 – Mean comprehension scores for the OSE and No OSE groups	248
Figure 7.8 – Mean vocabulary scores for the OSE and No OSE groups	249

INDEX OF APPENDICES

APPENDIX A – Vocabulary Size Test (VST)	337
APPENDIX B – Spanish and English reading speed tests	341
B.1 – Spanish reading speed test: text, questions and answers	341
B.2 – English reading speed test: text, questions and answers	344
APPENDIX C – Out-of-school exposure questionnaire	347
APPENDIX D – Video watching questionnaire	349
APPENDIX E – Episode-based comprehension and vocabulary tests	351
APPENDIX F – Single words and multiword units categorisation	395
F.1 – Categorisation of new target words	395
F.2 – Categorisation of known target words	396
F.3 – Categorisation of distractors	397
APPENDIX G – Single words and multiword units per episode	399
G.1 – New and known target words, multiword units and distractors per episode	399
APPENDIX H – Vocabulary profile of the <i>Curious George</i> episodes / running time of episodes	403
H.1 – Vocabulary Profile of the <i>Curious George</i> episodes	403
H.2 – Running time of the <i>Curious George</i> episodes	404
APPENDIX I – Transcripts from <i>Curious George</i> (E9 and 19)	405
APPENDIX J – Item type and idea units of <i>Curious George</i> Episodes	413
APPENDIX K – Difficulty and Discrimination Index	415
K.1 – Difficulty Index of the multiple choice exercise	415
K.2 – Percentages of Difficulty Index of the multiple choice exercise	418
K.3 – Difficulty Index of the True / False task	419
K.4 – Difficulty Index percentages of the True / False task	420
K.5 – Discrimination Index of the True / False task	421
K.6 – Discrimination Index percentages of the True / False task	422
K.7 – Discrimination Index of the multiple choice task	423
K.8 – Discrimination Index percentages of the multiple choice task	426
APPENDIX L – Idea units and items created for episode 16	427
L.1 – Idea units and items for the True / False, multiple choice and sequencing exercise in episode 16	427

INDEX OF ABBREVIATIONS

CG	Control Group
CL	Cognitive Load
CLT	Cognitive Load Theory
CTML	Cognitive Theory of Multimedia Learning
DCT	Dual Coding Theory
E	Episode
EFL	English as a Foreign Language
ER	Extensive Reading
ESL	English as a Second Language
FL	Foreign Language
LTM	Long-Term Memory
L1	First Language
L2	Second Language
L1S	L1 Subtitles
L2S	L2 Subtitles
MC	Multiple Choice
MLAT	Modern Language Aptitude Test
OSE	Out-of-School Exposure
Q	Question (in the questionnaires)
RM	Repeated Measures
RQ	Research Question
RS	Reading Speed
SLA	Second Language Acquisition
T / F	True or False
T1	Time 1
T2	Time 2
TW	Target Word
VS	Vocabulary Size
VST	Vocabulary Size test
WF	Word Family
WM	Working Memory
WPM	Words Per Minute

CHAPTER 1 – INTRODUCTION

In this chapter, information about the context of this dissertation will be provided (1.1), focusing on the role of input for L2 learning. Next, some introductory notions on using TV series for second language (L2) learning will be summarised (1.2), followed by the reasons why this thesis has been conducted (1.3). Finally, the last section presents the organisation of the thesis (1.4).

1.1. Contextualisation

Input has been studied extensively in Second Language Acquisition (SLA). According to Gass and Mackey (2007, p.177), “input is the *sine qua non* of acquisition” and “it refers to the language the learner is exposed to in a communicative context”. Research has proved it plays an essential role in L2 learning (among others, see for example Nagy et al., 1985; Laufer, 1989; 2003; Hulstijn, 1992; Coady, 1997; Nation & Wang, 1999; Pulido, 2004; VanPatten, 2004; Pigada & Schmitt, 2006).

SLA theories support that language input is needed in order to acquire the L2 (Ellis, 2008). The Input Hypothesis by Krashen claims that input is “the essential ingredient for SLA” (1985, p.4): the learner will only progress in the acquisition of the L2 if the input s/he receives is comprehensible (i.e., it contains messages that learners can understand). However, this input must also be slightly above the learners’ proficiency level ($i+1$). If input is not a bit challenging, learners’ proficiency will not develop.

According to the Affective Filter Hypothesis (Krashen, 1985), L2 learners with high motivation, low anxiety and good self-confidence are better equipped for SLA than learners who present a lack of motivation and a high anxiety level, which create an affective filter that also acts as a barrier for L2 learners, preventing them from acquiring the language. A low affective filter will facilitate input becoming intake, i.e., input that stays in the stores of long-term memory (LTM) (Pawlak, 2011).

Muñoz (2008a, 2010), among others, argues that depending on whether the L2 learning environment is naturalistic or formal, there is a huge contrast in the amount of L2 input that learners receive. She claims that in formal settings where contact with input is reduced to very few hours a week, different options to provide learners with larger amounts of authentic input should be considered. Research conducted with children in L2 formal settings has shown that language acquisition through reading (Horst et al., 1998; Paribakht & Wesche, 1999; Waring & Tatsuki, 2003; Pigada & Schmitt, 2006; Pellicer-Sánchez & Schmitt, 2010), listening (Elley, 1989; Senechal & Cornell, 1993; Van Zeeland & Schmitt, 2013) and reading-while-listening (Elley & Mangubhai, 1983; Brown et al., 2008) is favoured when a tension-free environment with a low-affective filter is provided. The rapid growth of new technologies and internet TV can also provide a vast range of comprehensible input to be used for L2 learning, both inside and outside the classroom (Rodgers & Webb, 2011).

1.2. TV series for L2 learning

The use of TV programmes for learning new languages has been considered to be positive by several researchers. Krashen in the 80s already stated that “the Input

Hypothesis makes the plausible prediction that television can be helpful if it is comprehensible” (1985, p.34). Rodgers (2013) favours the use of audiovisual materials, especially TV programmes, claiming that they contain huge quantities of input and facilitate language learning due to its entertaining and familiar content. Vanderplank (2016) also sustains that audiovisual materials are a great source of comprehensible input in foreign language (FL) learning and he also argues that the use of these materials is in line with the Input Hypothesis by Krashen (1985). Furthermore, he claims that learners’ ability to infer ideas from context and their previous linguistic and cultural knowledge while exposed to multimodal input will also aid L2 learning.

The range of audiovisual materials available in different languages, especially English, has increased in the past few years, hence providing huge quantities of authentic input ready for L2 students to use (Rodgers & Webb, 2011). People can spend many hours watching TV at home every day: the increasing accessibility of the internet globally has provided a great diversity and choice of streaming services and satellite broadcasters. Due to that, young learners and teenagers from around the world tend to watch many audiovisual materials at home. They usually access them through their own PCs or laptops. Often children, teenagers or adults find themselves hooked on their screens and they are able to binge-watch full seasons of TV series in very short periods of time. Some would also watch new releases before they are translated into the desired language: when this is the case, they may include first language (L1) or L2 subtitles (L2S) to facilitate understanding. In this way, subtitles can help to improve comprehension (as in Garza, 1991; Baltova, 1999; Huang & Eskey, 1999; Guichon & McLornan, 2008; Danan, 2004; Vanderplank, 2010; 2013; Montero-Pérez et al., 2013;

Hsu et al., 2013; Rodgers & Webb, 2017). The high demand for subtitled materials has led online streaming services such as Netflix, HBO, Amazon Prime or Disney+ to offer subtitles in different languages for most of their films and series so that viewers can easily choose the most suitable language combination for them.

Research has shown that audiovisual materials can help L2 learning (e.g., Danan, 2004). Most of the TV series and programmes offered by streaming services are filmed and produced in English-speaking countries, which makes them a great asset for language learning. Hence, the lack of L2 exposure for L2 learners in formal settings may have turned cable TV and online streaming services into a promising resource for L2 learning. They are becoming extremely popular among young and adult students, who often resort to subtitles to understand as much as possible. Even though Spain is a dubbing country and people are not used to watching subtitled movies, the demand for subtitles has increased considerably over recent decades (Danan, 2015) and many young and teenage learners are getting familiar with watching subtitled videos.

On the other hand, studies have been conducted on extensive viewing in English as a foreign language (EFL) classes (e.g., Rodgers, 2013), as it can be an engaging activity. Extensive viewing involves the regular uninterrupted use of FL television to promote language learning (Webb, 2015). Watching TV series offers many advantages for L2 learning, as it provides authentic input and access to elements that are only present in oral language such as pauses and use of informal language (Talaván, 2007; Gesa, 2019). Young low-proficiency L2 learners may also benefit from watching TV series in the FL classroom, even though many aspects that will be

considered in detail in the following chapters would need to be considered. TV series have been claimed to enhance comprehension (Rodgers & Webb, 2017) and vocabulary learning (Rodgers, 2013) but, even though research on subtitled materials has increased over the last few decades, not much attention has been paid to extensive viewing and we usually find one-off studies in the literature instead of longitudinal ones. Research on young low-proficiency L2 learners is also scarce (with some exceptions, such as Galimberti & Miralpeix, 2018; Gesa, 2019; Avello, 2023).

There are also two individual variables (ID's) that should be considered with young L2 learners' acquisition: aptitude and proficiency. The language aptitude and the low L2 proficiency level of young learners may be a limitation to their understanding of authentic multimodal input. Due to that, special care should be taken when selecting the audiovisual materials that will be used for L2 learning. Cartoons may be considered a good option for these learners, as they contain more repetition than other types of TV programme and require less cognitive processing (Bahrani & Soltani, 2011; Bahrani & Sim, 2012). However, very few longitudinal studies have been performed with cartoons and young low-proficiency learners (Avello, 2023).

Extensive viewing may lead to positive effects for L2 learners, but there is a need to conduct longitudinal studies with L2 learners watching several episodes of the same TV series in order to find out whether this practice leads to potential benefits in populations different from adult L2 learners. Eventually, familiarity with this practice introduced in the formal setting can motivate learners to do the same during leisure time, increasing out-of-school exposure (OSE) to English, especially in dubbing countries, such as Spain. In non-dubbing countries such as Sweden (Sylvén &

Sundqvist, 2012; Sundqvist & Sylvén, 2014) or Belgium (Kuppens, 2010), exposure to English through (subtitled) TV has been shown to promote L2 learning.

1.3. Motivation for the present dissertation

This doctoral dissertation aims at exploring the effects that extensive viewing of animated TV series may have on L2 learning in children (11-12 year-olds). It was decided to focus on young low-proficiency L2 learners as the body of research on this specific audience when viewing cartoons is very limited (Alexiou, 2015; Avello, 2023) and longitudinal studies conducted with other populations such as adults or teenagers offer promising results (Rodgers, 2013; Gesa, 2019). Therefore, the present study seeks to determine the impact of a pedagogical intervention in the primary classroom where participants were tested immediately after watching a cartoon episode on a weekly basis. The treatment lasted two trimesters (6 months), which allowed us to obtain longitudinal data and analyse participants' L2 comprehension and word-form recognition development through sustained exposure to multimodal input.

More specifically, this thesis seeks to determine whether there are quantitative differences regarding L2 comprehension and written word-form recognition by young low-proficiency L2 learners when watching the cartoon *Curious George*, subtitled either in the L1 or the L2. Secondly, it attempts to examine whether linguistic aptitude and L2 proficiency affect students' performance when being exposed to multimodal input. Thirdly, it wants to find out whether this treatment influences how young low-proficiency L2 learners watch subtitled TV series: i.e., whether regular watching produces any change in the way they perceive/learn from multimodal input between

the beginning and the end of the intervention. Finally, it also aims to clarify whether participants' previous OSE to multimodal input influences comprehension and vocabulary scores obtained by learners when watching this animated series.

In order to meet these objectives, our participants are young low-proficiency L2 learners in Grades 5 and 6 of primary school in Catalonia (Spain), who watch a weekly episode of the *Curious George* cartoon TV series. One group watches the series with L1 subtitles (L1S) and the other with L2S (the audio is always in English) and comprehension and written word-form recognition are assessed immediately after watching each episode over a 20-week period. The data is obtained through the episode-based comprehension and vocabulary tests specifically devised for this intervention. The possible influence of the IDs of aptitude and proficiency is examined by means of the results of the LLAMA_B tests, Reading Speed (RS) tests and Vocabulary Size Test (VST) the students took prior to the intervention, as well as their English class grade. In order to analyse the viewing experience of our participants (and see if it varied throughout the intervention), the answers from a video watching questionnaire, also specially created for the study, are analysed. Lastly, the answers provided in the OSE questionnaire are examined to see whether OSE played a role in the L2 comprehension and vocabulary results obtained by these young learners.

1.4. Organisation of the thesis

The present doctoral dissertation contains nine chapters. Chapter one has introduced the background of the thesis and presented the context and aims of the study. Chapter two provides an overview of the literature on the role of input in L2 learning and

describes the existing input modalities and well-known multimodal cognitive processing theories. It also contains a description of how media and audiovisual materials can be used in L2 learning and an overview of subtitling in L2 learning. Chapter three focuses on the effects of multimodal input on L2 comprehension and the research performed in this area. Next, chapter four centres on L2 vocabulary learning from multimodal input and the research conducted on the effects of L1/L2 subtitling on L2 word-form recognition from watching audiovisual materials in populations of different ages. The review of the literature finishes with chapter five, which presents the results on research conducted so far on the effects of aptitude and proficiency in L2 learning. Chapter six introduces the research questions (RQs) as well as the methodology of the study: participants, instruments, procedure and the data analyses conducted to give answer to our proposed RQs. Then, chapter seven presents an overview of the results, which in chapter eight are discussed and related with previous findings in the literature. Finally, the reader will find the conclusions of the study in chapter nine, together with the limitations and several suggestions for further research, followed by several pedagogical implications.

CHAPTER 2 – MULTIMODAL INPUT IN L2 LEARNING

This chapter deals with input and its role in L2 learning, with a special focus on multimodal input. We will first describe the role of authentic input for L2 learning (2.1). Next, we present an overview of different input modalities (2.2), including unimodal, bimodal and multimodal input. Subsequently, the three main theories of learning from verbal and non-verbal input which frame this dissertation will be presented (2.3), followed by a review of how media and audiovisual materials can aid L2 learning (2.4), paying special attention to cartoons, which is the input we use in this study. Finally, we will concentrate more specifically on subtitling in L2 learning (2.5): for example, on the different types of subtitles available, as well as on the possible advantages and disadvantages of subtitling that the previous literature on the topic has identified.

2.1. The role of authentic input in L2 learning

Authentic language is defined as spoken or written materials which have not been designed for language teaching purposes and can be in the form of news, films, TV series, cartoons or songs among others (Nunan, 1999). Gilmore (2007), in his state-of-the-art article on the authenticity of language learning materials, provides some arguments in favour of their use. He claims that authentic input can be found in natural environments and contains the language spoken by native speakers of the language; he also argues that audiovisual materials supply the learner with more potential input than traditional textbooks, exposing them to a vast array of language components. Textbooks are stated to portray modified input, which is an inadequate representation of the real English used by native speakers (Crystal & Davy, 1975). Several

researchers have traditionally argued that using modified instead of authentic input may hinder language learning, as it reduces the amount of linguistic and extralinguistic cues present in the input (Clarke, 1989). Modified input presented in textbooks is very different from real input in multimedia materials, but teachers need to take into account the proficiency level of the learners, their interests and needs to enable the selection of adequate materials to promote learning (Khan, 2015).

The number of studies comparing modified and authentic input is limited, although the few that do exist present favourable results for the latter (Young, 1999). One of the multiple advantages of authentic materials is that they provide great amounts of input that teachers can use to supply their students with a unique experience and help them boost their L2 learning (Shabani & Zanussi, 2015). Authentic input presents L2 learners with real life situations including some of the sounds, sights and gestures they will find when they are immersed in a real L2 environment (Berwald, 1979; 1986; Kramersch, 1983). On the other hand, research has also suggested that authentic input should be presented in its original form to L2 students at any proficiency level, avoiding any type of modifications for teaching purposes (Gilmore, 2007; Bahrani & Soltani, 2011; Benavent & Peñamaría, 2011). A large number of studies have argued that authentic input and audiovisual materials can increase young learners' interest if there is a balance between the difficulty of the input and the support provided (Pinter, 2017) and, therefore, enhance L2 learning (Baltova, 1999; Danan, 2004; Vanderplank, 2010; Ghia, 2012; Bahrani & Sim, 2012; Rodgers, 2013; Montero-Pérez et al., 2014; Becker & Sturm, 2017).

Previous research has also identified a few disadvantages when using authentic materials, and these are related to learners' proficiency. Zanon (2007) states that authentic videos pose a challenge for low-level learners, as they present real language at a normal speed. However, she claims that these low-level students can learn from easier video sequences spoken at a slower rate when exposed to real communicative conditions. Bacon and Finneman (1990) believe that audiovisual materials might be challenging for beginner learners due to their lack of familiarity with authentic materials and their proficiency level. Martínez (2002) also claims that authentic language materials may be too difficult for beginner L2 learners, considering the mixed structures and difficult vocabulary they might contain.

2.2. Input modalities

Different input modalities have drawn attention in the past few years due to the extended use of multimedia materials for SLA (Sydorenko, 2010). The availability of materials supporting L2 learners with multimedia listening activities has also raised the question of how different input modalities, including unimodal, bimodal and multimodal input, are processed by L2 pupils (Jones & Plass, 2002).

2.2.1. Unimodal and bimodal input

When L2 learners acquire the language through unimodal input, they only receive information through one channel. Research has been conducted on vocabulary learning with unimodal input through reading (Rott, 1999; Pigada & Schmitt, 2006; Brown et al., 2008) and listening (Van Zeeland & Schmitt, 2013). There is research

available comparing the efficacy of two different types of unimodal input on incidental vocabulary acquisition. For example, Vidal (2011) carried out a study with 230 university students who were divided into three groups: those reading academic texts, those listening to three lectures and a control group (CG), which only performed the tests but did not hear the lectures or read the readings. The reading and listening groups obtained significantly higher scores when comparing the pre-test and immediate post-test gains. However, when the reading and the listening conditions were compared, those students in the reading group obtained significantly higher differences, especially at low levels in the reading group. The author claims that these results might be because L2 learners can go back and re-read the words they do not understand as many times as they want, while listening is more challenging and does not offer the possibility to review the content.

Although L2 learning through unimodal input yields favourable results, it has been claimed that more remarkable improvements could be obtained by using input processed through more than one channel (Paivio, 1986). In bimodal input situations, the L2 learners receive the same input through two different channels; in the case of reading and listening, sound and text are combined so that learners can read the words at the same time as they hear them. In this way students get support in the challenging task of deciphering the meaning of the message conveyed (Bisson et al., 2014). Bimodal input has also been shown to enhance speech segmentation (Charles & Trenkic, 2015). A study where unimodal and bimodal input were examined and compared was conducted by Brown et al. (2008): they analysed lexical acquisition through reading while listening and compared it with two unimodal conditions: reading alone and listening alone. Significantly higher results were obtained for the reading-

while-listening condition compared to the listening-only mode. However, no significant differences were obtained between the reading and reading-while-listening condition, although higher results were achieved by the latter.

Other studies have been conducted comparing different bimodal input combinations, such as static images and sound, with unimodal input. Results seem to indicate that the use of images to support aural input leads to better listening comprehension in low-proficiency learners (Mueller, 1980; Maleki & Safaee Rad, 2011). Mueller (1980) analysed listening comprehension in adult beginner learners of German. Participants were divided into three groups and were exposed to audio input only (CG) or viewed a static picture for 30 seconds before or after listening to the recorded excerpt. Students who viewed the picture performed significantly better than the ones who were exposed just to the oral prompt. Maleki and Safaee Rad (2011) conducted a study on listening comprehension with low and high-proficiency English as a second language (ESL) learners. The students took three batteries of IELTS (International English Language Testing System) listening tests, including visual support, text support or no support. Results indicated that low-proficiency learners scored significantly higher when the visual input was available and high-proficiency learners obtained significantly higher results when the textual support was included in the treatment.

On the other hand, authentic videos provide extra visual input to enhance listening comprehension (Robin, 2007). There are some studies analysing the effects of video viewing, where learners get input through aural (sound) and visual (images) channels (Brett, 1997; Ockey, 2007) without any on-screen text. For example, Ockey (2007) compared the comprehension of low-intermediate, intermediate and advanced ESL

university learners who viewed still images or videos while listening to two lectures. The author concluded that watching videos leads to better comprehension than only viewing still images. In a more recent study, Tragant and Pellicer-Sanchez (2019) examined the appropriateness of two types of multimodal materials, audiobooks and video with L2S, for young learners. They concluded that participants processed both the written and the aural input, proving that both materials are appropriate for EFL learning, even though more research on comprehension is advised. In this study, though, subtitles in videos were used, which is a form of multimodal input, as will be seen in the following section.

2.2.2. Multimodal input

Multimodal input consists of the presentation of oral and written input alongside images and video. For instance, subtitled videos (either in the L1 or L2) are a recurrent example of multimodal input. It implies delivering the same information through more than two channels, which may facilitate comprehension (Plass & Jones, 2005). Multimodal input has been thought to be beneficial for FL learning (Mayer, 2005) and the great variety of audiovisual materials available nowadays offers a vast array of options.

Watching subtitled TV programmes provides the viewer with multimodal input: audio (sound) and visual information (images and written text). It has been pointed out that the presentation of input through images, sound and text at the same time helps FL learners to form mental connections and the presence of three different types of input can lead to better content processing (Halliday, 1989; Vanderplank, 1993; Danan,

2004). It has also been observed that the viewer often switches their attention from the image to the written text effortlessly and automatically (d'Ydewalle & Gielen, 1992). Research has been conducted with subtitled videos and it has usually been seen as positive for L2 comprehension (e.g., Garza, 1991; Huang & Eskey, 1999; Winke et al., 2010; Latifi et al., 2011; Etemadi, 2012; Gowhary et al., 2015) and L2 vocabulary learning (e.g., Koolstra & Beentjess, 1999; Markham, 1999; Sydorenko, 2010; Nagira, 2011; Montero-Pérez et al., 2014), although participants in these studies are often adult learners. Research on L2 comprehension and L2 vocabulary learning from multimodal input will be thoroughly described in the sections below, as they are key areas for the present dissertation.

2.3. Theories on learning from verbal and non-verbal input

During the past decades, several theories have tried to interpret the role of bimodal and multimodal input in language learning. Three main authors propose three different theories that have had a significant influence on the field. These are The Dual Coding Theory (1.3.1) by Paivio (1986), The Cognitive Load Theory (1.3.2) introduced by Sweller (1988) and the Cognitive Theory of Multimedia Learning (1.3.3) by Mayer (2009).

2.3.1. Dual Coding Theory (Paivio, 1986)

The Dual Coding Theory (DCT) deals with human cognition and it examines how information is processed by the human brain. Paivio (1986, p. 53), states that

"human cognition is unique in that it has become specialized for dealing simultaneously with language and with nonverbal objects and events. Moreover, the language system is peculiar in that it deals directly with linguistic input and output (in the form of speech or writing) while at the same time serving a symbolic function with respect to nonverbal objects, events, and behaviors. Any representational theory must accommodate this dual functionality." (p. 53)

Paivio (1986) argues that there are two independent cognitive processing subsystems, verbal and nonverbal, and he seeks to give equal weight to both of them. The verbal system deals with language and has logogens as its specific representation units, which correlate with verbal entities that operate sequentially. On the other hand, imagens are "representations from which mental images are generated under appropriate conditions" (1986, p.53) and they belong to the non-verbal system, which processes non-verbal objects. The DCT claims that, although they are processed separately, the verbal and non-verbal systems are connected and interact with each other. This interaction of the verbal and non-verbal subsystems results in better recall, which implies that input will be better processed when images and verbal information are presented together.

According to the DCT, there are three distinct types of mental processing: representational, referential and associative processing. In representational processing, the verbal or non-verbal systems are activated directly throughout the learning process whereas in referential processing, the verbal processing system

activates the non-verbal one (or vice-versa). Finally, in associative processing, representations are activated within the same verbal or non-verbal systems. Depending on the task performed, any or all of the three processes may be required. The DCT has been used to explain the outcomes of several studies with audiovisual materials, as proof that processing through different channels enhances L2 comprehension (e.g., Brett, 1997; Jones & Plass, 2002; Plass & Jones, 2005; Guichon & McLornan, 2008) and L2 vocabulary knowledge (e.g., Sydorenko, 2010).

2.3.2. Cognitive Load Theory (Sweller, 1988)

Sweller (1988; 2005) bases the Cognitive Load Theory (CLT) on Baddeley's model of working memory (WM) and argues that learning will take place if the brain's cognitive capacity is not overloaded. As the amount of items that short-term memory can hold simultaneously is limited, learning will be more effective when LTM is not overloaded. The LTM is composed of schemas, defined by Sweller as "a cognitive construct that organises the elements of information according to the manner with which they will be dealt" (1988, p.296). These schemas allow the learner to think, notice and deal with any difficulties that might arise. Some schemas are already present in the learner's LTM whereas others are acquired throughout lifelong learning processes. When new data is included in the processing system, it is arranged around this set of schemas, which leads to easier learning and processing.

In terms of proficiency, low-level learners do not have the required schemas to perform demanding tasks adequately, whereas advanced students can show their high language ability due to the appropriate schematic structures that they have already

acquired. Teachers should provide learners with specific activities that are not highly demanding in terms of WM load, and pre-teach or train their students before dealing with complex contents. Following the assumptions of the CLT, the fact that audiovisual materials offer input through three different yet complementary channels may facilitate WM load and help boost comprehension and language acquisition (Mayer, 2009). However, authors such as Kruger et al. (2013) also state that adding subtitles to videos can increase cognitive load (CL) in educational settings.

2.3.3. Cognitive Theory of Multimedia Learning (Mayer, 2009)

Mayer's Cognitive Theory of Multimedia Learning (CTML) claims that “people learn more deeply from words and pictures than from words alone” (2009, p.47) and it is based on three principles: (1) input is delivered through visual and auditory channels and processed separately, (2) these channels have a limited capacity and, (3) the language learning process can filter, select, organise and integrate the input, taking into account the previous knowledge of the learner.

Mayer distinguishes between three memory stores: sensory memory, WM and LTM. The sensory memory store can briefly (.25 of a second) hold the stimuli (images and text) received. On the other hand, input processing and integration to create mental constructs (schema) takes place in the WM store; it does not process the whole material, only a few pieces, and it usually takes less than thirty seconds (Mayer, 2010). Finally, the LTM comprises the archive of the whole learner's knowledge infinitely. That is why Sweller claimed that WM overload should be avoided.

The CTML, therefore, implies that learners boost their language learning when linguistic input is presented together with images. This combination of images and text will reinforce the previously acquired mental representations and enhance the creation of new ones. As in the DCT, assuming that there is one cognitive subsystem for words and one for pictures, it is more probable that viewers process information better when the two subsystems work at full capacity (e.g., when one is exposed to multimodal input).

2.4 Media and audiovisual materials for L2 learning

Audiovisual materials have usually been considered an entertainment or a form of distraction (Bahrani & Sim, 2012). However, recent research has shown that audiovisual materials can also be useful for L2 learning (Webb, 2015). The wide range of multimedia materials available can help teachers introduce or revise different aspects of the L2 and its culture in the classroom, and provide language learners with visual and contextual support in different tasks (Becker & Sturm, 2017). Multimedia technology has become a great asset for L2 teachers, who try to teach the language through real-life experiences and they intend to insert it as much as possible into the L2 classroom setting (Markham, 1999). Students of an L2 can also benefit from multimodal input (Sherman, 2003; Webb, 2010) to improve, for instance, their L2 listening comprehension skills (Vandergrift, 2007).

Watching TV is a very popular activity among young learners (Rice & Woodsmall, 1988) and provides them with a great amount of L2 input (Patterson, 2002; Anderson & Hanson, 2009) through hundreds of audiovisual programmes offered by an

increasing number of TV channels and streaming platforms. Viewing TV programmes can be a good way to learn the L2 (Hanf, 2015) and may also provide exceptional gains and different chances that traditional tools cannot offer for L2 learning (Lin & Siyanova, 2015). The advances in technology have facilitated young learners to have access to these programmes through the internet and their smart phones, which has increased its audience (Holmes, 2008).

Many TV programmes (e.g., news, documentaries, soap operas or cartoons) and films broadcast daily on TV channels from European countries, are originally recorded in the UK or the USA in English and so they constitute a good resource for EFL learners (Koolstra & Beenjes, 1999). Videos offer several advantages over books like speech rhythm and contextual cues, while the images of a TV programme, film or documentary can aid learners in anticipating the next scenes or infer ideas from what they see on the screen (Canning-Wilson, 2000). The potential of audiovisual materials and their usage by EFL students of any age provides rich data for research, which has focused on this topic in recent decades. Several studies suggest that they can enhance listening comprehension (Garza, 1991; Huang & Eskey, 1999; Winke et al., 2010; Latifi et al., 2011; Etemadi, 2012; Gowhary et al., 2015) and vocabulary knowledge (Neuman & Koskinen, 1992; Pavakanun & d'Ydewalle, 1992; d'Ydewalle & Pavakanun, 1997; d'Ydewalle & Van de Poel, 1999; Koolstra & Beentjes, 1999) as a result of the large amounts of spoken L2 input they contain (Webb, 2015).

Despite the advantages these materials may offer, teachers need to carefully select the most adequate ones for their students (Khan, 2015). According to Webb (2011), watching TV programmes in the L2 can be challenging when language learners start

with this practice, due to the speed and the unknown vocabulary that may appear in the videos. He argues that the initial difficulties of watching TV programmes in the L2 will be eased if the programmes are related to each other, like TV series which include several seasons with related content and recurring characters that reduce the lexical burden and expand the background knowledge of the learners (Rodgers, 2013). When L2 learners are exposed to multimodal input they enjoy, they are usually able to learn in a relaxed way (Sherman, 2003); thus, the process of acquiring the language may turn into a fun and entertaining process (Khan, 2015). However, this practice may also cause stress if the input provided is too challenging (Danan, 2004).

Several types of audiovisual material can be used as a source of input for language learning, amongst which we find cartoons. They are usually designed for a young audience and, due to that, they require less cognitive processing than other types of TV programme like news or films. We focus on cartoons in the next section, as it is the multimodal input used in the experiment in the present dissertation.

2.4.1. TV cartoons and L2 learning

Cartoons or 'animated television series' are films that use animation techniques to portray a sequence of drawings rather than objects or real people. They can be a powerful tool for L2 learning when used with FL learners who enjoy watching this type of programme (Danan, 2004). Cartoons, which usually include basic and simple vocabulary, are more accessible and entertaining than other TV genres, especially for low-level learners (Bahrani & Sim, 2012). L2 learning settings have used cartoons as a teaching tool (Bahrani, 2014; Akcan & Demirhan, 2016) and, although they were

originally considered a low form of educational entertainment, they have proved to be a potential linguistic tool (Curtis, 2015).

Bahrani (2014), who conducted research on cartoons as pedagogical tools for L2 learning, claims that these TV programmes are quite different from other audiovisual materials, as they include some exclusive features: for instance, they contain bright colors and exaggerated intonation, which attract the attention of the viewer. Another noticeable aspect is that the language of the characters includes plain and complete dialogues with constant recurrences at a low speech rate. Due to all these features, Bahrani claims that cartoons are able to attract and maintain the attention of the audience and create a relaxed atmosphere in the classroom.

Further advantages are mentioned by Bahrani and Soltani (2011). One example is that the information conveyed through cartoons would take longer to be transmitted in a written text, and the comprehension range offered by audiovisual materials would be much more limited with only written words. They also argue that exposure to cartoons helps L2 learners increase confidence and motivation, due to the low affective filter environment for language learning that these audiovisual materials create. They also suggest that exposure to cartoons helps L2 learners' memory when they try to make connections between the new information presented and their background knowledge. Koolstra and Beentjes (1999), who used a documentary to analyse vocabulary acquisition, also claimed that cartoons might be more useful for children when compared to other types of audiovisual input. Their participants were young learners attending primary school who watched one L1-subtitled episode of a documentary series. The results showed significant positive results for L2 vocabulary acquisition.

However, they argued that if the programmes in future studies were more motivating for children and included more simple language with clear images matching with text words, scores could be higher and gains more remarkable. They concluded that for this specific type of audience, cartoons might be the ideal type of audiovisual programme to enhance language learning.

Eye-tracking studies have been conducted with cartoons to study the reading behavior of children when watching these programmes (d'Ydewalle & Van Rensbergen, 1989; d'Ydewalle & De Bruycker, 2007) as well as how children process TV cartoons as compared to audio-books (Tragant & Pellicer-Sánchez, 2019). d'Ydewalle and Vanrensbergen (1989) explored how children in Grades 2, 4 and 6 of primary school were reading the L1S while watching the videos. The students watched two different cartoons, *Garfield* and *Popeye*, and their eye movement patterns while viewing the L1 subtitled cartoons were analysed. Results indicated that 8-year-old children could read the subtitles in the same way as the adults. d'Ydewalle and Bruycker (2007) examined the reading behavior of primary children attending Grades 5 and 6. Participants watched 15 minutes of a Swedish cartoon movie and the results showed no significant differences between the two groups, although fixations were longer and saccades shorter for young learners.

A more recent study by Tragant and Pellicer-Sánchez (2019) also explored young EFL learners' processing of multimodal input. The participants, who were 10-11 years old, attended Grade 5 in two different semi-private schools from Barcelona. They were divided into two groups and exposed to the same content, but it was presented in different formats. The materials selected were an episode from the series *Charlie and*

Lola, entitled *We honestly can look after your dog*, which was available as a video (cartoon) and as a picture storybook. One group read the illustrated storybook with audio support, whereas the other group watched the video episode with L2S. The eye-tracking results showed that the visual information did not distract students from processing the text, which leads the authors to claim that both materials were appropriate for L2 learning.

Other studies have focused on incidental L2 vocabulary acquisition through authentic input in the form of cartoons, without subtitles, with very young learners (Alexiou, 2015; Kokla, 2016). Alexiou (2015) analysed how Greek preschoolers (4-6 years old) recognised English vocabulary in four episodes of the British TV cartoon series *Peppa Pig*. The episodes included basic vocabulary topics such as colours, weather or food. There was no pre-teaching of the target words (TWs) before watching the episodes and their only exposure was through the videos. After watching each episode twice, the students performed an aural recognition test, where they had to point to the picture representing the word the teacher had said aloud. The promising results in this study confirmed that authentic input in the form of cartoons aided L2 vocabulary recognition, as participants identified 21 words in total, a third of the words that were presented to them in the test.

In another study in Greece, Kokla (2016) examined receptive and productive L2 vocabulary in toddlers who watched eight episodes of the American TV cartoon series *Dora the Explorer*. Although the original soundtrack was in English with some basic words in Spanish, in Greece this TV series was dubbed into Greek and the Spanish words were translated into English in order to promote EFL learning. The students

performed various vocabulary tests individually (immediate recall), which included a total of 62 items, 43 of which were tested receptively and 19 productively. Results confirmed that receptive word knowledge was enhanced by this practice and an effect of age was observed, as older learners obtained higher scores than the younger ones. However, the validity of the results has been questioned because the authors did not actually control whether participants had seen the cartoons before.

The use of subtitled cartoons for L2 learning has been examined recently by Avello (2023). She analysed vocabulary learning and the development of receptive language skills amongst young learners who were exposed to multimodal input in the EFL classroom. 120 participants from Grades 4 and 5 attending a primary school in Chile watched 11 episodes from the cartoon TV series *Charlie and Lola* and answered a written questionnaire and a semi-structured interview at the end of the treatment. Results showed significant L2 gains from extensive subtitled video viewing.

During the 1990s, some research was also conducted using cartoons to analyse vocabulary acquisition in adults (Pavakanun & d'Ydewalle, 1992; d'Ydewalle & Pavakanun, 1995; 1997), as animated TV series, originally addressed to children (Curtis, 2015), are said to be easy to understand by L2 learners of all ages and proficiency levels (Sherman, 2003). In these studies, researchers analysed the vocabulary knowledge of adult L2 learners who watched cartoons (approximately 15 minutes long) with L1 or L2 subtitles and, immediately afterwards, they were tested on vocabulary recognition. The results showed incidental L2 vocabulary acquisition had taken place through watching cartoons with either L1 or L2 subtitles.

There are also studies comparing the effects of watching cartoons and other types of subtitled TV programmes on adults' L2 learning, such as Bravo (2008) or Bahrani and Sim (2012). Bravo (2008) conducted two studies in her PhD thesis: in the first one she compared the effects of watching authentic audiovisual materials from different genres on L2 comprehension and vocabulary learning. The 32 adult L2 Portuguese learners at a university, with an average age of 23, watched 12 videos, six with L2S and six without subtitles. The episodes the participants watched ranged from 4 to 6 minutes and they belonged to the following categories: news, feature film, musical, documentary, soap opera and animated cartoons. The L2 learners took comprehension and vocabulary tests in the L2 (Portuguese) after watching the videos (it was not possible to find a lingua franca due to the diversity encountered in the L1 of the participants). The results showed higher significant differences for the students watching the videos with L2S in all genres, except for animated cartoons. In this particular case, the scores were slightly higher for the no subtitles group, although there were no significant differences when compared with the L2S group. The author explains these results arguing that due to the specific features of cartoons, such as the attractiveness of the images and its slow pace, cartoons provided a better understanding than other TV genres and the participants relied more on watching the images than on reading the subtitles.

In Bahrani and Sim (2012), 60 low-proficiency adult learners were divided into three groups. Each group consistently watched news, films or cartoons without subtitles once a week for four months. The participants took a sample of the IELTS language proficiency test as a pre and post-test to check if there was any improvement in their L2 knowledge. The group watching cartoons proved to have the highest gains,

followed by the group who watched films: those watching TV news did not obtain any significant gains. Results from these studies show that TV cartoons can be effective in enhancing proficiency for low-level adult learners, too.

Finally, it should be acknowledged that in recent decades specially designed cartoons for adults have appeared and TV series like *The Simpsons* or *Family Guy* have encountered great success. The effect of watching these TV programmes (with or without subtitles) on L2 learning has been analysed in several studies, such as Karakas and Sariçoban (2012), Vulchanova et al. (2015) and Peters et al. (2016). Karakas and Sariçoban (2012) performed a study on vocabulary acquisition with upper-intermediate adult EFL students watching two episodes of *Family Guy*. The scores did not show any significant differences between the L2S and the no subtitles group. Vulchanova et al. (2015) used the same cartoon as Karakas and Sariçoban and analysed comprehension with 16 intermediate and 17 advanced ESL students after they watched one episode of the same American TV series. The results showed significant differences in favour of the L1S and L2S groups when compared to the no-subtitles group.

Peters et al. (2016) examined the effect of L1S and L2S on different aspects of L2 vocabulary learning at different proficiency levels (beginner, pre-intermediate and intermediate adult participants). In their research, adult ESL students of a vocational school watched one episode of “The Simpsons” with L1S or L2S. The L2 learners were tested on different vocabulary measures with the scores indicating that the L2S group gained more than the group watching the video with L1S for written form recall and written form recognition, although no significant differences were obtained for the

latter. The authors highlight the need to perform more studies to determine the best type of audiovisual input (e.g., cartoons, news, films) for incidental FL vocabulary learning. Indeed, studies with sustained exposure to TV programmes in the L2 classroom setting with young populations are scarce and further research on this area is needed.

2.5. Subtitling for L2 learning

Subtitles (also called ‘Closed Captions’ in North America and ‘Teletext Subtitles’ in the UK) were first devised to offer an aid for deaf and hearing-impaired people to understand videos that were only presented with visual and aural input (Vanderplank, 1999). The study on the use of subtitling by the European Commission (2011) claims that subtitles can also be beneficial and stimulating for L2 learning. Furthermore, they are also being used in some European countries (e.g., Sweden, Finland, The Netherlands, Denmark, Malta, Portugal), as an alternative to dubbing. On the other hand, countries such as Spain, France, Switzerland, Austria, Germany and Italy use dubbing in broadcasting programmes.

Subtitles are presented in a fixed spatial position at the bottom of the screen and they are usually displayed in two lines for a limited amount of time (usually for six seconds). Attention also needs to be paid to the visual input that appears on the screen while the learner is reading. There are some rules that should be followed when adding subtitles: there should be up to two lines with 32 characters and spaces each (the maximum permitted) and they should appear on the screen for about six seconds (d’Ydewalle & Bruycker, 2007). However, this might vary if the video is addressed to children who

are still learning how to read and, in that case, an ‘eight-second rule’ might be applied (Koolstra et al., 2002). Rules have not been established for the speed of subtitles in programmes addressed to children, although it has been recommended to present them at a slower pace due to their poorer reading skills (Fresno, 2018).

The European Commission report *Multilingualism: An Asset for Europe and a Shared Commitment* (2008) claims that media could aid language learning in an informal way. Even though Spain is referred to as a dubbing country in the latest European Commission Report on the use of subtitling (European Commission, 2011), the increased accessibility and availability of subtitled videos in European countries (including Spain) over the past few years has made it easier to watch the same programme while being exposed to several languages. In addition, there is an increasing number of viewers who choose to view TV series and films in an L2 with L1S (Birulés-Muntané & Soto-Faraco, 2016). Muñoz (2017) argues that students, named ‘Digital Natives’ by Prensky (2001), are used to working with and experiencing digital technologies in their everyday lives; so they will also benefit from subtitled audiovisual programmes when shown in formal L2 settings, improving their language proficiency levels.

In the past few years, publishing houses such as Oxford University Press have developed instructional audiovisual materials which contain the option to include textual support in the form of subtitles. Due to that, teachers can now include subtitles when teaching an L2. In addition, textbooks for primary schools like the “Incredible English Kit” and “Ace” series include cartoons of the stories illustrated in the books, which may facilitate young learners’ understanding of the L2 content.

There are different kinds of on-screen text: L1S (L2 audio, L1S) and L2S (L2 audio and subtitles). Other types of on-screen text include reversed subtitles (L1 audio, L2 subtitles) and keyword captioning (when just key words appear as subtitles). Zanon (2007) claims that although L1S are the most commonly used subtitling mode, L2S are frequently used in classroom settings. The use of L1S, L2S and keyword subtitles for language learning will be more deeply discussed in the sections that follow.

2.5.1. L1 Subtitles (L1S)

L1S can also be referred to as ‘standard’ or ‘interlingual’ subtitles. Watching a video with L1S implies that the soundtrack is in the L2, whereas subtitles provided through on-screen text are in the L1 (Lertola, 2015). L1S are claimed to help L2 students in figuring out the meaning of unknown L2 words, hardly understandable through listening alone, and improving comprehension (Caimi, 2006). Although the presence of the L1 in the written input can boost L2 learners’ confidence, it has also been pointed out that learners may not listen to the input in the L2, due to their focused attention on reading the text with L1S (Danan, 2004). However, Koolstra and Beentjes (1999), among others, find L1S appropriate at low levels, as they help learners accessing input that would not be understandable otherwise.

2.5.2. L2 Subtitles (L2S)

L2S, also referred to as ‘captions’, ‘bimodal’ or ‘intralingual’ subtitles include the L2 in both the soundtrack and the subtitles (L2S + L2 soundtrack). King (2002) lists several advantages for L2S: they strengthen L2 comprehension of context-bound expressions

and vocabulary learning, support the students in understanding the plot more easily and can reinforce the learners' aural input with the written form of the words that appear in the subtitles. Caimi (2006) expands on the benefits proposed by King and claims that L2S aid in understanding the L2 cultural and linguistic context: "the contact with a different linguistic and cultural context opens a window on cultural self-definition, which is processed in relation to what is perceived as different from one's own cultural identity" (p. 4).

L2S are claimed to enhance different aspects of L2 learning: Vanderplank (2016) suggests that L2S are a good resource for enhancing our reading or listening skills and Etemadi (2012) concludes that pronunciation can also be improved when L2S are included in the input. Other important benefits of L2S include the assistance they provide to L2 learners in segmenting the speech stream and identifying words as separate units (Danan, 2004). L2S can also help word recognition when the word is presented with written (subtitles) and aural input in the same language and, due to that, L2 learners can become more aware of the words presented to them (Garza, 1991; Peters et al., 2016). Research with L2S has argued that they are more useful than other types of subtitles in order to understand certain details about the characters or the story line (Chung, 1999). Although they might be more difficult for L2 learners, research has proved that L2S may also help comprehension (Garza, 1991; Huang & Eskey, 1999; Bravo, 2008; Winke et al., 2010; Etemadi, 2012; Rodgers, 2013; Gowhary et al., 2015; Ebrahimi & Bazaee, 2016) and L2 vocabulary recognition (Neuman & Koskinen, 1992; Markham, 1999; Sydorenko, 2010). Danan (2004) has pointed out that L2S will only be appropriate for low-proficiency L2 learners if the materials are carefully selected and adapted.

2.5.3. L1 Subtitles (L1S) vs. L2 Subtitles (L2S)

A strong need to find the best type of subtitles for SLA surged in the 2000s, coinciding with a wider use of the internet around the world, and the number of studies comparing L1S and L2S has increased since then (Matielo et al., 2015). However, results are still inconclusive and further research is needed (Garza, 1991; Borrás & Lafayette, 1994; Bianchi & Ciabattini, 2008; Winke et al., 2010; Sydorenko, 2010; Rodgers, 2013; Frumuselu et al., 2015; Vulchanova et al. 2015; Peters et al., 2016). Several authors have claimed that the proficiency variable is determinant in deciding which is the best type of subtitles for beginner, intermediate and advanced learners (d'Ydewalle & Pavakanun, 1995; Danan, 2004; Vanderplank, 2010), although the amount of studies where different proficiency levels are taken into account is scarce (Montero-Pérez et al., 2013). Danan (2004) claims that L1S might be better suited for low-level learners and L2S for advanced students. Lin and Siyanova (2015) support Danan in favouring the use of L1S to enhance comprehension in beginner learners and they recommend a progressive move to L2S when the proficiency level increases. There is also a study by Lavaur and Bairstow (2011), which analysed L2 comprehension in teenagers with different proficiency levels and claimed that both L1S and L2S were too challenging for beginner learners. Intermediate learners did not obtain higher results when subtitles were added to any of the film versions in the study and the advanced learners did not need L1S or L2S due to their high proficiency level.

Studies comparing the effects of L1S and L2S on L2 comprehension in adult learners have found mixed results: some have found significant results with adult and intermediate level learners in favour of L2S (Hayati & Mohmedi, 2011; Birulés-Muntané

& Soto-Faraco, 2016) and others for L1S (Markham et al., 2001; Markham & Peter, 2003; Bianchi & Ciabattini, 2008; Latifi et al., 2011). Other studies also performed with intermediate adult learners have found no significant differences between the two conditions (Guichon & McLornan, 2008; Basaran & Köse, 2012). In another study, Vulchanova et al. (2015) obtained no significant differences between L1S and L2S with adult intermediate and advanced-proficiency students. Most of the studies mentioned above analysing L2 comprehension when watching audiovisual materials with L1S and L2S used short videos for the treatment (maximum 15 min. long) in a very limited number of sessions, which ranged from one to six (Markham et al., 2001; Markham & Peter, 2003; Guichon & McLornan, 2008; Bianchi & Ciabattini, 2008; Winke et al., 2010; Sydorenko, 2010; Hayati & Mohmedi, 2011; Latifi et al., 2011; Birulés-Muntané & Soto-Faraco, 2016).

Vocabulary acquisition from subtitled videos (either with L1S or L2S) with adult L2 learners has also been examined (Stewart & Pertusa, 2004; Hui, 2007; Frumuselu et al., 2015; Peters et al., 2016; Galimberti & Miralpeix, 2018) and most of the studies have obtained favourable results in favour of L2S. Hui (2007) analysed word recognition in 182 adult Chinese ESL learners (90 advanced and 92 beginners) at an Institute of Technology in China. The participants watched a 16-minute National Geographic documentary with L1S, L2S and no subtitles and the results were significant in favour of the L2S group. Frumuselu et al., (2015) conducted their longitudinal study with adult beginner, intermediate and advanced L2 learners who watched thirteen episodes from the popular TV series *Friends*. They took a multiple choice (MC) vocabulary test and answered fifteen open questions to assess word meaning. Results showed significant differences for the L2S group, who scored higher.

Peters et al. (2016) performed two different studies; in the first one, they selected teenagers who watched an excerpt of a 13-minute documentary and took a spoken form recognition and meaning recall test. In the second study, the adult participants with a proficiency level ranging from beginner to intermediate watched one episode from *The Simpsons* (20 minutes) and they were tested on written form recall and written form recognition through MC tests. The L2S group scored significantly higher on the test and no significant differences were observed for the second.

Even if watching subtitled TV programmes provides huge quantities of input, which is beneficial for L2 learners (Danan, 2004; Webb & Rodgers, 2009a; Vanderplank, 2010; 2013, Winke et al., 2013; Birulés-Muntané & Soto-Faraco, 2016), it is not yet clear which subtitle condition (L1/L2) is most appropriate for different proficiency learners in different areas (Garza, 1991; Borrás & Lafayette, 1994; Biachi & Ciabattini, 2008; Hsu et al., 2013; Vulchanova et al., 2015) and further research is needed, especially with beginner learners (Matielo et al., 2015). Moreover, the diversity of the measures and tests used in these studies makes it hard to obtain generalisable and reliable results, either for L2 vocabulary acquisition or listening comprehension (Montero-Pérez & Desmet, 2012). We will focus on L1/L2 subtitles in multimodal input for learners at different proficiency levels in sections 2.2 (listening comprehension) and 3.4 (vocabulary learning).

2.5.4. Keyword subtitles

Garza (1991, p.246) explains that keyword subtitles, instead of providing the transcription of the full text, just supply “the essential word or words in an utterance”

which convey the meaning of the input. They have been said to reduce the CL for low-proficiency learners. Some authors like Vanderplank (2016) have pointed out that this subtitling mode has some relevant disadvantages. He emphasises the deaf community's rejection of them in the 1970s and he also insists that they imply extra work for anyone who wishes to use them. However, it has also been claimed that keyword subtitles reduce textual density (Guillory, 1998). The effect of keywords instead of full subtitles for L2 learning has been analysed in some studies, although findings are inconclusive, for example, Guillory (1998) and Montero-Pérez et al. (2014).

In Guillory (1998), the 202 French adult beginner participants were divided into three groups and watched a videotape from a textbook with full subtitles (L2S), keyword subtitles and without any written text support. The students were tested immediately after watching the videos through fourteen short-answer content comprehension questions (seven for each clip). Even if the full subtitles group obtained higher results than the keyword subtitles one, significant differences were not found.

A more recent study analysing the effects of keyword subtitles was conducted by Montero Pérez et al. (2014). They analysed comprehension and vocabulary acquisition in 133 adult upper-intermediate French students who watched three short videos from a current affairs programme. They were divided into four groups and watched the materials with L2S, L2S with highlighted words, keyword subtitles or without subtitles. The results for L2 comprehension showed similar scores between groups and no significant differences were obtained. The authors claimed that this could be due to the fact that the questions were not appropriate for the learners.

However, it is important to note that the scores were not very high, implying that the content was hard to understand by participants. Regarding vocabulary, the results for the groups watching the videos with L2S, L2S with highlighted words and keyword subtitles were higher than the no subtitles group, showing that subtitles provide crucial support in the enhancement of L2 word recognition, but not proving the efficacy of keyword subtitles.

2.5.5. Advantages and disadvantages of subtitling

The increasing availability of subtitled TV programmes and films through TV, DVDs, online streaming and video on-demand has prompted their use for L2 learning. Autonomous learners can benefit from subtitles, as they present extra support to control and understand input in the L2 (Koolstra & Beentjes, 1999; Danan, 2015). When videos are used in an instructional setting, a more dynamic environment can be created. Learners may acquire the L2 in a relaxed way, more similar to informal settings (Matielo et al., 2015). There is also an important number of teachers who encourage their students to watch subtitled videos at home in order to improve their language acquisition (Shabani & Zanussi, 2015).

Hanf (2015) enumerates some benefits of subtitled materials: for example, they can enhance language awareness and provide instant feedback to FL learners, who are able to see the written transcription of words or expressions they are hearing. He asserts that subtitles aid L2 learners in decreasing their anxiety level, which is accomplished when there is a low affective filter. Thus, the L2 classroom learner may feel more comfortable watching subtitled movies than answering questions from the

teacher, a very common situation with low proficiency levels (Gowhary et al., 2015). Hence, subtitles present the audience with additional linguistic cues (Becker & Sturm, 2017). The increasing use of subtitled videos might be motivated by other reasons than only L2 acquisition (Neuman & Koskinen, 1999). De Bot et al. (1986) argue that subtitles can also be helpful for maintaining the L2 proficiency level when included in television programmes. They acknowledge that the exposure of L2 learners to subtitled audiovisual programmes will provide the necessary contact with the L2, which will allow them to preserve their proficiency level.

Redundancy and authenticity are among the advantages recounted by Koolstra et al. (2002) for subtitling. They claim that multimodal input offers redundancy, due to the L2 learners' exposure to three different input modalities that include the same information and, at the same time, benefit learning. Another asset is that authentic input offers access to real language through the original voices of the actors while the written transfer of the spoken dialogue appears at the bottom of the screen without voices being dubbed.

However, Hatim and Mason (2000) also list some facts that need to be taken into account when dealing with subtitled videos. They argue that when the information from the soundtrack is transcribed into the subtitles, there is a loss of information due to the absence of certain speech characteristics that will not be depicted, such as non-standard dialect, emphatic devices (intonation), code-switching, style-shifting and turn-taking. Another issue is the limited physical space available and the pace of the dialogue presented in the soundtrack, which can make it challenging to include all the information in the subtitles and match it with the visual input that appears on the screen

at that moment; all the missing information that subtitles are not able to provide can be conveyed through the aural and the visual input (Bravo, 2008). It has also been suggested by Koolstra et al. (2002) that another disadvantage is that subtitles in any language can often be considered distracting: they can aesthetically harm the images of the video shown on the monitor, as subtitles will cover a part of it, and it has been claimed that if L2 students are focused on reading the subtitles, they may not see the important actions that occur in the images (Borrás & Lafayette, 1994; Diaz-Cintas & Remael, 2007).

Even if the use of subtitles with children is controversial, it has been seen that young learners can also benefit from watching subtitled audiovisual materials as multimodal input, considering that the presence of subtitles can increase vocabulary acquisition (Koolstra & Beentjes, 1999) and improve word decoding (Koolstra et al, 1997). Lambert and Holobow (1984) point out that a key feature for success in L2 acquisition with subtitled materials is that these L2 learners need to be familiar with watching TV with subtitles in order to avoid distractions that would compromise the benefits. It has been claimed that children might not have the required reading skills to follow the subtitles due to their speed (Van Lil, 1988), although eye-movement studies have proved that subtitles are automatically read and processed by both young and adult learners (d'Ydewalle et al., 1987; Pavakanun & d'Ydewalle, 1992; d'Ydewalle & Gielen, 1992; Tragant & Pellicer-Sánchez, 2019), as we will see in the next section.

2.5.6. Reading behaviour with subtitles

Viewers' reading behaviour with subtitles has been examined because there is a desire to find out how L2 learners process subtitles (whether they read them fully or only partly) and in determining where and how learners focus their attention when they are presented with visual (images and text) and aural input (sound) at the same time (Winke et al., 2010). Eye-movements are defined by Frenck-Mestre (2005, p.1) as “the complex trace of saccades, fixations and regressions that the eyes make while taking in a line of text”. She argues that eye-movement recordings are considered the richest record for understanding the cognitive processing of reading. Research has seen that learners are able to process, assimilate and recall the information coming from the three different channels that multimodal input offers (Pavakanun & d'Ydewalle, 1992). This idea is supported by d'Ydewalle and Gielen (1992, p.425), who argue that when subtitles are presented together with the soundtrack, the L2 learners split their attention between both conditions, using smart strategies to integrate the information: “when people watch TV, the distribution of attention between different channels of information turns out to be an effortless process. Viewers seem to have developed a strategy that allows them to process these channels without problems and in which reading the subtitles occupies a major place”. However, they also claim that more time is usually devoted to processing the subtitles due to the complexity of this type of input.

Gielen (1988) states that if subtitles are displayed in different genres of audiovisual programmes (films and news reports), they are processed and read even when they are not needed to understand the meaning of the input. On the other hand, Garza (1991) claims that the presence of subtitles aids the connection between the sound

and the images. Furthermore, subtitles assist in deciphering dubious input and recognising unknown words or phrases (Neuman & Koskinen, 1992; Danan, 2004; Winke et al., 2010). It has been proved that when watching subtitled audiovisual materials with L1S and L2 soundtrack, this information is processed almost in parallel, which may lead to L2 acquisition (Danan, 1992; D'Ydewalle & Van de Poel, 1999; Koolstra & Beentjes, 1999; Danan, 2004; Perego et al, 2010).

As has been previously mentioned, eye-movement studies have proved that reading subtitles is an automatic behavior (d'Ydewalle & Gielen, 1992; Pavakanun & d'Ydewalle, 1992) and children are able to read them in the same way as adults (d'Ydewalle & Vanrensbergen, 1989; d'Ydewalle & Bruycker, 2007). However, in a recent pioneering eye-tracking study with young low-proficiency and intermediate learners carried out by Muñoz (2017), results indicated that children (beginners), due to their lower L2 level compared to adults, required more fixations on L2S than L1S, although both types of subtitles were attended to. Beginner learners needed to focus more on both L1S and L2S, due to their low proficiency level, although L1S received less fixations. Intermediate learners resembled beginners and also spent more time on L2S than on L1S, although the numbers were lower compared to the beginner group. The scores indicated that proficiency and age mattered for reading behavior with subtitled videos.

CHAPTER 3 – MULTIMODAL INPUT AND L2 COMPREHENSION

In this chapter, we present previous research on multimodal input and L2 comprehension. We will first provide some context with some general previous considerations on listening comprehension in an L2 (3.1.1), followed by a description of the processes employed for L2 comprehension: bottom-up and top-down. Subsequently, we will focus on the most typical ways in which comprehension has been assessed. Next, studies on L2 comprehension from reading and listening are introduced (3.1.2) in an attempt to provide the background for the most recent studies on multimodal input and L2 comprehension, where simultaneous reading and listening are involved. Then, we focus on studies on L1/L2 subtitling for L2 comprehension from videos, with special attention to those dealing with adults (3.2.1), adolescents (3.2.2) and young learners (3.2.3), as the findings in this previous research are relevant for the present study.

3.1. Previous Considerations

3.1.1. Listening comprehension in the L2

As we have seen earlier, the Input Hypothesis by Krashen (1985) claims that in order to acquire an L2, the learner needs to be provided with comprehensible input that is a little beyond their proficiency level ($i + 1$). Listening and reading input are processed through cognitive operations which help learners build mental representations of the information received (Chun & Plass, 1997). Furthermore, listening is an inferential process in which the learner constructs meaning through linguistic and non-linguistic

sources (Buck, 2001; Vandergrift, 2007), although many factors can make comprehension challenging (Pinter, 2017).

Listening comprehension is an active cognitive process where learners guess and infer meaning (Noblitt, 1995). Difficulties in the listening area could be due to the short-term memory of the learners, which impedes them from decoding and understanding everything they have just heard, as new input is provided in real time (Goh, 2000). Listening is temporal and not controlled by the student but by the speaker (Brown, 2001). When listening, L2 learners are only provided with aural input, which implies that they must deal with the challenging procedure of understanding input instantly (Staer, 2009). On the other hand, a written text facilitates comprehension as learners can read the printed words as many times as they need to understand it (Garza, 1991).

In order to develop and improve L2 listening comprehension, teachers usually include listening activities in their classes. However, these tasks often contain only aural input that leads to poor comprehension (Jones & Plass, 2002). The cognitive processes involved in bottom-up and top-down listening comprehension are crucial to understanding how L2 learners use their strategies in order to acquire an L2.

3.1.1.1. Bottom-up and top-down processing in listening comprehension

In order to interpret an oral message, linguistic sources (phonological, lexical, syntactic, semantic and pragmatic) as well as non-linguistic sources (topic, world and context knowledge) are used to infer meaning (Buck, 2001). This variety of knowledge types is employed using bottom-up and top-down processes for comprehension. A

bottom-up process is a type of decoding where the sound stream is automatically divided into significant parts to understand the message conveyed to language learners, who will rely on their linguistic knowledge to guess the context. On the other hand, a top-down process is a perception mechanism that uses non-linguistic sources such as background knowledge and context to understand the meaning of the input (Vandergrift & Goh, 2012). If these two processes do not interact with each other, learners may have problems in understanding the message conveyed (Staer, 2009). The use of top-down and bottom-up strategies is assessed in comprehension tests, which measure topic and understanding (Rodgers, 2013).

Goh (2000) claims that this harmonious relationship between bottom-up and top-down processes can be affected by the proficiency level of the learners. She argues that those with a modest language ability will tend to use bottom-up processes, whereas more skilled learners will focus on top-down mechanisms. The scarce ability that low-proficiency learners have in the L2 makes them face serious difficulties when processing listening input, as they can only use lexical or syntactic units (Rubin, 1994). This is problematic in young learners, who face a challenging task when trying to decipher aural input due to their early age, their lack of prior knowledge of the target language and its cultural background, which lead them to employ bottom-up processes better. To sum up, children might have more difficulty applying linguistic and non-linguistic strategies effectively in discourse comprehension than older learners (Pinter, 2017).

Support for low-level learners to enhance L2 comprehension can be provided by imagery, (e.g., included in audiovisual materials), which will supply non-linguistic input

and background knowledge to these less skilled learners. Hence, the use of top-down strategies will be more easily activated (Vandergrift, 2007; Rodgers, 2013). Research analysing L2 comprehension with visual and textual support has reported positive results (Brett, 1997; Jones & Plass, 2002; Guichon & McLornan, 2008). Furthermore, experiments performed with multimodal input comparing the L2 comprehension scores while watching videos with and without subtitles have obtained better results for subtitled videos (Garza, 1991; Borrás & Lafayette, 1994; Markham & Peter, 2003; Danan, 2004; Winke et al., 2010; Etemadi, 2012; Rodgers, 2013; Gowhary et al., 2015; Ebrahimi & Bazae, 2016).

3.1.1.2. Measuring listening comprehension

There are various features of listening comprehension that can be measured. According to Buck (2001, p.14), listening comprehension comprises three abilities:

- “processing extended samples of realistic spoken language, automatically and in real time,
- understanding the linguistic information that is unequivocally included in the text, and
- making whatever inferences are unambiguously implicated by the content of the passage”.

Taking into account these three aspects, Rodgers (2013) suggests that to obtain useful data on comprehension, optimal tasks can be true or false (T / F) questions and MC

questions, as well as item sequencing. Students are used to these types of question and both top-down and bottom-up processes are assessed when answering T / F and MC questions (Richards, 1990).

Children learning an L2 are characterized differently from adult learners, even if they are at the same proficiency level. This needs to be taken into consideration when designing comprehension tests. Pinter (2017) claims that when designing tasks for young learners, their linguistic knowledge should be considered. She also suggests that these students should be provided with not too challenging tasks (such as translating), because more demanding tasks might cause demotivation when performing the tests. She affirms that the best kick-off for children would be a 'listen and do' activity, which is what MC, T / F tests and item sequencing exercises offer, as they are receptive tests without high demands on production by low-proficiency learners.

3.1.1.2.1. Multiple choice (MC) tests

MC tests have been widely used in comprehension studies (Huang & Eskey, 1999; Markham et al., 2001; Markham & Peter, 2003; Bianchi & Ciabattini, 2008; Winke et al., 2010; Hayati & Mohmedi, 2011; Latifi et al., 2011; Etemadi, 2012; Başaran & Köse, 2012; Rodgers, 2013; Vulchanova et al., 2015; Gowhary et al., 2015; Ebrahimi & Bazaee, 2016; Galimberti & Miralpeix, 2018). The most typical MC tests contain a stem (the question), followed by a number of possible options, which include the correct answer (usually one) and the distractors (Bradbard et al., 2004). When deciding how many possible options are needed for each MC question, Rodríguez (2005)

recommends that only the necessary distractors should be added for each question, as introducing many wrong options may lead to the inclusion of information that might be relevant for answering other questions on the same test. Rodgers (2013) argues that three options are enough for test takers, who will be more confused if this amount is increased.

Although MC tests have been criticised for not portraying the natural settings of listening (Buck, 2001; Hughes, 2003), the practical advantages that they offer like "high marker reliability, ease of marking and objective scoring" turns them into widely-used tests for the analysis of listening comprehension (Hemmati & Ghaderi, 2014, p.639). According to Spaan (2007), these tests are useful when evaluating the comprehension of large amounts of input and they are thought to be more reliable than tests with open-ended questions (In'nami & Koizumi, 2009), which can lead to different scores depending on the researcher who is evaluating the answers (Brindley, 1998).

3.1.1.2.2. Item sequencing

Item sequencing is a task in which the learner is required to place a number of items into the 'correct' order. It aims at finding out if the participants have fully understood the text as a whole, measuring its overall comprehension (Richards, 1983). It is essential to understand the order of events in listening comprehension and that is the reason why this type of test is needed when testing listening skills (Brett, 1995). In the case of reading, the comprehension of a text implies that a competent reader is able to identify how the ideas are related with each other and also its order in the story line (Alderson & Banerjee, 2002). Exercises on sequencing items were

also used by Rodgers (2013) and Gesa (2019) to test learners' listening skills and evaluate their competence, and proved to be useful in the assessment of content comprehension in multimodal input.

3.1.1.2.3. True / False (T / F) questions

A T / F question consists of a statement that is either correct or incorrect and it requires a true or false response. Other variations such as: “yes” or “no”, “correct” or “incorrect”, and “agree” or “disagree” statements are also used in the literature. Ebel and Frisbie (1991, p.135) define T / F items as “simple declarative sentences of the kind that make up most oral and written communications”. They enumerate the main arguments in favour of using this type of exercise to collect data, arguing that “all verbal knowledge can be expressed in propositions” and that “a proposition in any sentence can be said to be true or false”.

However, this type of exercise has been claimed to be “especially susceptible to guessing” (Ebel, 1975, p.2), although empirical data has provided reliable test results that contradict this belief. T / F questions are frequently used in SLA textbooks (Haladyna et al., 2002) and L2 students are used to performing them in L2 classes and exams. Rodgers (2013) claims that they are a quick and effective way to test L2 learning and obtain a reliable measurement of comprehension.

3.1.2. Bimodal input and L2 comprehension

Bimodal input entails processing information coming from two different channels and it can be described as the simultaneous presentation of text and sound. Learners can be reading a book while listening to its audio file or someone can read it aloud to them. In the literature, dynamic video pictures have proved to be more effective than static pictures accompanied by sound when the aim is to help comprehension (Herron et al., 1995; Ockey, 2007).

Research on bimodal input has been mainly conducted in studies analysing reading while listening or watching still pictures while reading (Mueller, 1980; Ockey, 2007; Brown et al., 2008; Maleki & Safaee Rad, 2011). Mueller (1980) analysed bimodal and unimodal input on listening comprehension in adult beginner learners of German. The three groups of participants were exposed to audio input only or viewed a static picture before or after listening to the recorded excerpt. Students who were exposed only to bimodal input performed significantly better than the ones who were exposed to unimodal input. Ockey (2007) examined the comprehension of 2 low-intermediate, 2 intermediate and 2 advanced adult ESL university learners. The participants viewed either still images or videos while listening to two lectures. After analysing the results, the author determined that comprehension was better when watching videos rather than viewing still images.

Furthermore, Brown et al. (2008) analysed the vocabulary acquisition of 35 pre-intermediate to intermediate adult Japanese students attending a university in Japan. The three groups of participants were in one of these three conditions: reading,

reading-while-listening or listening to three stories in a graded reader form. The reading-while-listening condition obtained significantly higher results than the listening-only mode, whereas no significant differences between the reading and the reading-while-listening condition were obtained. The outcomes also indicated a significant decrease in time in retention time for the reading and listening conditions but not for the reading-while-listening one, which implied that bimodal input may provide a longer retention than unimodal input. Maleki and Safaee Rad (2011) analysed the effect of unimodal and bimodal input on listening comprehension with low and high proficiency ESL learners. The students took three batteries of IELTS (International English Language Testing System) listening tests, including visual, text or no support in the assessment. The results showed that both low and high-proficiency participants scored significantly higher when exposed to bimodal input.

In recent years, the introduction of audiobooks in EFL classrooms has also attracted the attention of researchers in order to find out if they are useful for EFL learning. A study by Tragant and Pellicer-Sanchez (2019) examined the appropriateness of bimodal input in the form of audiobooks with young learners. The results evidenced that the participants processed both the written and the aural input and confirmed that bimodal input is suitable for EFL learning.

Finally, several studies comparing the effects of watching videos with and without subtitles (i.e. sound + image or sound + image + text) have obtained higher results for subtitled videos in L2 comprehension (Garza, 1991; Borrás & Lafayette, 1994; Markham & Peter, 2003; Danan, 2004; Winke et al., 2010; Etemadi, 2012; Rodgers, 2013; Gowhary et al., 2015; Ebrahimi & Bazaei, 2016) and L2 vocabulary acquisition

(Pavakanun & d'Ydewalle, 1992; Neuman & Koskinen, 1992; Koolstra & Beentjes, 1999; Stewart & Pertusa, 2004; Winke et al., 2010; Sydorenko, 2010; Frumuselu et al., 2015; Peters et al., 2016).

These outcomes lead us to conclude that even though bimodal input is beneficial for L2 learning, the inclusion of multiple input types (sound + image + text) can lead to more favourable results. Watching subtitled TV programmes or films may entail a better understanding than only listening to text (or listening + reading in audiobooks), as Webb and Rodgers (2009a) also notice.

3.2. Multimodal input and L2 comprehension

When L2 learners watch a TV programme they assemble the understanding of the plot from diverse sources such as visual input (imagery), which acts as a support to aural input (sound), enhancing L2 reading and listening comprehension skills (Garza, 1991; Borrás & Lafayette, 1994; Danan, 2004). Moreover, the choice of visual input is crucial to obtain positive outcomes as L2 learners might not engage with a TV programme or film if it does not meet their expectations (Rubin, 1994).

Webb and Rodgers (2009b) and Rodgers and Webb (2011) also report the effects of multimodal input on language learning, arguing that L2 comprehension increases if multiple episodes of the same TV series, like one full season, are watched successively; in this case, learners will not need to learn about new characters, themes or settings for each episode, due to the background knowledge already acquired in the first episodes of the season, and this will facilitate comprehension. They argue that

in a regular season of a TV series, the number of word families (WFs) will be lower, due to the recurring cast, repeated scenarios and connected plots that will appear in most of the episodes, rather than watching random episodes from different TV programmes. Despite the potential that TV series have in offering comprehensible authentic input for L2 learners, research has not traditionally focused on analysing L2 comprehension of full-length episodes (Rodgers, 2013).

One study in L2 comprehension research with audiovisual input that is important to mention is the one conducted by Rodgers (2013), which contains a longitudinal design. Indeed, the present study was inspired by it. In the first study of his PhD thesis, Rodgers analysed L2 comprehension through viewing ten episodes of the American TV series *Chuck*. The participants, 282 male and 133 female Japanese ESL university students ranging from a pre-intermediate to intermediate level, took comprehension tests for each episode, which included MC, T / F and item sequencing questions. The results showed significant comprehension gains from the 1st until the 10th episode, which implied that the background knowledge that the students had accumulated since the first episode was useful for them to comprehend the other episodes better. However, scores were episode-dependent.

Several studies have analysed the effects of subtitled videos for L2 comprehension at different proficiency levels. Table 3.1 presents a chronological summary of these studies, including the number of participants, their proficiency level, type of audiovisual input participants were exposed to, subtitling condition used and a brief summary of the results. We now turn to describe the tendencies found according to the different age groups (adults, adolescents and children).

The research presented in Table 3.1 was conducted to study the L2 comprehension of audiovisual materials (with or without subtitles) and shows that generally speaking, both L1S and L2S aided comprehension when included in the video (Guichon & McLornan, 2008; Başaran & Köse, 2012; Vulchanova et al., 2015; Galimberti & Miralpeix, 2018), although significant differences were not always found. When L2S are added to the input, comprehension is significantly enhanced (Garza, 1991; Huang & Eskey, 1999; Winke et al., 2010; Etemadi, 2012, Rodgers, 2013; Gowhary et al. 2015; Ebrahimi & Bazaee, 2016). The studies that compared L1S and L2S obtained mixed results: the outcomes of some studies showed significant differences in favour of L1S (Markham et al., 2001; Markham & Peter, 2003; Bianchi & Ciabattoni, 2008; Latifi et al, 2011), whereas others showed significantly higher scores for L2S (Hayati & Mohmedi, 2011; Etemadi, 2012; Birulés-Muntané & Soto-Faraco, 2016). Studies comparing L2S, keyword subtitles and no subtitles did not find significant differences, even though higher scores were obtained when the subtitles were included in the input (Guillory, 1998; Montero-Pérez et al., 2014). In addition, Taylor (2005) did not obtain significant differences for L2 comprehension and the outcomes showed negative results for L2S.

As can be observed in Table 3.1, most of the studies conducted on L2 comprehension with subtitled audiovisual materials have been carried out with adults: Garza (1991); Guillory (1998); Huang and Eskey (1999); Markham et al. (2001); Markham and Peter (2003); Taylor (2005); Bianchi and Ciabattoni (2008); Guichon and McLornan (2008); Bravo (2008); Winke et al. (2010); Hayati and Mohmedi (2011); Latifi et al. (2011); Etemadi (2012); Rodgers (2013); Montero-Pérez et al. (2014); Frumuselu et al (2015); Gowhary et al. (2015); Birulés-Muntané & Soto-Faraco (2016); Ebrahimi and Bazaee

(2016). Furthermore, very few studies have been conducted with teenagers (Baltova, 1999; Bravo; 2008; Lavour & Bairstow, 2011; Vulchanova et al., 2015; Pujadas, 2019) or children (Başaran & Köse, 2012; Galimberti & Miralpeix, 2018). In addition, there is a study conducted by Gesa (2019) who compared three different age and proficiency groups (children, teenagers and adults) on content comprehension.

Table 3.1. Previous research on L1/L2 subtitling and comprehension (L1S: L1 subtitles, L2S: L2 subtitles)

Study	Participants	Proficiency Level	Video type	Comprehension test	Subtitles' Condition	Results
Garza (1991)	Adults (19-22 years old)	Advanced (n=140)	5 segments of American or Russian videos on different genres: drama, news, comedy, music and animation. (2-4 minutes each)	Ten MC questions for each segment	L2S / No subtitles	L2S scores significantly higher than the no subtitles group. L2S > no subtitles
Guillory (1998)	Adults College Students	Beginners (n=202)	Videotape from the textbook: <i>Parallèles</i> (2 segments)	Seven open-ended questions	L2S / No subtitles / keyword subtitles	No significant differences, although subtitled groups obtained higher scores.
Baltova (1999)	Teenagers (Grade 11)	Beginners (n=93)	Scientific Documentary (7.5 minutes)	8 open-ended questions (taken twice)	L1S / L2S / No subtitles	L1S and L2S groups significantly outperformed the no subtitles group.
Huang and Eskey (1999)	Adults	Intermediate (n=30)	Instructional video series <i>Family Album</i> (21 minutes)	MC test (listening).	L2S / No subtitles	L2S scores significantly higher than the no subtitles group.
Markham et al. (2001)	Adults College students	Intermediate (n=169)	Short episode on the Apollo 13 NASA mission to space (7 minutes)	MC comprehension test and a written summary.	L1S / L2S / No subtitles	L1S significantly outscored L2S. No subtitles group scored significantly worse than L1S and L2S.
Markham and Peter (2003)	Adults College students	Intermediate (n=213)		20 MC questions in the L2.	L1S / L2S / No subtitles	
Taylor (2005)	Adults College students	Beginners (n=85)	Instructional video on typical Spanish and Latin food <i>Ricos Sabores</i> (10 minutes)	15 MC questions in the L1 and free recall.	L2S / No subtitles	No significant differences. On free recall, learners with 3-year exposure significantly higher than those with 1-year exposure.

Chapter 3 – Multimodal input and L2 comprehension

Study	Participants	Proficiency Level	Video type	Comprehension test	Subtitles' Condition	Results
Bianchi and Ciabattoni (2008)	Adults (18 – 45 years old)	Beginners (n=17) Intermediate (n=45) Advanced (n=23)	Short clips of <i>Harry Potter and the Philosopher's Stone</i> and <i>Fantasia</i> .	A MC test on L2 content comprehension and a post-test.	L1S / L2S / No subtitles	Significant differences in <i>Fantasia</i> : L1S better than L2S and 'no subtitles' the lowest. No significant differences for <i>Harry Potter</i> .
Guichon and McLornan (2008)	Adults (20 years old)	Intermediate (n=40)	News report on a Franco-American couple living in the suburbs of Washington, DC. (3 minutes)	Written summary in the L2.	L1S / L2S / No subtitles / Sound only	No significant differences between L1S and L2S. Results higher when subtitles are present.
Bravo (2008)	1 st study Adults (20 – 66 years old; average age 23)	Elementary (n=12) Intermediate (n=11) Advanced (n=9)	6 excerpts of different programme genres: news, feature film, musical, documentary, soap opera, cartoons. (4 - 6 minutes each)	Content comprehension questions and two vocabulary recall tests.	L2S / No subtitles	L2S scored significantly higher, usually doubling the no subtitles condition for all genres except for cartoons.
	2 nd study Teenagers (13 and 14 years old)	Intermediate (n= 75)	10 episodes of <i>The Fresh Prince of Bel-Air</i> (15 minutes each)	MC / item sequencing comprehension and vocabulary questions in the L2 and translation of lexical phrases.	L1S / L2S	L1S scored significantly higher than L2S. Scores always higher for L1S since week 4.
Winke et al. (2010)	Adults (20 – 22 years old)	Beginners (n=150)	3 short animal documentaries (3 – 5 minutes)	MC comprehension test.	L2S / No subtitles	L2S scored significantly higher than no subtitles.
Hayati and Mohmedi (2011)	Adults (Average age: 22)	Intermediate (n=90)	<i>Wild Weather</i> (6 episodes of 5 minutes each)	Ten MC questions.	L1S / L2S / No subtitles	L2S significantly better than L1S and no subtitles.
Latifi et al. (2011)	Adults (17-30 years old)	Intermediate (n=36)	30' of film <i>Alvin and the Chipmunks</i> divided in 15 sessions. (2 minutes each)	Ten MC comprehension questions.	L1S / L2S / No subtitles	L1S scored the highest, followed by L2S. The no subtitles group scored the lowest.

Study	Participants	Proficiency Level	Video type	Comprehension test	Subtitles' Condition	Results
Lavaur and Bairstow (2011)	Teenagers (15 – 18 years old)	Beginners (n=30) Intermediate (n=30) Advanced (n=30)	Short clip of Hitchcock film <i>North by Northwest</i> (≈ 8 minutes).	Comprehension test with half of the questions referring to images and half to dialogue	L1S / L2S / No subtitles	Significant differences with beginner learners in dialogue: L2S>no subtitles, L1S>no subtitles.
Etemadi (2012)	Adults (20 - 27 years old)	Advanced (n=44)	2 documentaries: <i>Dangerous Knowledge</i> and <i>Where is my robot?</i> (20 and 30 minutes)	MC comprehension test.	L2S / No subtitles	L2S scores significantly higher than the no subtitles group (L2S>'no subtitles').
Başaran and Köse (2012)	Children (Grade 8, average age: 14 years old)	Intermediate (n=10) Low-Intermediate (n=20)	<i>Harry Potter and the Order of the Phoenix</i> (19 minutes only)	MC comprehension test.	L1S / L2S / No subtitles	No significant differences among the three conditions.
Rodgers (2013)	Adults College students	Pre-intermediate to intermediate (n=73)	TV series <i>Chuck</i> 10 episodes (≈ 40 minutes)	Item sequencing, MC and T / F tasks.	L2S / No Subtitles	L2S scores significantly higher than the 'no subtitles' group. L2S > 'no subtitles'
Montero Pérez et al. (2014)	Adults Mage= 17.98 years	High-Intermediate (n=133)	3 short videos from a current affairs programme (≈ 3, 4 and 3 minutes)	T / F, open ended questions and a task which combined general and more specific questions	No subtitles / L2S / Keyword subtitles / highlighted keywords	No significant differences obtained.
Frumuselu et al. (2015)	Adults	Beginner to advanced (n=40)	13 episodes of the TV series <i>Friends</i> (≈ 25 minutes each)	30-item pre- and post-test: 15 MC questions 15 open questions	L1S / L2S	L2S group scored significantly higher than L1S.
Vulchanova et al. (2015)	Teenagers (16 – 17 years old)	Intermediate (n=16) Advanced (n=17)	Cartoon: <i>Family Guy</i> 1 episode (≈ 20 minutes)	MC L2 comprehension test.	L1S / L2S / No subtitles	L1S and L2S significantly different from the 'no subtitles' group.

Chapter 3 – Multimodal input and L2 comprehension

Study	Participants	Proficiency Level	Video type	Comprehension test	Subtitles' Condition	Results
Gowhary et al., (2015)	Adults at a Language Institute	Beginners (n=22) Intermediate (n=22) Advanced (n=22)	Instructional video series: <i>Connect with English</i> (10 episodes of 20 minutes each)	30 MC questions from the TOEFL test as a pre and post-test.	L2S / No subtitles	L2S scores significantly higher than the no subtitles group in all proficiency levels.
Birulés-Muntané & Soto-Faraco (2016)	Adults (21 – 28 years old)	Intermediate (n=60)	1 hour episode of <i>Downtown Abbey</i> .	Eight comprehension questions about the story and a vocabulary test of 15 definition matching items.	L1S / L2S / No subtitles	Significant differences on listening comprehension: L2S > 'no subtitles'. Significant differences on plot comp.: L1S> 'no subtitles'.
Ebrahimi and Bazaee (2016)	Adults (20 – 27 years old)	Advanced (n=44)	Documentary: <i>Dangerous Knowledge</i> (30 minutes)	10 MC content comprehension and 10 MC vocabulary questions.	L2S / No subtitles	L2S scores significantly higher than the no subtitles group. L2S> 'no subtitles'.
Galimberti and Miralpeix (2018)	Children (12 years old)	Beginners (n=52)	1 episode of <i>The Suite Life of Zack and Cody</i> (22 minutes)	T /F questions, sequencing items and MC questions.	L1S / L2S / No subtitles	No significant differences.
Gesa (2019)	Children (Grade 6) Teenagers (Grade10) Adults (University)	Beginner Intermediate Advanced (n=158)	Several episodes of the following TV series: Grade 6: <i>The Suite Life of Zack and Cody</i> and <i>The Wizards of Waverly Place</i> . Grade 10 and university: <i>I love Lucy</i> and <i>Seinfeld</i> .	Comprehension test including Item sequencing, T / F and MC questions.	L1S (Grade 6) L2S (Grade 10 and university)	-Time was found to have a significant effect for university students and for children in terms 2 and 3. -Pairwise comparisons between episodes showed significant differences but did not follow a regular pattern.

Chapter 3 – Multimodal input and L2 comprehension

Study	Participants	Proficiency Level	Video type	Comprehension test	Subtitles' Condition	Results
Pujadas (2019)	Teenagers (Grade 8)	Beginners (n= 98)	24 consecutive episodes of <i>Fresh off the boat</i> . (≈ 21 minutes each episode)	5 T / F items and 5 MC items.	L1S / L2S	L1S scored significantly higher than L2S in all episodes for content comprehension.

3.2.1. Studies with adult learners

Garza (1991) evaluated the effectiveness of subtitled audiovisual materials on L2 comprehension (Russian and English) with 140 college students (19-22 years old). The participants, with an advanced proficiency level, were divided into two groups and watched five authentic video excerpts in English or Russian. This input belonged to different genres (drama, comedy, news, animation and music), and participants viewed them with or without L2S. The results of the MC comprehension tests showed significant differences favouring the groups watching the videos with L2S, indicating that L2S “bridge the gap between the learners’ competence in reading and listening” (p. 239).

Another study that analysed the effects of subtitles on L2 comprehension is the one conducted by Guillory (1998), although he focused on adult beginner learners attending university. The 202 participants were divided into three groups (L2S, keyword subtitles and no subtitles) and watched two clips from a textbook. After watching the excerpts, the students answered fourteen content comprehension questions (seven for each clip). Results indicated that the groups watching the input with L2S and keyword subtitles outscored the group viewing it without subtitles. In addition, the keyword subtitles group obtained higher means than the full subtitles group, although no significant differences were found between them. These results suggest that either L2S or keyword subtitles helped improving L2 comprehension.

Similarly, Huang and Eskey (1999) did not use authentic materials but an instructional video series specially designed for ESL teaching purposes. The 30 adult intermediate

participants were divided into two groups and watched one episode (seven minutes) of *Family Album* twice, with L2S or without subtitles. Immediately after watching it, they all took a MC comprehension test and the results indicated significantly higher scores for the L2S group.

Another study which compared the effects of L2S vs. No subtitles was conducted by Winke et al. (2010). The participants were 150 L1 English (except one with L1 Kannada) adult beginner learners of Arabic, Chinese, Spanish, and Russian as a L2 attending second and fourth year at an American university. They watched (with L2S or without subtitles) three short videos selected from three animal documentaries (3-5 minutes each) twice. Participants answered some MC L2 comprehension and vocabulary questions immediately after viewing the videos. The analysis of results, which focused on the second-year L2-Spanish students, indicated that the L2S group obtained significantly higher scores than the no subtitles group on L2 comprehension and vocabulary. These outcomes are consistent with the results of previous studies (Garza, 1991; Guillory, 1998; Huang & Eskey, 1999).

Also in line with these results, Etemadi (2012) obtained significant outcomes for L2S for the 44 adult subjects with an advanced English level who attended a university in Iran and participated in the study. Input consisted of two documentaries, *Where is my robot?* and *Dangerous Knowledge*, and each group watched them either with L2S or without subtitles. Immediately after watching the videos, the participants answered a MC comprehension and vocabulary recognition test. The scores showed significant differences in comprehension, favouring L2S when presented together with the videos. The students achieved higher results in L2S when compared to the no subtitles

condition, concluding that subtitles aided comprehension. However, there were no significant differences for vocabulary recognition, a finding that will be reported in the vocabulary section below.

Regarding the study we have previously introduced as a source of inspiration for the present thesis, Rodgers' (2013) PhD thesis comprised five studies and analysed the L2 comprehension and vocabulary acquisition of Japanese university EFL learners with a pre-intermediate and intermediate level. In the fifth study, 73 students watched ten episodes of the American TV series *Chuck* with L2S or without subtitles on a weekly basis. The episodes were divided into six sections and the students, who were tested immediately after viewing each one of them, were able to read the questions before they started watching each video. The comprehension tests included MC, T / F and item sequencing questions. On the other hand, the vocabulary tests comprised MC tests in which the participants had to mark the translation of the TWs presented. Results indicated increasing scores and significant results for both groups in comprehension between the first episode and the last one. When analysing each episode individually, scores were consistently higher for the L2S group. There were no significant differences between both groups for vocabulary acquisition, although some significant correlations between vocabulary knowledge and lexical gains were obtained for the L2S group but not for the no subtitles group.

Another study examining the effectiveness of L2S was conducted by Gowhary et al. (2015), although the materials used in their research were not authentic. The 66 adult beginner, intermediate and advanced EFL learners attending a private institute in Iran watched 10 episodes (20 minutes each) of *Connect with English*, a TV series specially

designed for EFL learning with L2S or without subtitles. Results indicated that the L2S group outperformed the no subtitles group. Furthermore, the scores of the MC post-test confirmed the existence of significant differences for the L2S group, concluding that subtitles aided L2 comprehension for all proficiency levels, although the advanced learner group obtained higher scores than the other groups.

Finally, a more recent study analysing the effects of L2S for L2 comprehension was performed by Ebrahimi and Bazae (2016), who examined L2 comprehension with 44 Iranian adult EFL advanced learners who attended a language institute in Iran. In one session, the participants watched the documentary *Dangerous Knowledge*, the same authentic video used by Etemadi (2012) in her study, with L2S (English) or without subtitles. Immediately after viewing the input, they performed a MC comprehension test and significantly higher scores for L2S were obtained when compared with the no subtitles group.

Bravo performed two different studies in her PhD thesis. In the first study, she analysed the L2 comprehension and vocabulary learning of 32 adult L2 Portuguese learners attending a university in Portugal. They belonged to different nationalities and their proficiency levels ranged from elementary to advanced. Before starting the treatment, they answered a questionnaire on their TV viewing habits and another one when they had finished watching the videos and taking the tests. They viewed twelve 4-6 minute excerpts (six with L2S and six without subtitles) of authentic TV programmes or films which belonged to different genres (news, feature film, musical, documentary, soap opera or animated cartoons) either with or without subtitles. After each video, participants answered content comprehension questions. The L2S group scored

significantly higher, usually doubling the score of the no subtitles condition for all genres, except for animated cartoons. This last genre obtained the lowest scores, and the no subtitles condition obtained a slightly higher score (39%) over L2S (36%). The author suggests that when watching the animated video, participants relied more on the images due to their attractiveness, while the slow rhythm of the video allowed for a better understanding than the other TV genres analysed. The results of the two vocabulary recall and retention post-tests indicated that the advanced students scored the highest (72% and 74%), followed by the intermediate and the elementary learners (50% and 53%), who scored the lowest results (39% and 42%).

Taylor (2005) conducted a study that obtained conflicting results: the effects of subtitled videos on the L2 comprehension of 85 beginner learners of Spanish attending university were analysed and the strategy used was also assessed. The four groups of participants watched *Ricos sabores*, an instructional video on typical Spanish and Latin food, with L2S or without. The students took a MC test in English and a free recall test where they had to write all the information they remembered about the programme. The results showed no significant differences on the *t*-tests between the two groups. However, the participants with three years of Spanish obtained significantly higher scores on free recall when compared to the students with only one year. Significant differences were also obtained when comparing the two groups with one year of exposure to Spanish. In this case, the results favoured the students watching the video without subtitles, suggesting that L2S might be harmful. Furthermore, the students with less exposure to Spanish found L2S more distracting than the students who had had more instruction in the target language, although all the students considered L2S a helpful tool. Regarding the students with three years of

exposure to the Spanish language, L2S did not aid or harm their language learning. After analysing the results of this study, some authors have expressed their concerns about these results: Vanderplank (2010) believes the study is not reliable because the materials selected were not appropriate for the low-proficiency learners of the study. Furthermore, he claims that the use of L2S with beginner students could not have been adequate because they might not be ready to process them. In addition, Rodgers (2013) holds that the detrimental results were due to a wrong procedure. He argues that the use of free recall to measure L2 comprehension might not have been an adequate form of assessment if the students were not motivated enough.

The study by Montero Pérez et al. (2014), a subsequent study from Montero Pérez et al. (2013), is also worthy of mention. It analysed the L2 comprehension and vocabulary acquisition of 133 adult high-intermediate French students who watched 3 short videos (2.5–4.5 minutes) from a Belgian and Swiss current affairs programme twice. The first clip selected was about French brewery strategies for production and export and the second and third clips about the history of the Lego factory. The participants, divided into four groups (L2S, L2S with highlighted words, keyword subtitles and no subtitles), were tested for L2 comprehension and vocabulary learning and they also performed a questionnaire at the end of the treatment. After viewing each clip, participants completed three comprehension tests: short open-ended questions, T / F items and a combination task where the students had to combine different statements with their corresponding picture. Statistical tests showed that there were no significant differences between groups (they all obtained similar scores), and the findings implied that vocabulary size (VS) was related to comprehension. The authors claim that this development could be due to a lack of appropriateness of the comprehension

questions or a mismatch between the difficulty of the videos and the level of the learners. It is important to note that the scores were not very high, implying that the content was not obvious to the participants. Further research with longitudinal studies including longer videos in different subtitling conditions is recommended.

The studies reported above compare the effects of subtitled (L2S) and non-subtitled videos and there is a tendency for participants watching videos with L2S to score higher than their peers (Garza, 1991; Guillory, 1998; Huang & Eskey, 1999; Winke et al, 2010; Etemadi, 2012; Rodgers, 2013; Gowhary et al., 2015; Ebrahimi & Bazae, 2016) with the exceptions of Bravo (2008) and Taylor (2005). However, we also find research comparing the effects on comprehension of watching L1S, L2S and non-subtitled videos. The study performed by Bianchi and Ciabattini (2008) showed favourable results on L2 comprehension with multimodal input but especially with L1S. They divided the 85 participants of their study (adult Italian ESL learners attending an English course at university) according to their proficiency level (beginner, intermediate or advanced). Comprehension and the interaction of variables such as video type, topic familiarity, language complexity, language level of viewer, and familiarity with subtitling were analysed. Participants watched various clips of *Fantasia*, a Disney animated film, and *Harry Potter and the Philosopher's Stone* with L1S, L2S or no subtitles and they took a MC test on L2 content comprehension after watching the videos. The students also took a post-test one week after watching the videos. Results on content comprehension showed that the L1S group scored the highest at the three different proficiency levels for the two films. Significant differences were found for *Fantasia*, favouring L1S, although there were no significant differences for the Harry Potter movie. Even though the results were consistent for L1S, some

differences were found between the L2S and the no subtitles group at the intermediate level: the L2S group outscored the no subtitles group in the Harry Potter film but, on the contrary, the no subtitles group outscored the L2S group for *Fantasia*. The authors claim that the different videos might be the cause for the conflicting results and further research is recommended.

On the other hand, a study by Guichon and McLornan (2008) did not give significant differences for L2 comprehension when comparing groups watching subtitled videos with L1S vs. those with L2S. They assessed the L2 comprehension of 40 intermediate adult ESL learners attending a French university. The participants were divided into 4 groups and watched a news report on a Franco-American couple living in the suburbs of Washington, D.C. twice, under one of these four conditions: L1S, L2S, no subtitles or sound-only. At the end of the video, they were asked to write a summary of its content using the notes they had taken. When subtitles were included in the input, the results were higher but no significant differences were found between L1S and L2S. Authors argue that the small number of participants used for this study did not allow for the results to be generalised.

The research conducted by Markham et al. (2001) tested the effects of watching a short authentic video on the Apollo 13 NASA mission to space on L2 comprehension with 169 intermediate participants at university level. They watched the video with L1S (English), L2S (Spanish) and no subtitles and immediately after watching it the subjects wrote a summary. The same material and the same conditions were used in Markham and Peter (2003) to check for L2 listening and reading comprehension. The participants, 213 university students with an intermediate proficiency level, took a MC

test. The results of both studies showed significant differences favouring L1S when compared with L2S. Significant differences between the subtitles (L1S and L2S) and the no subtitles groups were also obtained, with the latter scoring significantly worse. The authors conclude that in order to enhance comprehension, ESL students should start watching videos with L1S, move to L2S when their proficiency level improves and end up watching videos with no subtitles when the language is mastered.

The effects of L1S, L2S and no subtitles on L2 comprehension were also compared in research conducted by Hayati and Mohmedi (2011). 90 participants, intermediate EFL students attending a university in Iran, watched six episodes of a BBC documentary film called *Wild Weather*, dealing with the Earth's climate and natural disasters. After each episode, the subjects took a MC comprehension test and the results showed significantly higher scores for L2S, followed by L1S and finally the no subtitles group. The authors concluded that watching videos with L2S aided the students significantly better than viewing them with L1S or no subtitles. L2S helped the students understand the content of the video and develop their comprehension in the L2 much better than the other conditions.

Subsequently, a study performed by Latifi et al. (2011) analysed the effects of L1S, L2S and video only on 36 adult intermediate ESL learners with Persian as their L1. The material selected was 30 minutes of the film *Alvin and the Chipmunks*, which was divided into 15 segments, each of 2 minutes' duration, and watched in 15 consecutive sessions over a four-week period. After each session, the participants took a MC comprehension test with ten questions dealing with the main ideas of the clip they had just watched. The results confirmed that the groups watching the clips with L1S and

L2S outperformed the one watching the materials without subtitles. The L1S group obtained higher significant differences than the no subtitles group, however, the differences in scores between L1S - L2S and L2S with the no subtitles group were not significant. The authors point out that the small number of participants could be the reason why they obtained such results and call for further research in the area with more proficiency levels in order to generalise the results.

Finally, there is a study dealing with subtitles and authentic input performed by Birulés-Muntané & Soto-Faraco (2016). In this experiment, the 60 participants, intermediate EFL learners from Spain, watched one episode of the British TV series *Downton Abbey* with L1S, L2S or no subtitles and took a comprehension test immediately after it. After analysing the data, it was concluded that the participants watching the episode with L2S obtained significantly better scores (17%) than the other groups in L2 listening comprehension. The no subtitles group also obtained significantly improved scores (7%), although much lower than L2S, but no significant differences were found for L1S (0%). The high scores for L2S revealed them as the best option for improving L2 listening comprehension. The results for plot comprehension were different: L1S obtained higher significant scores, followed by L2S and the no subtitles group, which scored the lowest. According to the authors, this outcome proves that the participants achieved higher results when they could read the subtitles in their L1 and benefit from their high competence in their native language.

Taken as a whole, the research conducted with adult learners on L2 comprehension suggests that the inclusion of written input in the videos selected enhanced learners' L2 comprehension: studies comparing the effectiveness of L2S vs. non-subtitled

videos show that L2S significantly improved the results. However, no definite conclusions have been drawn on the best subtitle type for L2 learning (L1S or L2S), due to the many differences between the studies conducted (in terms of participants' proficiency levels, ages, type of video used or ability tested).

3.2.2. Studies with adolescent learners

Research on the effects of L2 comprehension with subtitled audiovisual materials conducted with teenagers has not attracted much attention. Table 3.1 shows that there is a limited amount of studies with L2 learners at this age (Baltova, 1999; Bravo, 2008; Lavaur & Bairstow, 2011; Vulchanova et al., 2015; Pujadas, 2019) compared to adult learners, although more studies have been conducted on adolescents recently.

The study conducted by Baltova (1999) with low-level learners in her PhD thesis is one of the first with adolescents. The participants were L2 French beginner learners with English as their L1, who attended a school (Grade 11) in Ontario, Canada. They were divided into three groups and watched a scientific documentary with English audio and French subtitles (L2S), French audio and L2S or no subtitles with the audio in French. They took a short answer comprehension test immediately after watching the video and they repeated it again two weeks later. The students from the groups watching the video with L2S obtained significantly higher scores for comprehension compared with the group watching it with aural input only.

Nearly a decade later, Bravo (2008) performed another study with teenagers. This research is detailed in the second study in her PhD thesis, conducted with 75 ESL learners aged 13 and 14 years old who attended Grade 9 at a school in Portugal. At

the beginning of the treatment, the students answered a questionnaire in which they were asked about their TV viewing habits. In addition, before watching each episode the L2 learners received a handout covering 10 comprehension and 10 vocabulary items. They watched ten episodes from season one of the American TV series *The Fresh Prince of Bel-Air* with L1S or L2S. The videos were divided into two sections and the participants answered comprehension and vocabulary questions through MC and item sequencing tests immediately after viewing each one of them; they were also asked to provide translations of lexical phrases three times. There were consolidation tests after weeks five and ten and a final one which included items from the ten weeks. Results indicated that L1S obtained higher scores than L2S for comprehension and vocabulary learning, “although the differences were not very significant” (p. 196). It should be noted, though, that from week four onwards, L1S scores were always higher than L2S. The tests performed to analyse vocabulary recall and retention indicated that the L1S group performed better than the learners watching the videos with L2S, who “had more constant scores” (p. 195). For the vocabulary recall and retention consolidation tests, the results indicated a significant effect for the L1S group between the first and the final test. The author claims that L1S are more efficient than L2S for L2 comprehension and lexical learning.

Lavour and Bairstow (2011) also examined adolescents’ performance. The participants of their study were 90 French high school ESL students who were divided into three groups depending on their proficiency level: beginner, intermediate and advanced. The students had no previous experience viewing subtitled films and they watched a short clip of the film *North by Northwest* by Hitchcock in one of the following conditions: L1S, L2S or no subtitles. The findings concluded that the L2 beginner

learners processed the visual content better when subtitles were not present, although dialogue comprehension scores were significantly higher with L1S and L2S than without subtitles. Intermediate students still relied on the subtitles, although they did not score higher for L2 comprehension in any of the subtitled conditions. In addition, the advanced group also achieved high comprehension results when they watched the video without L1S or L2S. Overall, subtitles were considered distracting, especially if participants (advanced) did not need them to understand the message conveyed because their knowledge of the language was good enough to process the information without them. It is interesting to note that subtitles did not have a negative effect for the viewers who were already used to watching subtitled films or TV programmes, but the lack of familiarity of some students with subtitled films might have affected their results.

Another study with teenagers is the one performed by Vulchanova et al. (2015). They conducted a study with 114 Norwegian teenagers (16-17 years old), at an intermediate and advanced level, learning ESL at high school. The learners watched one episode of the cartoon series *Family Guy* with L1S (Norwegian), L2S (English) or no subtitles and completed a MC L2 comprehension test immediately after watching it. The results obtained showed significant differences for the L1S and L2S groups compared to the no subtitles group, although there were no significant differences regarding the language used in the subtitles. The findings obtained proved that L2 comprehension was enhanced when viewing a single TV episode either with L1S or L2S in the input.

Finally, the longitudinal study conducted by Pujadas (2019) examined the comprehension of low L2 proficiency learners at Grade 8 in a high school in Barcelona

(Spain). Participants were divided into two groups and watched 24 consecutive episodes of the TV series *Fresh off the boat* either with L1S or L2S. After watching each episode, the learners took a comprehension test, which consisted of five T /F items and five MC items. Results showed significant differences for comprehension in all episodes, favouring the L1S group.

The few studies available with adolescents presented in this section show diverse results. The research conducted by Baltova (1999), Lavour and Bairstow (2011) and Vulchanova (2015) show higher significant scores for the videos watched with L1S and L2S. However, Lavour and Bairstow only found significant results with the beginner learners in their study. Finally, Bravo (2008) and Pujadas (2019) obtained favourable results for L1S when compared with L2S. This seems to point towards a facilitating effect of subtitles, especially when the level is low. However, this finding is not consistent in the few studies with adolescents available.

3.2.3. Studies with young learners

The amount of studies conducted analysing the effects of L2 comprehension on subtitled audiovisual materials with young learners is very scarce (Matielo et al., 2015). The studies reported below present few treatment sessions using different audiovisual materials (cartoons, movies and TV series) and different procedures. The results obtained were similar in Başaran and Köse (2012) and Galimberti and Miralpeix (2018), with no significant differences between groups using subtitles (either in the L1 or L2) and those not, although the small samples used and the limited amount of time

devoted to the treatment make it impossible to generalise the results and it is obvious that more research is needed.

In Başaran and Köse (2012), the participants were 30 Turkish students learning EFL at Grade 8 at a primary school in Turkey. They belonged to two different proficiency levels: low-intermediate (20 students) and intermediate (10 students) and they watched the first 19 minutes of the film *Harry Potter and the Order of the Phoenix* with L1S (Turkish), L2S (English) or no subtitles. They were divided into three groups (two low-intermediate and one intermediate) and they took a MC comprehension test immediately after watching the video. Although the results did not show any significant differences between the three conditions, the low-intermediate young learners who watched the videos with L1S or L2S obtained higher scores. Given these inconclusive findings, the author asks for further research at different proficiency levels.

Galimberti and Miralpeix (2018) worked with young low-proficiency EFL learners attending Grade 6 at an Italian public school and assessed L2 comprehension and vocabulary acquisition after watching an episode of a TV series. The 52 participants watched *The Suite Life of Zack and Cody* from the Disney Channel with L1S, L2S or no subtitles. The students were tested immediately after watching the episode. The results of the T / F, sequencing and MC tests indicated that the L1S group obtained better results for comprehension, although the scores were only slightly higher than the other groups and no significant differences were found. Meaning recognition followed the same pattern and no significant differences were found either. For vocabulary recall, the L2S group obtained significantly higher scores compared with the L1S group, although the results were similar to the no subtitles group. Further

research is recommended with beginner L2 learners using a longitudinal design over several sessions.

On the other hand, Gesa (2019) conducted his study with participants of different ages and proficiency levels. They attended primary (Grade 6) and high school (Grade 10) as well as university. At each grade learners were divided into two groups (experimental and control) and those in the experimental group watched a subtitled episode of a TV series each week throughout a whole school year while the CGs did not. The TV series selected were the following: *The Suite life of Zack and Cody* and *The Wizards of Waverly Place* for the primary group and *I love Lucy* and *Seinfeld* for those at high school and university level. The participants in experimental groups took a comprehension test (adapted from Rodgers, 2013) for each episode, which included a sequencing exercise, MC questions and T / F items. The results for L2 comprehension showed no clear pattern in the evolution of the scores throughout the year, although time was statistically significant for the university level students and in terms 2 and 3 for Grade 6 children. As also happened in Rodgers (2013) with adult learners, comprehension scores were episode-dependent.

Finally, a very recent study with low-proficiency L2 young learners was conducted by Avello (2023), where primary school children in Chile watched several episodes of the TV series *Charlie and Lola* with L2S. It seems that the highly supportive imagery of the TV series lowered the learners' lexical demands and compensated for the low L2 proficiency level of the children.

The studies reported above may suggest some positive effect of subtitling for L2 comprehension in children watching audiovisual materials. However, most studies do not find significant differences between conditions and no clear patterns are found. Although research on subtitling has become popular in the past few years, generalising findings from the studies conducted is difficult as there are still very few studies with young low-proficiency children, also because a wide variety of testing procedures and input materials have been used (Montero-Pérez, 2013). To the researcher's knowledge, there is only one existing study on extensive viewing conducted with children adopting a longitudinal approach (Avello, 2023). While most research has been conducted with intermediate and advanced university learners, there is a need to carry out research on extensive viewing with young learners (where the children are not just exposed to one video and the treatment lasts longer than a couple of sessions or weeks) in order to explore the effects of sustained exposure to multimodal input on children's L2 learning (Matielo et al., 2015).

CHAPTER 4 – MULTIMODAL INPUT AND L2 VOCABULARY

LEARNING

In this chapter, we will first focus on several issues related to vocabulary learning that are relevant to our research (4.1): we will see what it means to learn words following Nation's framework (2007), and that learning vocabulary implies much more than just acquiring single words. In addition, we will see that vocabulary can be learned both implicitly and explicitly. Studies on vocabulary learning from reading and listening are also presented (as well as factors that may affect learning, such as cognateness and frequency) in an attempt to provide the background for those more recent studies on multimodal input and L2 vocabulary learning where simultaneous reading and listening are involved. Then, we will concentrate on studies on L1/L2 subtitling for vocabulary learning from videos (4.2), which are presented with special attention to those dealing with adults (4.2.1) and young learners (4.2.2), as the findings in this research are relevant for the present study with EFL learners at primary school.

4.1. Vocabulary learning in the L2

FL vocabulary acquisition implies learning a huge amount of words in order to be fluent in the target language, a very demanding task for any L2 learner (Bisson et al., 2014). Learning new words is already difficult in the native language and it will obviously become even more challenging in the L2 (Lévesque, 2013).

Knowing a word entails knowledge of many different aspects. Nation (1990, 2001) distinguishes three: namely, knowledge of form, meaning and use (see Figure 4.1). Regarding each of these key dimensions, both receptive and productive knowledge should be taken into account as well. For example, knowing the written form of the word implies recognising its written form when we find it in writing (receptive knowledge), but also knowledge of how to spell it when we write it down (productive knowledge). Similarly, knowing the meaning of a word may imply recognising its meaning among others in a MC test (receptive knowledge) but also knowing what it means without being given any clues when we produce it in speech (productive knowledge). In this dissertation, we will mainly explore receptive knowledge of word forms that young learners may gain after being exposed to multimodal input.

Form	spoken	R	What does the word sound like?
		P	How is the word pronounced?
	written	R	What does the word look like?
		P	How is the word written and spelled?
	word parts	R	What parts are recognisable in this word?
		P	What word parts are needed to express the meaning?
Meaning	form and meaning	R	What meaning does this word form signal?
		P	What word form can be used to express this meaning?
	concept and referents	R	What is included in the concept?
		P	What items can the concept refer to?
	associations	R	What other words does this make us think of?
		P	What other words could we use instead of this one?
Use	grammatical functions	R	In what patterns does the word occur?
		P	In what patterns must we use this word?
	collocations	R	What words or types of words occur with this one?
		P	What words or types of words must we use with this one?
	constraints on use (register, frequency ...)	R	Where, when, and how often would we expect to meet this word?
		P	Where, when, and how often can we use this word?

Note: In column 3, R = receptive knowledge, P = productive knowledge.

Figure 4.1. What it means to know a word. From Nation (2001, p.27)

4.1.1. Single words vs. multi-word units

Even though vocabulary learning research has focused on examining single-word items, it has been widely accepted that multiword units (MWUs) should be taught separately from single word units (Pellicer-Sánchez, 2020). A MWU is a vocabulary item which consists of a sequence of two or more words. They usually reduce language processing time and effort and improve learners' fluency and idiomaticity. A compound word ('ice cream') or a collocation ('do exercise') are examples of MWUs. Moreover, a study by Conklin and Schmitt (2012, p. 46) indicates that formulaic language "makes up between one third and one half of discourse" and Martinez and Schmitt (2012, p.299) claim that "research has now established that formulaic language is fundamental to the way language is used, processed, and acquired in both the L1 and L2". MWUs have been shown to be crucial in language learning (Ellis, 1996; 2002; Wray, 2002; Schmitt, 2004; Boers, 2020), as high percentages of these lexical units are included in everyday discourse (Conklin & Schmitt, 2012). It has also been argued that memory is enhanced when learners are exposed to lexical chunks (Robinson, 1988).

In order to identify MWUs, Granger and Paquot (2008, p.4) claim that the key lies in the distinction between free combinations, "which are only governed by semantic co-occurrence restrictions and are thus considered as falling outside the realm of phraseology" and MWUs "whose co-occurrence cannot be accounted for by semantics and qualify as phraseological units or phrasemes." Granger and Paquot (2008, p.15) propose a categorisation and suggest a division of MWUs into three main blocks: referential, textual and communicative. They argue that referential MWUs "convey a

content message: they refer to objects, phenomena or real-life facts” (e.g., *heavy rain, bed and breakfast, black hole*). In addition, they claim that textual MWUs are “typically used to structure and organise the content of a text or any type of discourse” (e.g., *in addition to, so that, in other words*) and communicative MWUs “are used to express feelings or beliefs towards a propositional content or to explicitly address interlocutors, either to focus their attention, include them as discourse participants or influence them” (e.g., *good morning!, you’re welcome, I think that*). Table 4.1 lists the MWUs having a referential, textual and communicative function according to the Granger and Paquot (2018) categorisation.

Table 4.1: Categorization of MWUs by Granger and Paquot (2008).

MWUs		
Referential Function	Textual Function	Communicative Function
(Lexical) Collocations Idioms Irreversible bi- and trinomials Similes Compounds Grammatical collocations Phrasal verbs	Complex prepositions Complex conjunctions Linking adverbials Textual sentence stems	Speech Act Formulae Attitudinal formulae Commonplaces Proverbs Slogans

There are several factors that might influence the learning of single words and MWUs, for example, frequency of occurrence. Webb and Nation (2017) claim that the repeated number of encounters and the quality of attention given to each encounter will affect language learning, the latter depending on the input mode, which will determine the level of engagement. The study by Webb et al. (2013) examined the role of frequency on Taiwanese university students learning ESL and results showed that the number of encounters needed to learn collocations might be similar to the number needed for

single words. Furthermore, Boers (2020) argues that it would be easier for learners to recall familiar words than new word forms. He also suggests that the type of input might influence the rate of acquisition of MWUs, as the combination of aural and written input may compensate the challenging task of understanding the MWUs in real-time listening.

There are not many studies on the acquisition of single and MWUs, as usually single words are chosen as target items to be learned. Peters (2014) analysed the frequency of occurrence required for the learning of single items and the learning of collocations. Significant differences were found for both single items and collocations appearing five times compared to those that were introduced only once. The results also suggested that collocations were more difficult to learn when compared to single items. Another study by Alali and Schmitt (2012) reported similar results. In this case, the Kuwaiti students were taught single words and idioms using the same methodology. The outcomes showed that the single words and idioms that received an extra exposure through a written task were better recalled. This data suggests that amount of exposure is crucial to learn both single and MWUs and the type of input received might be decisive for the results. The following sections present the different types of MWUs that will be examined in our study.

4.1.1.1. Collocations

Collocations can be defined as “a group of words that occur together more often than by chance [...] and they cover word pairs and phrases that are commonly used in language, but for which no general syntactic or semantic rules apply” (McKeown &

Radev, 2000, p.1), for example, ‘shown live’ and ‘out of order’. It has been argued that collocations are learnt incidentally in the same way as single word items (Mackin, 1978). Nation (2001, p.321) also claims that “language knowledge is collocational knowledge”. In their study investigating the incidental learning of collocation, Webb et al. (2013, p.107) indicated that “encountering collocations when reading while listening contributed to incidental learning of collocation”. Moreover, they state that frequency of exposure was a key variable for learning collocations in context. In a more recent study, Gonzalez Fernandez and Schmitt (2015) examined the relationship between the productive knowledge of English collocations and watching films, extensive reading (ER), listening to music and the use of social media. They obtained significant findings for the watching films and ER variables, which strongly correlated with collocational knowledge.

4.1.1.2. Compounds

Compounds can be described as combinations of a minimum of two words (Ebeling & Hasselgård, 2015). There are three types of compound words: closed form (the words are written together), hyphenated form (a hyphen is written between the words) and open forms (the words are written separately). Examples of compounds can be ‘goldfish’ and ‘dark-haired’. Research on compound noun processing has shown differences. While Badecker and Allen (2002) and Juhasz (2007) argue that these type of formulaic sequences are processed separately, Mondini et al. (2002) have claimed that compounds can be accessed as a whole unit.

4.1.1.3. Phrasal verbs

Phrasal verbs are also included in the referential block of MWUs by Granger and Paquot (2008). They can be defined as a lexical verb that when combined with a preposition or an adverb, or both, lead to a different meaning from the meaning its separate parts have. According to Liu (2011), learning phrasal verbs can be challenging for English learners and teachers. This task can also be hard for advanced learners when they try to learn them in classroom settings (Wray, 1999). There is a huge amount of phrasal verbs with multiple meanings and a complex semantic and syntactic structure, which makes them difficult to master (Liao & Fuyuka, 2004). Even though some learners might try to employ one-word verbs and avoid phrasal verbs in their discourse, phrasal verbs are important to learn as they are widely used in everyday language (Garnier & Schmitt, 2016).

There are a few studies performed in relation to phrasal verbs as MWUs (Schmitt & Redwood, 2011; Chen, 2013) in language learning. Schmitt and Redwood (2011) obtained significant results when analysing the relationship between watching English language movies and phrasal verbs knowledge, as exposure to multimodal input promoted phrasal verb learning. However, the results were not that clear when participants were listening to music. In a more recent study, Chen (2013) examined the relationship between frequency (number of occurrences of phrasal verb forms in a corpus of essays) and the knowledge of several phrasal verbs in Chinese EFL learners. A positive correlation was found between the two variables.

4.1.1.4. Linking adverbials

Linking adverbials “are instrumental in providing cohesion in both speaking and writing” (Liu, 2008, p.491) and its function is “to make semantic connections between spans of discourse of varying lengths” (Biber et al. 1999, p.558). According to Granger and Paquot (2008, p.17), linking adverbials “include various types of phrasemes such as grammaticalized prepositional phrases, adjectival phrases, adverbial phrases, finite and non-finite clauses that play a conjunctive role in the text”. Some examples of linking adverbials are: ‘in other words’, ‘last but not least’, ‘more accurately’, ‘what is more’, ‘to conclude’. In order to find out which linking adverbials should be taught in ESL classrooms, Liu (2008; 2012), who has performed corpus studies on linking adverbials, recommends introducing first those that are most frequent. Some studies have examined the role of linking adverbials in L2 acquisition for reading (Kremmel et al., 2017) and writing (Chen, 2016). However, they are not often included as TWs in multimodal input studies.

4.1.1.5. Speech act formulae

Speech act formulae or routine formulae are “those sequences that are used frequently by speakers in certain prescribed social situations” (Bardovi-Harlig, 2009, p.757). Some examples of speech act formulae are greetings such as ‘how are you’ or conventional expressions like ‘thanks for having me’. If L2 learners demonstrate a good command of routine formulae, communication will be easier and their processing load will also be reduced, which is essential especially for low-proficiency learners (Roever, 2011). However, research shows that there is a lack of pragmatic knowledge

of these formulae by L2 learners, which makes it difficult to handle social situations (Edmonson & House, 1991). Some studies have examined the exposure and recognition of routine formulae (Roever, 2005; Bardovi-Harlig et al., 2008; Bardovi-Harlig, 2009; 2010).

In Roever (2005), for example, 267 participants performed three tasks aimed at improving pragmatic competence. The outcomes showed that the participants who received 3 months' exposure to situational routines obtained significant results when compared to the English learners who did not. However, the participants' answers did not show if they did not know the expression or they found it inappropriate. Bardovi-Harlig et al. (2008) tried to improve Roever's study by including a written recognition task inspired by the vocabulary research performed by Meara (1989), in which learners had to circle the familiar expressions from a list. The written recognition task was paired with a production task and the scores showed higher results for the former. However, the outcomes also indicated that the participants, who had different proficiency levels, selected most of the expressions. The interpretation of the data revealed that it was impossible to tell if the learners selected many familiar expressions either because the task was too easy or because they circled some expressions even though they did not know them.

In a subsequent study Bardovi-Harlig (2009) improved the written recognition task performed in Bardovi-Harlig et al. (2008) and used an aural recognition task to look into the aural distinction of conversational expressions and corresponding grammatically modified expressions. The participants were adult L2 learners who performed the task, accompanied by a rating scale in which the participants had to

indicate if they heard the expression often, sometimes or never. The results were significant for the authentic conversational expressions and proved that learners were sensitive to frequency. In addition, the higher the proficiency level of the learners, the more authentic expressions they recognised, which showed that the interaction between group and proficiency level was significant, too. Bardovi-Harlig (2010) continued performing research in order to improve the study published in 2009. In this new study, the 171 participants performed a recognition and a production task. Results proved learners who obtained higher results were the ones with a higher proficiency level, whereas the lowest scores were obtained by the lower level participants. Due to the growing interest in this area, the research performed on routine formulae has increased in pragmatics studies in the past few years (Bardovi-Harlig, 2012).

4.1.2. Implicit / Explicit and Intentional / Incidental L2 vocabulary learning

The terms ‘implicit’ and ‘explicit’ learning come from the field of psychology (Rieder, 2003). Implicit learning is claimed to be the “acquisition of knowledge about the underlying structure of a complex stimulus environment by a process which takes place naturally, simply and without conscious operation”, whereas explicit learning is defined as a “more conscious operation where the individual makes and tests hypotheses in a search for structure” (Ellis 1994, p.1). Reber (1976, p.93) characterised implicit learning as “a primitive process of apprehending structure by attending to frequency cues” and he interpreted explicit learning as “a more explicit process whereby various mnemonics, heuristics, and strategies are engaged to induce a representational system”. In addition, Dekeyser (2008) argues that the key feature that distinguishes implicit from explicit learning lies in awareness: implicit learning

takes place without consciousness whereas in explicit learning the learner is aware of the learning process. He also points out that there is scarce research in the area of implicit and explicit L2 learning in order to determine which is the best method to acquire the L2.

Krashen (1982; 1985) and Reber (1976) claimed that implicit L2 learning through a considerable amount of comprehensible input is more advantageous for the acquisition of challenging language forms than explicit learning. In an effort to find out the most effective type of learning, Norris and Ortega (2000) conducted a meta-analysis in which they analysed 14 studies comparing implicit and explicit instruction and found out that the latter was more effective. They also conclude that the methods used to obtain the results in the research examined favoured explicit instruction and they point towards a lack of classroom studies focusing on implicit and explicit FL learning.

In relation to vocabulary, Beglar and Hunt (2002) claimed that explicit instruction is better for L2 beginner and low-intermediate level students with limited vocabularies (L2 children and unskilled adult learners), as they do not yet have a proficient reading ability that allows them to pick up words when reading. This is the reason why explicit vocabulary learning may be responsible for most L2 lexical learning at these low levels (Laufer, 1991, 2001; Laufer & Paribakht, 1998; Webb, 2008). On the other hand, Beglar and Hunt argue that ER and listening might receive more attention from more proficient intermediate and advanced students as implicit learning is more likely to take place. Several authors have claimed that explicit learning is also beneficial at all levels

(Schmidt & Frota, 1986; Dekeyser, 1998; Swain & Lapkin, 1995), especially for learning certain language rules (Dekeyser, 2008).

According to DeKeyser (2000; 2008), the processes children and adults follow when learning L2 vocabulary are completely different due to their distinct available SLA mechanisms: children will greatly benefit from implicit learning procedures, while adults will learn the L2 better when the language is taught explicitly. The reason why adult learners cannot use their implicit skills to the fullest extent has been explained by the Critical Period Hypothesis by Lenneberg (1967), which states that the ability children have to acquire the language successfully disappears after puberty. This theory would explain why adults find it very challenging to learn the L2 implicitly and the need they have to rely on explicit learning mechanisms to acquire it. Even if Lenneberg proposed the Critical Period Hypothesis for the L1, Johnson and Newport (1989) adapted and extended it to the L2, proposing two different theories: the exercise and the maturational hypothesis. The former argues that the strong ability to learn languages that young learners possess will disappear or deteriorate as the child grows up unless the learner uses it. If exercised, it would remain intact for SLA. On the other hand, the maturational hypothesis claims that the stronger L2 learning ability that young learners show before puberty might vanish or weaken after maturation, which may cause the need for older learners to build on explicit learning through instruction.

In L2 literature, implicit has often been taken as a synonym for 'incidental' and explicit for 'intentional' learning. However, some researchers do not agree on this synonymy and they argue that learning the L2 incidentally implies explicit and implicit mechanisms (Ellis, 1994; Huckin & Coady, 1999; Hulstijn, 2001). In vocabulary

studies, Rieder (2003), for example, claims that there is a confusion between the terms implicit and explicit learning and intentional and incidental vocabulary learning and she concludes that incidental L2 vocabulary learning involves both implicit and explicit learning processes. Ellis (1994) also points out that both implicit and explicit learning mechanisms are involved in incidental vocabulary acquisition, although they are in charge of different processes. He claims that receptive and productive aspects of word forms will be learned through implicit learning mechanisms and without awareness, whereas semantic aspects like a word's semantic properties and meanings will be learned through explicit learning, in which the learner is aware of the whole acquisition process.

Incidental L2 vocabulary acquisition has been defined by Hulstijn (2001, p.270) as the "learning of vocabulary as the by-product of any activity not explicitly geared to vocabulary learning", whereas intentional L2 vocabulary learning is "any activity geared at committing lexical information to memory". However, Hulstijn (2001) also claims that incidental learning can turn into intentional learning if students are told that they will be tested after receiving the input. Hulstijn (2001) makes a distinction between implicit learning and implicit vocabulary learning. He claims that implicit learning takes place without conscious awareness but for implicit lexical learning, the learner's attention is needed in order to learn the word forms and meaning.

Huckin and Coady (1999) list several advantages and disadvantages for incidental and intentional L2 vocabulary learning. For instance, they claim that when lexical items are learnt incidentally, reading skills and vocabulary acquisition work together to infer meaning. The opportunity to learn vocabulary in context is also an advantage,

although there are difficulties and challenges in guessing, as guesses are not always correct. Learning incidentally can be a time consuming strategy and can lead to misinterpretations that the learner will overcome when he/she has enough vocabulary knowledge and sufficient proficiency in the necessary reading strategies.

When comparing incidental and intentional learning, Ahmad (2011) summarises that incidental L2 vocabulary learning is more successful than intentional L2 vocabulary learning, due to the ER and guessing processes that learners at all proficiency levels face when inferring meaning. He also believes that, while learning words intentionally, L2 learners do not endure any cognitive process and just perform different tasks repeatedly, which result in the acquisition of a low amount of vocabulary. However, this view is not shared by many others, as it has been claimed that intentional learning through instruction significantly contributes to vocabulary development (Nation, 1990; Paribakht & Wesche, 1997).

Although in formal settings the L2 is usually learned intentionally, researchers agree that incidental vocabulary learning should be encouraged and incorporated into formal L2 learning environments (Hunt & Beglar, 2005; Nation, 2001; Schmitt, 2000; Waring & Takaki, 2003). Webb (2008) also states that incidental learning should be fostered, although the process might be slow and entail small steps, by building upon learners' previous knowledge through repeated encounters with target vocabulary until it is known.

In this thesis, we will be assessing lexical knowledge after participants have been exposed to a cartoon episode every week. We consider that the activity is not specially

geared towards vocabulary learning, but that the learning of vocabulary will be a by-product, as the focus of the task is on understanding the messages conveyed in the episode. In addition, even if participants do a comprehension exercise and a vocabulary exercise after watching the videos, these are not exams and they are not graded so students are not pressured to have a good mark (the researcher collects the exercises for research purposes). Even if authors such as Hulstijn (2001) consider that exam announcement is a way of turning incidental learning into intentional, in this case students are attentive because they enjoy watching the cartoons (not because they want a high mark). Also Ellis (1994) and Hulstijn (2001) state there is some degree of attention in incidental learning.

4.1.2.1. L2 vocabulary learning and young learners

Regarding implicit learning, children are at an advantage when they are exposed to large amounts of input, but the limited input provided to them in formal settings is not enough for them to learn in an implicit way (Muñoz, 2008b). Research has shown that children possess an outstanding capacity for acquiring the words they may encounter incidentally in speech dialogues, TV series or films and storybooks (Nagy et al., 1987). It has also been suggested that children with larger vocabularies will be able to learn the L2 faster than children with more limited vocabulary knowledge (Elley, 1989).

Children studying English in primary schools across Europe are usually taught explicitly and in a form-focused way, making them very often aware of the new vocabulary they are learning. Most schools teach L2 vocabulary explicitly, due to the low level of the students, who tend to receive EFL classes about two or three times a

week. Therefore, in FL settings, there is limited opportunity for contact with English and vocabulary is often taught explicitly, even if, according to Nation (1990) it may not be the most fashionable approach. However, it would be beneficial to explore other ways of promoting vocabulary acquisition in order to improve rate of L2 learning, given that vocabulary acquisition is a process that takes time. Dockrell et al. (2007, p.578) try to explain the intricate L2 vocabulary acquisition process that children follow: “when children acquire a new word they need to identify the sound in the speech stream, encode a corresponding phonological representation and then establish a mapping between this word form and the world; ultimately, a detailed semantic representation is developed for the new term with an indication of its morphosyntactic features”. It can be hard to do all of this incidentally, and that is why most vocabulary is taught explicitly to beginner learners.

However, the inclusion of technology in primary schools across Europe has facilitated the spread of audiovisual materials, which have lately been portrayed as tools that promote L2 vocabulary learning. TV series, films and cartoons have progressively been incorporated by L2 teachers as a part of their L2 curriculum in formal settings, leading to the conclusion that the activity of watching videos, which supply the same input through different channels, provides advantages for young learners when compared to exposure to aural input only (Lévesque, 2013).

4.1.2.2. Incidental vocabulary learning from reading and listening

Incidental vocabulary acquisition takes place when learners focus on comprehending the message conveyed and not on the word as an individual unit (Ellis, 1994). In order

for incidental L2 vocabulary learning to occur, the repeated exposure of the L2 words in meaningful collocations and contexts is needed, although there is no agreement on the number of exposures required (Paribakht & Wesche, 1997; Huckin & Coady, 1999). However, research agrees that an increased exposure can lead to better learning rate (Barclay, 2017). Even though it has been thoroughly investigated, many questions remain unanswered about incidental vocabulary learning (Gass, 1999; Paribakht & Wesche, 1997).

L2 incidental vocabulary learning through reading (unimodal input) has been shown to take place (Chun & Plass, 1996; Hulstijn et al., 1996; Zimmerman, 1997; Horst et al., 1998), although teachers need to supply learners with a varied number of opportunities to acquire each word (Beglar & Hunt, 2002), especially at low levels of proficiency, where more encounters will be needed for lexical items to be acquired (Zahar et al., 2001). If a word is only encountered once, the learner has only got a 10% chance of learning its meaning from context (Nagy et al., 1985). In addition, the process of L2 incidental vocabulary learning through reading is slow and it leads to different outcomes for each L2 learner (Parry, 1993).

Woodinsky and Nation (1988) claim that incidental L2 vocabulary learning through ER can benefit learners at all levels. Vocabulary knowledge is critical for understanding the message conveyed through a written text, as the reader needs to have enough competence to understand the words included in the input (Laufer, 1997). L2 learners will be able to understand the lexis of a text if they know the meaning of the majority of the words. This means knowing 95% to 98% of the vocabulary (Nation, 2001). In the case of children or adult beginners who do not master this high percentage,

pictorial glosses might be an option (Kost et al., 1999). Hwang and Nation (1989) have suggested that beginner learners will especially benefit from narrow reading, which contains texts with recurrent characters and related storylines, compared to independent texts that do not share any topics or plots.

Research has suggested that incidental L2 vocabulary learning through reading (using modified input such as graded readers) with adult intermediate or advanced learners can successfully take place (Horst et al, 1998; Waring & Tatsuki, 2003; Pigada & Schmitt, 2006). Brown et al. (2008) included different input modes in their study: participants were divided into three different groups and followed the treatment in one of the three conditions: reading, reading-while-listening or listening to graded readers. Results in this study suggested that learning through bimodal input (reading and listening at the same time) may provide more durable L2 word-form recognition than unimodal input. Finally, Pellicer-Sánchez and Schmitt (2010) selected authentic input (a novel) for their research and tested participants on various aspects of L2 vocabulary learning. The results from the MC word recognition spelling test showed that lexical recognition through reading was significantly taking place.

Even though most studies on incidental L2 vocabulary learning have been conducted on reading, incidental L2 vocabulary learning through listening has also shown to enhance L2 lexical learning, although it has received little attention (Vandergrift, 2007). Several authors claim that when trying to make sense of aural input, L2 learners definitely make use of their linguistic and non-linguistic sources (Buck, 2001; Vandergrift, 2007). In order to understand a listening excerpt, Staehr (2009) explains that the L2 learner must apply an intricate combination of bottom-up and top-down

cues. He also claims that the increased difficulty of an aural text is due to its real-time nature and the limited processing capacity of WM, which make the job harder for low-proficiency learners. He also states that the L2 learners' scarce knowledge of the language will allow them to focus mostly on syntactic or lexical units.

Most of the studies featuring L2 vocabulary learning through listening have centred on adult learners (e.g., Brown et al., 2008; Vidal, 2011; Van Zeeland & Schmitt, 2013). The study conducted by Brown et al. (2008) has already been explained in detail in previous paragraphs. Results showed higher significant differences for the reading-while-listening condition when compared to the listening-only mode. In a later study focusing on L2 vocabulary acquisition through listening, Vidal (2011) analysed the relationship between incidental L2 vocabulary acquisition and the number of exposures needed for 230 EFL students attending a university in Spain at different proficiency levels. They were divided into three different groups and treatments: reading three academic texts, listening to three lectures or receiving no input. Significant differences were obtained for vocabulary learning in the reading and listening groups (larger significant differences tended to be seen for the reading condition). The low-proficiency learners obtained the lowest scores as the input they were exposed to might have been too challenging for them. Finally, Van Zeeland and Schmitt (2013) analysed the effects of incidental L2 vocabulary learning through listening with thirty upper-intermediate to advanced students learning English at a British university. Learners were divided into two groups: one group listened to four passages and performed an immediate vocabulary post-test, whereas another group listened to the input and took the same test after two weeks. The results comparing

the immediate and delayed post-tests showed higher significant differences for word form recognition, but not for meaning recognition.

The few existing experiments with young learners on vocabulary learning through listening to stories have shown that L2 learning through listening does occur. Some studies have focused on L2 word meaning through listening to stories with primary school learners and have concluded that children can acquire new words through listening while viewing the images (Elley, 1980; Elley & Mangubai, 1983; Elley, 1989) or only listening (Ashehri, 2014). In addition, research performed with preschoolers listening to stories has also obtained positive results. In Senechal and Cornell (1993), participants listened to a picture storybook and significant improvement was seen between the immediate pre and post-tests and also between the immediate and delayed post-test. Finally, in a study conducted by Collins (2010), teachers read aloud eight similar storybooks three times a week over three consecutive weeks to young learners (4 and 5 years old) and also found significant improvement between the pre- and post-tests.

To summarise, the findings of these studies suggest that L2 vocabulary learning through listening is less successful than that which takes place when reading for incidental L2 vocabulary acquisition. However, as has also been pointed out in the literature, we should take into account that the measures used might not be the most appropriate for gauging knowledge improvement and may not actually show all the knowledge gained by the L2 learners (Van Zeeland & Schmitt, 2013).

4.1.2.2.1. Incidental word-form recognition from reading and listening

Previous research on word knowledge and its dimensions like word-form recognition has been measured through the skills of reading (Waring & Tatakı, 2003; Horst, 2005; Webb, 2007; Pellicer-Sánchez & Schmitt, 2010) and listening (Van Zeeland & Schmitt, 2013; Alshehri, 2014).

Waring and Tatakı (2003) analysed incidental L2 word-form recognition on 15 adult low-intermediate to advanced English learners at university who read a 400 headword graded reader. Participants were given a word-form recognition task that contained twenty-five invented TWs and seventeen distractors. Learners were required to circle any words they recognised from the text. Learners took the test three times (immediately after reading the text, a week after the intervention and three months later). Results showed a significant decrease when examining the data: many of the words participants recognised in the immediate post-test were forgotten: participants only recognised 54.90% of the items after three months.

Another study focusing on word-form recognition through reading was performed by Horst (2005). 21 ESL learners with different proficiency levels (elementary to upper-intermediate) in Montreal were divided into two groups and read graded readers for a period of six weeks, although they did not read the same amount of books and the number differed widely, with a mean figure of 10.52 books. The learners performed 100-item pre and post-tests where they had to indicate if they recognised the word or not with three rating options (NO, NS, YES). The results confirmed significant

differences between the pre- and post-tests, concluding that incidental L2 word-form recognition through ER was taking place.

The research conducted by Webb (2007) is also important for the current study. He analysed frequency of occurrence on L2 word-form recognition with 121 Japanese intermediate EFL learners who were divided into four experimental groups and one CG. The participants read various sentences, extracted from different pages of some titles that belonged to the Oxford Bookworm graded reader series. They were tested on different receptive and productive vocabulary measures after 1, 3, 7 or 10 encounters. In order to measure the receptive knowledge of form, participants took a MC test that contained the correct TW and three distractors in each question. Results on word-form recognition showed a significant increase at seven and ten encounters when compared to only one or three. Pellicer-Sánchez and Schmitt (2010) analysed L2 word-form recognition and meaning with 20 adult advanced ESL learners attending a university in Spain. In order to assess word-form recognition, learners took a MC test that included, in each question, the correct spelling of the word, three distractors (similar spelling to the correct TW) and a 'don't know' option. Participants did an individual interview where they gave the answers (orally) to the researcher. Results indicated that lexical recognition increased significantly the more encounters participants had with the words when reading.

Furthermore, there are also some studies focusing on L2 word-form recognition through listening (Van Zeeland & Schmitt, 2013; Alshehri, 2014). Van Zeeland and Schmitt (2013) analysed the listening effects on different dimensions of L2 vocabulary knowledge such as word form, grammar recognition and meaning recall. Word-form

recognition was measured through an aural MC test with distractors. Results showed significantly higher differences for L2 word-form recognition compared to grammar recognition and meaning recall in the immediate post-test. The participants scored the highest on L2 word-form recognition, indicating that listening was an effective method for L2 vocabulary learning.

Finally, Alshehri (2014) also tested L2 word-form recognition through listening. The participants received aural input through a story that was read to them. Different aspects of L2 vocabulary knowledge were tested: spoken word-form recognition, meaning recognition and meaning recall. In the word-form recognition tests, the learners listened to 20 words twice and circled 'yes' or 'no' according to whether they recognised them from the story or not. In the final results, the four experimental groups significantly outperformed the CG in the three aspects of vocabulary knowledge that were analysed. The results proved that incidental L2 vocabulary learning through listening is an effective approach to perform in formal settings with young low-proficiency L2 learners.

Studies analysing L2 word-form recognition from reading or listening in authentic input show that L2 incidental word-form recognition takes place (Waring & Tatsaki, 2003; Horst, 2005; Webb, 2007; Van Zeeland & Schmitt, 2013; Alshehri, 2014). It is also worth mentioning that the higher results obtained for L2 word-form recognition in most of the studies described above, compared to other lexical dimensions (e.g., meaning recognition), confirms that acquiring the meaning of a word is more difficult than mere word-form recognition, which can be less challenging.

4.1.2.3. Incidental L2 vocabulary learning from multimodal input

Although research on incidental L2 vocabulary acquisition has mainly focused on reading and listening, in the last few years many studies have also analysed L2 vocabulary acquisition from audiovisual materials. Paivio (1986) claims that associating an image with a word aids L2 learners in remembering the lexical item better than if it is presented through text only. Multimedia materials can also be an option to promote vocabulary learning in the classroom (Danan, 2004; Vanderplank, 2010, 2013; Montero-Pérez et al., 2013). In addition, they can engage students in continuing to watch videos in the L2 at home (Kuppens, 2010). According to Rodgers, (2013, p.61), there are three main features of audiovisual materials that aid incidental L2 vocabulary acquisition: “the combination of aural and visual input, the comprehension-focused nature of viewing television and the serial nature of episodes of television”. It has been suggested that watching TV may provide equal gains to reading (Neuman & Koskinen, 1992), highlighting the powerful benefits of L1 and L2 incidental vocabulary learning through subtitled authentic input (Rodgers & Webb, 2011; Montero-Pérez et al., 2013). Several studies have proved that watching subtitled audiovisual programmes leads to larger L2 lexical gains than viewing the same video without subtitles (e.g., Pavakanun & d’Ydewalle, 1992; Neuman & Koskinen, 1992; Koolstra & Beentjes, 1999; Stewart & Pertusa, 2004; Winke et al., 2010; Sydorenko, 2010; Frumuselu et al., 2015; Peters et al., 2016).

As has been mentioned in the previous chapter, watching TV episodes may be more effective than watching films, as the former provides L2 learners with background knowledge in each episode (Webb, 2015). Previous research has proved that

background knowledge (Pulido, 2003) and topic familiarity enhance vocabulary learning (Pulido, 2004). Viewing episodes from the same TV programme provide language learners with a “cumulative build-up knowledge” that they use throughout the whole season of the TV series and help them in understanding the plot (Rodgers & Webb, 2017, p.23). However, the amount of research focusing on the effects of extensive viewing, where learners watch full-length episodes of the same TV series is scarce. We briefly summarise four main studies on extensive viewing and vocabulary learning (Bravo, 2008; Rodgers, 2013; Frumuselu et al., 2015; Gesa, 2019).

Research conducted by Bravo (2008) in her PhD thesis included two studies (the first one has already been described in the previous section). Her second study was conducted with 75 teenage ESL learners who attended Grade 9 at a school in Portugal. The students filled in a questionnaire about their TV viewing habits and they watched ten episodes from season one of the American TV series *The Fresh Prince of Bel-Air* with L1S or L2S. The episodes were divided into two halves and participants took comprehension and vocabulary tests (MC and item sequencing questions) immediately after viewing each video. In addition, before watching each episode, they received a handout that included 10 comprehension and 10 vocabulary items. There were three consolidation tests after weeks five and ten and a final one, which included items from the ten weeks. The vocabulary tests indicated that the L1S group performed better than the learners watching the videos with L2S, who “had more constant scores” (p. 195). For the vocabulary recall and retention tests, results showed a significant effect for the L1S group between the first and the final test. The author claims that for L2 vocabulary learning, L1S are more effective than L2S.

As we have seen previously, Rodgers (2013) performed his study with Japanese university pre-intermediate and intermediate EFL learners who watched ten episodes (one per week) from the first season of the American TV series *Chuck* with L2S and without subtitles and were assessed for L2 comprehension and vocabulary learning. The vocabulary tests comprised MC tests (where the L2 learners had to mark the translation of the TWs presented). The author concludes that there were no significant differences between both groups on L2 vocabulary acquisition, although some significant correlations between vocabulary knowledge and lexical gains were obtained for the L2S group.

More research analysing incidental L2 vocabulary learning through multimodal input has been conducted in the past few years. Frumuselu et al. (2015) performed a longitudinal study on L2 vocabulary learning through watching subtitled audiovisual materials. The 40 beginner to advanced ESL university students, who were divided into two groups, watched thirteen episodes from season one of the American TV series *Friends* for a period of seven weeks (two sessions per week except for the first one) with L1S or L2S. The participants took pre and post MC tests to check recognition of specific TWs and expressions; moreover, they were also tested for word meaning through fifteen open questions. Higher significant scores were obtained by the L2S group in the post-test for all proficiency levels, proving that vocabulary learning was enhanced when subtitled audiovisual materials were watched in the same language as the soundtrack (L2S).

Gesa (2019) conducted his study with participants attending primary, high school and university. At each level, these learners were divided into two groups (experimental

and control) and both were taught some TWs weekly before experimental groups watched the TV series episode subtitled in the L1 (Grade 6) or the L2 (Grade 10 and university). The participants took pre- and post-tests in order to find out their lexical knowledge and meaning of the TWs at the beginning and at the end of each term during the treatment. In addition, a post-test was passed to the primary and high school students eight months after the treatment. The vocabulary results suggest a significant beneficial effect of watching TV series for the primary and high school group, especially in the long term and for intermediate students.

Therefore, most research so far on extensive viewing and vocabulary learning has been conducted with adolescents and adults, with the exception of Gesa (2019), who pre-taught TWs before watching the videos. The mixed results obtained can be attributed to the differences in the design of the studies and the instruments used. It remains to be explored if and how extensive viewing may help beginner learners in acquiring L2 vocabulary.

4.1.3. Factors affecting word learning: evidence from studies on reading, listening and TV watching

Frequency is defined as “the number of occurrences of a linguistic item in a text or corpus” (Richards & Schmitt, 2002, p.232). This means that incidental L2 vocabulary acquisition will gradually occur when a word is encountered in context multiple times, although the required number of repetitions in order to master vocabulary knowledge is still uncertain (Huckin & Coady, 1999; Webb, 2007). However, it has been demonstrated that the words receiving more repeated encounters are more likely to

be acquired than words with a low frequency of exposure (Horst et al., 1998; Webb, 2007; Rodgers, 2013). Therefore, frequency should be considered when teaching L2 vocabulary, especially to young learners (Kokla, 2016). On the other hand, the effect of cognateness and relevance on incidental L2 vocabulary learning has also been examined in previous studies. However, it should be noted that frequency of occurrence has proved to be an essential factor for incidental L2 vocabulary learning (Horst et al., 1998; Webb, 2007; Pellicer-Sánchez & Schmitt, 2010).

Several researchers have tried to find out the ideal number of exposures needed to learn word items. Nation (1990) argues that in order to acquire vocabulary in a successful way, words need between five and sixteen exposures. On the other hand, Horst et al. (1998) suggest that eight encounters are enough to learn a word whereas Webb (2007) expands this number and claims that ten encounters are needed. Laufer (2005) claims that young learners in formal L2 settings are taught explicitly, paying special attention to form and being consciously aware of what they are learning; thus, he concludes that the ideal number of exposures for these learners ranges from six to ten. Furthermore, it has been claimed that word repetition will be more favourable in related texts with recurrent characters, enhancing vocabulary learning through narrow reading (Hwang & Nation, 1989; Matsuoka & Hirsch, 2010).

Hunt and Beglar (2005), among others, provide similar arguments to Nation (1995) about teaching vocabulary explicitly and highlight the need to practice the new words a repeated number of times, featuring frequency of occurrence as an important variable for vocabulary acquisition. Explicit L2 vocabulary learning has been considered the main source for L2 vocabulary learning in formal L2 settings (Laufer &

Paribakht, 1998; Webb, 2008), although research also advocates for the inclusion of incidental vocabulary learning in L2 environments (Nation, 2001; Hunt & Beglar, 2005; Webb, 2008). However, it has been acknowledged that rehearsal is crucial to teach low-proficiency learners explicitly, through repeated encounters over a repeated number of sessions and using several tasks and strategies (Pinter, 2017). This would be more useful than just introducing all the words in one lesson without rehearsal, as it would be more difficult for the student to acquire them, especially if the student is a child learning the L2. Furthermore, frequency has been argued to be significant in formulaic language learning and several studies have demonstrated that high frequency phrases and sentences are processed faster than low-frequency ones (Bod, 2000; 2001; Arnon & Snider, 2010);

The bulk of research available on incidental L2 vocabulary learning has been conducted on reading, which has shown positive results (Horst et al., 1998). However, teachers need to supply the learners with a repeated number of encounters for them to acquire each word (Beglar & Hunt, 2002), especially at low levels of proficiency where more encounters will be needed for words to be learned (Zahar et al., 2001). The number of encounters needed to acquire a word will also vary depending on the dimension assessed, such as word-form recognition, which will require fewer occurrences than meaning recognition (Webb, 2007; Pellicer-Sánchez & Schmitt, 2010). Research on frequency of occurrence has focused on reading (Rott, 1999; Waring & Tataka, 2003; Webb, 2007; Pigada & Schmitt, 2006; Pellicer-Sánchez & Schmitt, 2010) and listening (Van Zeeland & Schmitt, 2013). In addition, in order to find out the better option for L2 vocabulary learning, Vidal (2011) performed one study comparing both skills. Finally, recent studies have also focused on the effects of

frequency of occurrence when learners are exposed to multimodal input (Webb & Rodgers, 2009a; Bisson et al., 2014), as we will see below.

Rott (1999) examined incidental L2 vocabulary learning through reading with learners attending the University of Illinois. Learners took different tests immediately after reading, one week later and after four weeks, which included a recognition exercise, a production task and a MC test. Results showed that although there were significant gains for two encounters, there were higher significant scores for six encounters. The author concludes that six encounters might be enough for incidental L2 vocabulary learning from reading, although frequency has the biggest effect on low-proficiency learners. Another study that measured word-form recognition is the one performed by Pellicer-Sánchez and Schmitt (2010). They measured L2 word-form recognition, spelling, word class and meaning recall. Results revealed that fewer encounters were needed to master L2 word-form recognition compared to the other skills such as word-class recognition or spelling accuracy, implying that more repetitions would be required to learn the other lexical features more appropriately.

Furthermore, the research by Vidal (2011) examined the relationship between incidental L2 vocabulary learning and the number of exposures required, either when reading, listening or receiving no input. Results showed that reading proved to be more effective on L2 vocabulary learning, especially for low-level learners. Frequency of occurrence was the highest contributing variable to L2 vocabulary recognition when reading, with important gains starting at two or three occurrences. On the other hand, participants that listened to the lectures needed more repetitions to acquire the words. The author concludes that vocabulary acquisition through listening might be harder than word learning through reading. She claims that when listening, learners cannot

recheck the input, as they can do when reading a written text. Finally, Van Zeeland and Schmitt (2013) analysed the effects of L2 vocabulary knowledge through listening on L2 word-form and grammar recognition and also on meaning recall. Participants with an upper-intermediate to advanced level listened to four passages from different genres and took a MC test to assess L2 word-form recognition. The results proved that the first ability learners were able to master was L2 word-form recognition, which scored the highest and needed fewer exposures than the other areas examined (grammar and meaning).

On the other hand, when watching television, the number of encounters needed to learn the words depends on different factors related to the input, such as “the clarity of the discourse, the speed of the discourse and the amount of semantic overlap between the imagery and the vocabulary” (Webb, 2011, p.130). Lévesque (2013) also claims that the context and the nature of the words may also play an important role in determining the number of encounters needed to acquire the lexical item. In addition, learners who are exposed to multimodal input might need a lower number of exposures for the words they need to master, due to the presentation of the same information through different channels (Bisson et al., 2014; Lin & Siyanova, 2015). There are a few studies that have analysed the role of frequency of occurrence on multimodal input (Bisson et al. 2014; Webb & Rodgers, 2009a). Bisson et al. (2014) tried to figure out the number of exposures that beginner university students needed to learn a word through audiovisual input. They concluded that incidental L2 vocabulary learning with low-proficiency learners takes place with only two exposures. They also confirm that a higher number of exposures implies better results and they attribute the encouraging results to the presence of pictures in the input that help

present the information in a clearer way. Subsequently, Webb and Rodgers (2009a) analysed the required number of exposures for vocabulary learning with low frequency items in 88 TV programmes and concluded that depending on the type and genre of audiovisual input, the number of exposures differed. On the other hand, children's programmes included a lower percentage (61%) of WFs and were the most accessible category for vocabulary acquisition.

Even though frequency of occurrence has been claimed to be a crucial factor for word learning (Horst et al., 1998; Webb, 2007; Pellicer-Sánchez & Schmitt, 2010), other factors such as cognateness and relevance might affect incidental L2 vocabulary learning (Peters & Webb, 2018). The studies by Vidal (2011) and Peters and Webb (2018) analysed the relationship between incidental L2 vocabulary learning and the variables of frequency of occurrence, relevance and cognateness. The research conducted by Vidal (2011) examined incidental L2 vocabulary learning on 230 participants studying English at a Spanish university. Learners were divided into three groups and read three academic texts or watched three lectures. The third group acted as a CG and only completed the tests. Results revealed a positive relationship for frequency of occurrence, especially for written input (smaller effect for aural input). In addition, there were greater gains for cognates than for other words that had no similarity and the outcomes were remarkably higher for aural input compared to written input. Finally, the relationship between relevance (technical words that were crucial to understand the context of the lecture) and L2 vocabulary acquisition was also demonstrated with high results in aural input.

The other study examining the factors affecting incidental L2 vocabulary learning is the one performed by Peters and Webb (2018). This experiment is very relevant to this thesis as the researchers analysed word learning through watching a single full-length TV programme and its relationship with frequency of occurrence, cognateness and relevance. Learners were tested on meaning recognition, meaning recall and form recognition and the results showed positive outcomes for meaning recall and meaning recognition. Moreover, the variables of frequency of occurrence and cognateness seemed to be related with word learning, with cognateness thought to play a more central role in aural input, in the same way as Vidal (2011). However, no relationship between incidental L2 vocabulary learning and relevance when watching a TV programme was found in the analyses, contrary to the outcomes obtained by Vidal (2011), probably due to the fact that word relevance was determined differently, which could have altered the outcomes.

4.2. Effects of L1/L2 subtitling on vocabulary learning from multimodal input

Even though learning from audiovisual materials may sometimes be difficult for low-proficiency learners, the introduction of subtitles may be a useful way to support them in learning and revising vocabulary through multiple modalities (Winke et al, 2010; Sydorenko, 2010). The presence of L2S may facilitate word recognition, as it will help decoding and isolating words (Winke et al., 2010; Frumuselu et al., 2015). However, L1S have also been shown to be useful (Koolstra & Beenjes, 1999; Lekkai, 2014) as they may help comprehension.

Word-form recognition is one of the many abilities that can be enhanced by subtitles in audiovisual materials (Huang & Eskey, 1999). The use of subtitles has been claimed to assist and reinforce L2 word-form recognition (e.g., Neuman & Koskinen, 1992; Koolstra & Beentjies, 1999; Markham, 1999; Hui, 2007; Sydorenko, 2010; Lekkai, 2014; Montero-Pérez et al., 2014; Peters et al., 2016). The repeated number of encounters is thought to help learners in becoming familiar with new words (Nation, 1990). Even though word-form recognition is just one of the facets involved in vocabulary learning, its significance should not be underestimated (Cameron, 2002). Table 2 presents a chronological summary of several studies analysing the effects of video / TV watching (with or without subtitles) on incidental L2 vocabulary learning and L2 word-form recognition. We have narrowed down vocabulary knowledge to word-form recognition as it is the focus of the present dissertation. The authors, participants, proficiency level, type of audiovisual input and subtitles used in the studies are listed in the table below, as well as a brief summary of the results.

Table 4.2. Previous research on L1 / L2 Subtitling for L2 vocabulary learning (L1S: L1 subtitles, L2S: L2 subtitles)

* Studies on word-form recognition

Study	Participants	Proficiency Level	Video type	Vocabulary test	Subtitles	Results
Neuman and Koskinen (1992) *	Children	Beginners (n= 129)	A science programme for children: 3-2-1 <i>Contact</i> . Three units divided into nine segments (5-8 minutes).	MC word-form recognition test, T / F and concept questions for meaning recognition (90 TWs).	L2S / No subtitles Text, sound / text only	L2S significantly higher than the no subtitles group.
Koolstra and Beenjes (1999)	Children	Beginners (n= 246)	One episode of the documentary series <i>The New Wilderness</i> (15 minutes) about grizzly bears.	MC translation and 30-item aural word recognition tests for learners exposed to English audio (35 TWs).	L1S / No subtitles	L1S significantly higher than the no subtitles group.
Markham (1999) *	Adults	Advanced (n=118)	Two segments of educational videos about whales and civil rights (13 and 12 minutes).	Listening word recognition test (50 TW).	L2S / No subtitles	L2S significantly higher than the no subtitles group.
Stewart and Pertusa (2004)	Adults	Intermediate (n= 95)	Two films: <i>Mujeres al borde de un ataque de nervios</i> (89 minutes), <i>Sexo por compasión</i> (109 minutes).	MC word-form recognition tests.	L1S / L2S	No significant differences but a modest higher score for L2S.
Hui (2007) *	Adults	Beginners (n=92) Advanced (n= 90)	A National Geographic documentary on the Science of Nature. (16 minutes).	Word-form recognition / spelling / meaning tests (10 TWs).	L1S/ L2S / No subtitles	L2S significantly higher for both proficiency levels in recognition and spelling.
Yuksel and Tanriverdi (2009)	Adults	Intermediate (n= 104)	One episode of the TV series <i>Seinfeld</i> . The episode watched "The jacket" (≈ 9 minutes).	Vocabulary Knowledge Scale (20 TWs).	L2S / No subtitles	Significant gains in both groups from pre- to post-test. No significant differences between groups.

Chapter 4 – Multimodal input and L2 vocabulary learning

Study	Participants	Proficiency Level	Video type	Vocabulary test	Subtitles	Results
Sydorenko (2010) *	Adults	Beginners (n= 26)	Russian comedy series: -three video clips (≈2 – 3 minutes each).	Written and aural form recognition tests + translation tests (L2-L1) (28 TWs).	L2S / No subtitles Subtitles w/ no sound	L2S scored significantly higher for written recognition and the no subtitles group for aural recognition.
Nagira (2011)	Adults	Beginners (n= 48)	Two Disney cartoons: <i>The Birthday Machine</i> from Little Einsteins and <i>Tinker Bell</i> (7 minutes each).	VST	L2S / No subtitles	L2S scored significantly higher for <i>Little Einstein's</i> episode but no significant differences for <i>Tinker bell</i> .
Etemadi (2012)	Adults	Advanced (n= 44)	Two documentaries: <i>Dangerous Knowledge</i> (20 min.) and <i>Where is my robot?</i> (30 minutes).	MC (10 TWs).	L2S / No subtitles	No significant differences for vocabulary recognition between groups.
Karakas and Sarıçoban (2012)	Adults	Upper Intermediate (n= 42)	Two episodes from the cartoon TV series <i>Family Guy</i> (≈ 20 minutes).	Vocabulary knowledge Scale (18 TWs).	L2S / No subtitles	No significant differences between groups but significant differences from beginning to end in both groups.
Montero Pérez et al. (2014) *	Adults	High- Intermediate (n= 133)	3 short videos from a current affairs programme (≈ 3 – 4 minutes each video).	Four tests: word-form recognition, clip association, meaning recall, meaning recognition (17 TWs).	L2S / Keyword subtitles / L2S + highlighted keywords / No subtitles	L2S, keyword subtitles and highlighted keywords scored significantly higher than the no subtitles group in form recognition.
Lekkai (2014) *	Children	Beginners (n=93)	A children's cartoon for TV (15 minutes).	Aural word-form recognition and MC translation test (35 TWs).	L1S / No subtitles	L1S scored significantly higher compared to the no subtitles group.

Chapter 4 – Multimodal input and L2 vocabulary learning

Study	Participants	Proficiency Level	Video type	Vocabulary test	Subtitles	Results
Montero Pérez et al. (2015) *	Adults	High-Intermediate (n= 51)	2 authentic short clips from a current affairs programme (≈ 3 and 7 minutes each).	Four tests: word-form recognition, clip association, meaning recall, meaning recognition (18 TWs).	L2S vs. keyword subtitles	Keyword subtitles scored significantly higher than L2S in word-form recognition.
Frumuselu et al. (2015)	Adults	Beginner to advanced (n=40)	Thirteen episodes of the TV series <i>Friends</i> (25 minutes each).	MC word-form recognition (15 TWs / expressions).	L1S / L2S	L2S obtained higher significant differences than L1S at all proficiency levels.
Peters et al.(2016) *	Adults	Study 1: Upper-Intermediate (n=28)	A documentary <i>Eating insects</i> (13 minutes).	Spoken word-form recognition and meaning recall (39 TWs).	L1S / L2S	L2S scored significantly higher.
		Study 2: Low-proficiency to pre-intermediate (n=18)	1 episode from <i>The Simpsons</i> .	Written form recall, written form recognition and written meaning recognition (18 TWs).		No significant differences but higher scores for L2S.
Montero Perez et al. (2018) *	Adults	High-Intermediate (n= 133)	3 short videos from a current affairs programme (≈ 3 – 4 minutes each video).	Four tests: word-form recognition, clip association, meaning recall, meaning recognition (18 TWs).	L2S, keyword, glossed keyword and no subtitles.	All subtitles groups significantly outperformed the no subtitles group in word-form recognition and clip association.
Galimberti and Miralpeix (2018)	Children	Beginners (n=52)	One episode from <i>The Suite Life of Zack and Cody</i> (22 minutes).	A vocabulary pre- and post-test and a MC vocabulary recognition test. (10 TWs).	L1S / L2S / no subtitles	No significant differences for meaning / form recall. Significant differences for vocabulary recall: L2S>L1S but not for the CG.

Chapter 4 – Multimodal input and L2 vocabulary learning

Study	Participants	Proficiency Level	Video type	Vocabulary test	Subtitles	Results
Gesa (2019)	Children Teenagers Adults	Beginner Intermediate Advanced (n=158)	Several episodes of the following TV series: - <i>The Suite Life of Zack and Cody</i> - <i>The Wizards of Waverly Place</i> - <i>I love Lucy</i> - <i>Seinfeld</i>	Active recall of word forms. MC test (40 / 110 / 120 TWs).	L1S / no subtitles L2S / no subtitles	-Significant differences for beginner learners in L1S -Significant differences for intermediate learners for L2S. -Significant differences between groups not found until Term 3.
Muñoz et al. (2022)	Adults	Intermediate (n=67)	One episode from <i>Fresh off the boat</i> (≈20 minutes).	Meaning recall and recognition, sound recognition and reading efficacy. (23 TWs: 9 single words and 14 MWUs).	L2S / no subtitles	Learners with higher results on the sound recognition test scored significantly higher than those with lower outcomes in meaning recognition.

Some of the studies listed above have compared the effects of L2S vs. no subtitles on L2 vocabulary learning (Yuksel & Tanriverdi, 2009; Nagira, 2011; Etemadi, 2012; Karakas & Sariçoban, 2012) and more specifically on L2 word-form recognition (Markham, 1999). Other experiments have analysed the effects of L2 word-form recognition with L2S vs. other conditions, such as eliminating the visual or aural input (Neuman & Koskinen, 1992), removing the sound (Sydorenko, 2010) or adding keyword subtitles (Montero-Pérez et al., 2014; 2015; 2018). On the other hand, the number of studies comparing the effects of L1S vs. no subtitles with multimodal input on L2 word-form recognition is more limited (Lekkai, 2014). Finally, very few studies have compared the effects of L1S vs. L2S on L2 word-form learning (Hui, 2007; Peters et al., 2016).

The studies described in Table 4.2 appear in chronological order but in this section we will classify and describe them according to the learners' age, starting with adult learners. As can be seen, most of the studies have been conducted with adult students (Markham; 1999; Stewart & Pertusa, 2004; Hui, 2007; Yuksel & Tanriverdi, 2009; Sydorenko, 2010; Nagira, 2011; Etemadi, 2012; Karakas & Sariçoban, 2012; Montero-Pérez et al., 2014; 2015; 2018; Frumuselu et al., 2015; Peters et al, 2016). As is often the case, very few studies examining L2 word-form recognition have been conducted with children (Neuman & Koskinen, 1992; Koolstra & Beentjes, 1999; Lekkai, 2014).

4.2.1. Studies with adult learners

One of the first studies analysing L2 word-form recognition is the one performed by Markham (1999). The 118 adult advanced ESL university students watched two

segments of two educational videos, with L2S or without subtitles, and answered 50 MC questions with four options each where they had to identify the correct answer from three distractors. The first video included information about whales (13 minutes) and the second one dealt with the history of the civil rights movement in the USA (12 minutes). Higher significant differences were obtained for L2S, indicating that L2 word-form recognition was enhanced through the use of L2S.

In addition, Stewart and Pertusa (2004) examined L2 vocabulary recognition with 95 adult intermediate Spanish learners attending conversation classes at university. The participants watched two Spanish films, *Mujeres al borde de un ataque de nervios* and *Sexo por Compasión*, divided into three segments each, with L1S (English) or L2S (Spanish) and the Spanish soundtrack. The participants performed pre- and post-MC vocabulary tests and the results showed no significant differences, although a modestly higher score favoured L2S. The authors remark that, although not obtaining significant differences, the learners responded very positively to L2S in the surveys they answered after watching the videos with Spanish subtitles.

On the other hand, the study by Hui (2007) is relevant because it also included L1S. He performed L2 word-form recognition tests on 182 adult students, (90 high-proficiency and 92 low-proficiency freshmen students) at an Institute of Technology in China. All the participants watched a 16-minute National Geographic documentary on the Science of Nature twice, with L1S (Chinese), L2S (English) and no subtitles, whereas the soundtrack was in English for all conditions. The scores of the aural word-form recognition test indicated significant differences between groups, although the L2S group scored significantly higher than the others, followed by the L1S group, and

then the no subtitles group for high and low proficiency levels. The author concludes that subtitled audiovisual materials are beneficial for L2 vocabulary acquisition by complementing the visual cues provided by the images.

Another study analysing the behavior of L2S on L2 vocabulary learning is the one conducted by Yuksel and Tanriverdi (2009). The 104 adult intermediate EFL learners attending a university in Turkey were divided into two groups and watched a segment of the episode *The jacket*, from the TV series *Seinfeld*, with or without L2S twice. The participants performed a Vocabulary Knowledge Scale before and after the treatment to identify if they had understood the lexical units. Significant differences comparing the pre- and post-test were found for both groups. Furthermore, the group watching the video with L2S obtained higher results than the no subtitles group, although in this case no significant differences were found between groups.

Furthermore, in the research performed by Sydorenko (2010), 26 adult beginners of Russian as a FL at a university in the USA watched three short video clips (2-3 min.) of a Russian comedy series with L2S, without subtitles, and with L2S but without sound. Immediately after watching the audiovisual input, the subjects were tested on aural and written L2 word-form recognition and other vocabulary aspects. Regarding the written word-form recognition test, learners had to select the words they recognised from the videos. Even though learners reported difficulties when reading the L2S, due to their speed, results proved that L2S facilitated written L2 word-form recognition, obtaining significantly higher results when compared with the other groups. Significantly higher scores were also found for L1S in aural recognition,

although further research with larger samples and longer videos is suggested before the results can be generalised.

On the other hand, the study conducted by Nagira (2011) used cartoons to measure L2 vocabulary learning. She analysed its effects on 48 adult Japanese beginner EFL university students. The L2 learners watched two segments (7 minutes each) of two Disney cartoons: *The Birthday Machine*, an episode from *Little Einsteins*, and *Tinker Bell* with L2S (English) or no subtitles. A modified version of the Vocabulary Knowledge Scale with only three options (the original version contains five items) was used to check L2 vocabulary learning. The results confirmed that L2S aided L2 vocabulary learning and significant differences were obtained for the *Little Einsteins* episode whereas the scores for *Tinker Bell* did not show any significant findings.

Subsequently, Etemadi (2012) also focused her research on university students. She analysed L2 vocabulary learning on 44 adult participants with an advanced English level who attended a university in Iran and results showed no significant differences between groups. Rodgers (2013) questions these results because there are only ten TWs in the study and the number of encounters is not reported. Moreover, Karakas and Sariçoban (2012) conducted their study on vocabulary development. Participants watched two episodes from the American cartoon TV series *Family Guy* with L2S and without subtitles. A Vocabulary Knowledge Scale was used to test vocabulary development on eighteen TWs and both groups obtained higher significant scores when comparing the results with the pre-tests, leading to the conclusion that subtitled audiovisual materials aid L2 vocabulary learning. However, results did not provide any significant differences between the L2S and the no subtitles group.

The research by Montero Pérez et al. (2014) is very relevant for this section. The form recognition test required learners to answer whether the words presented to them were used ('yes') or not ('no') in the clips they had watched. The groups watching the videos with L2S, keyword subtitles and L2S with highlighted keywords obtained significantly higher scores than the no subtitles groups for L2 word-form recognition and clip association. On the other hand, L2S with highlighted words and keyword subtitles significantly outperformed the other two groups on meaning recognition, although for meaning recall all groups performed in a similar way, obtaining no significant differences. The authors encourage further research with longitudinal studies and longer authentic videos. In a different study, Montero Pérez et al. (2015) examined the effects of word-form recognition, clip association, meaning recall and meaning recognition on 51 upper-intermediate learners at university. Participants watched two authentic short clips from a current affairs programme with L2S vs. keyword subtitles and the form recognition tests required them to answer the question 'Word used in the clips?' by marking 'yes' or 'no'. The findings showed significant differences for keyword subtitles on L2 word-form recognition.

When examining the bulk of research on L2 vocabulary learning with multimodal input, the study by Frumuselu et al. (2015) is worth mentioning. As has been previously described in this chapter, they performed a longitudinal study on vocabulary knowledge with 40 adult students attending a university in Catalonia. The scores obtained by the L2S group in the post-tests were significantly higher than the L1S group, confirming that the students at all proficiency levels improved their vocabulary acquisition when watching audiovisual materials.

The study conducted by Peters et al. (2016) also yielded important outcomes. In the first experiment, participants watched a documentary with L2S and in the form recognition test they were required to mark 'yes' or 'no' whether they had heard or seen the target item before. The findings indicated significantly better results for L2 word-form recognition and there was a positive interaction effect for VS, frequency of occurrence and spoken word-form recognition. In the second study, learners watched one episode of a TV series and performed a written word-form recognition test where they ticked off 'yes' or 'no' to indicate whether they had heard or seen the words before. Results indicated that L2S obtained higher results than L1S for written L2 form recall and written L2 word-form recognition, although no significant differences were obtained between the two groups with subtitles. Results were very similar for written L2 meaning recognition, although L1S obtained better results than L2S. Participants from both studies obtained, in general, better results for L2S and it is suggested that VS and frequency of occurrence might be relevant for L2 vocabulary learning, although further research is encouraged with longitudinal studies, different types of audiovisual materials and larger samples.

In a later study conducted by Montero Pérez et al. (2018), 227 university students watched three French videos with L2S, keyword subtitles, glossed keyword subtitles and no subtitles. Participants were tested on form recognition, clip association, meaning recall and meaning recognition. In the form recognition test, learners had to check whether they recognised ('yes') or not ('no') each of the TWs presented to them. Results showed significant differences for the three groups with captions, who outperformed the no subtitles group in form recognition and clip association. Furthermore, a

longitudinal study was conducted by Gesa (2019), who analysed L2 vocabulary learning with Grade 6, Grade 10 and university learners. Participants from primary and secondary school watched 22 and 24 episodes (respectively) of a TV series over a period of 10 months and university learners watched 8 episodes of a TV comedy. For each age group, the experimental condition watched the videos with L1S (Grade 6) or L2S (Grade 10 and university) and completed several tasks before, during and after viewing each episode. The pre-viewing activity involved teaching 5 target items from the episode, whereas the post-viewing tasks examined comprehension and vocabulary learning. On the other hand, CGs performed the pre-and post-viewing activities but did not watch the TV episodes. Results showed that experimental groups for Grade 6 and Grade 10 scored higher than CGs on form and meaning recall, but significant differences were not always obtained for all episodes. No significant differences were obtained when comparing the experimental and CGs of university learners.

Finally, a recent study by Muñoz et al. (2022) examined L2 vocabulary learning through repeated viewing of one episode from the TV series *Fresh off the boat*. The 67 participants were divided into three groups and watched the L2 subtitled episode with immediate or spaced repetition. A CG was not exposed to any multimodal input and only took the pre- and post-tests. Learners took a sound recognition test and meaning recall and recognition vocabulary tests and they were also tested on reading efficacy. Results showed that participants in the spaced repetition group scored significantly higher than the CG in meaning recognition. Moreover, time was found to be a significant factor for meaning recognition and meaning recall and learners with higher results on the sound recognition test scored significantly higher than those with lower outcomes in meaning recognition. Reading efficacy was not found to be significant,

even if it correlated with the students' results.

4.2.2. Studies with young learners

The studies described below, which have young learners as their participants, focus on the effects of multimodal input on L2 vocabulary learning and word-form recognition through multimodal input (as word-form recognition is assessed in the present dissertation). They have been conducted with either L1S (Koolstra & Beentjes, 1999; Lekkai, 2014; Gesa, 2019), L2S (Neuman & Koskinen, 1992) or comparing L1S vs. L2S (Galimberti & Miralpeix, 2018). In addition, different materials and tests have been used and to date they have reported different findings, so further research is needed that takes these findings into account.

Neuman and Koskinen (1992) compared the effects of watching a science programme for children: *3-2-1 Contact* under different conditions. Three units of this programme were divided into nine segments (5-8 minutes) and participants (129 bilingual low-proficiency young learners attending Grades 7 and 8 at a middle school in the USA), who were divided into four groups, watched it with L2S, without subtitles, text combined with sound or text only. Learners performed pre-tests to measure their knowledge of the TWs and after watching each segment, they took a L2 word-form recognition post-test and a writing test to measure the frequency in which they used the TWs in their writing. Results showed that the students who watched the videos with L2S scored significantly better than the other groups. A retelling task was also performed by participants to assess the frequency with which they used the TWs.

Furthermore, results also proved that vocabulary acquisition was enhanced when L2S were presented with other types of input in video materials.

Moreover, in a study considered one of the first to use subtitles for language learning, Koolstra and Beenjes (1999) analysed vocabulary learning for 246 young learners. They attended Grades 4 and 6 at a primary school in The Netherlands and were divided into three groups. They watched a documentary about grizzly bears in English (aural input) with L1S or without subtitles. The third group acted as the CG using another version of the video material with a Dutch soundtrack and no subtitles. The two groups exposed to the materials in English took an aural word recognition test and results indicated that multimodal input aided the subjects in obtaining the best results. The significantly higher scores obtained for L2 vocabulary learning by the L1S group also proved that subtitles in the learners' L1 did not imply any distraction for L2 vocabulary learning. Another remarkable aspect is that the proficiency level of the subjects mattered: Grade 6 learners, who had already received English language instruction, significantly outperformed the Grade 4 group (no previous English lessons at school).

Subsequently, the research conducted by Lekkai (2014) analysed L2 aural form and meaning recognition on 93 beginner young learners (9–12 years old), who attended primary school (Grades 4, 5 and 6) in Greece. Participants were divided into three groups and watched 15 minutes of a children's TV cartoon twice, with L1S or without subtitles and Italian (L2) soundtrack, whereas the CG watched the cartoon with Greek soundtrack but no subtitles. Immediately after watching the materials, participants took a test to check for L2 aural word-form recognition and a MC test to examine L2 vocabulary acquisition. The word-form recognition test contained thirty words (twenty

TWs plus ten distractors) and learners had to answer whether they recognised ('yes') or not ('no') the Italian words they heard. Results showed a significant effect for L2 vocabulary learning and recognition in favour of the L1S group, and the students' age had an effect on the results: grade 6 learners outperformed those in grade 4. This outcome suggested that more hours of instruction implied a higher score in the tests, although the scores were not significant for L2 aural word-form recognition.

A more recent study by Galimberti and Miralpeix (2018) was also conducted with low-proficiency EFL young learners attending grade 6 at an Italian public school. Participants watched a full (20-minute long) TV episode of a series with either L1S, L2S or no subtitles. No significant differences were obtained for meaning and form recall when comparing the pre- and post-tests. On the other hand, the L2S group obtained significantly higher scores than the L1S group in vocabulary recall.

One of the latest studies analysing the effects of multimodal input on L2 vocabulary learning with Grade 6 learners was performed by Gesa (2019). This study has already been explained in detail in the previous section. However, it would be useful to recall here that significant differences did not tend to be found until Term 3, when exposure to multimodal input had accumulated. Results suggested that sustained exposure over time is beneficial for L2 vocabulary learning. However, the results were not maintained after eight months when a delayed post-test was administered.

Although it has been proved that audiovisual support aids L2 vocabulary learning and more specifically, word-form recognition, more studies need to be conducted with authentic materials from different genres in order to find out what happens with low-level young learners and which type of subtitles would be better (Peters et al., 2016).

Even though L1S and L2S have proven to be effective for L2 vocabulary learning, the studies presented above are conducted with different types of input (authentic, modified or adapted for the learners' level) of different lengths and with learners at different levels. Several trends have been described, but it is difficult to generalise results until more evidence is provided. In addition, there is a lack of longitudinal studies, except for Rodgers (2013) with adult learners and Gesa (2019) or Pujadas (2019) in younger learners. Thus, the present dissertation focuses on an under-researched population (young children) and concentrates on checking the role of L1/L2 subtitles for comprehension and vocabulary learning in English.

CHAPTER 5 – INDIVIDUAL DIFFERENCES AND OTHER FACTORS THAT MAY INFLUENCE L2 LEARNING FROM MULTIMODAL INPUT

This chapter discusses the role of aptitude and age in L2 learning, as well as that of other variables such as L2 proficiency level, VS and RS, (focusing especially on learning from multimodal input). Section 5.2 deals with aptitude, how it can be measured and introduces research with this variable on L2 learning, with special emphasis on studies with children and an introduction to the few studies available on aptitude and learning from viewing. Next, section 5.3 reviews age and L2 learning, which is of crucial importance for this study as our participants are young learners. Furthermore, we introduce the influence that OSE, recently to audiovisual input, has been shown to have on L2 learning, especially for young learners. Finally, section 5.4 focuses specifically on the variables of L2 proficiency, VS and RS, and on their relationship with L2 learning from multimodal input.

5.1. Learners' Individual Differences (IDs): aptitude and age.

Research has shown that individual differences (IDs) play an important role in both L1 and L2 learning (Robinson, 2001). IDs are characteristics or traits that distinguish individuals from each other and are crucial in the whole process of acquiring a language (Dörnyei, 2005). Learners differ in their abilities to acquire the L2 and several variables such as age, aptitude and motivation have been identified as influences on learning (Hummel, 2009). Similarly, Selinker (1972) argues that IDs are decisive in L2 learning, although L2 research has focused more on the development of linguistic aspects rather than the non-linguistic features (Zafar & Meenakshi, 2012).

Skehan (2002) lists four main IDs influencing L2 learning: language aptitude, learning style, motivation and learning strategies, whereas Ellis (2008) lists seven: age, language aptitude, learning style, motivation, anxiety, personality, learners' beliefs and learning strategies. Among them, the present dissertation focuses on aptitude and age, as it examines learning from multimodal input in young learners, taking their aptitude into account.

Aptitude has been claimed to play an important role in L2 learning (Skehan, 1989; Muñoz, 2010) and it has been identified as “one of the most promising areas of SLA research” (Dornyei, 2005, p.63). As Grigorenko et al. (2000) mention, the role of L2 aptitude is evident as anyone can acknowledge that some people are more talented than others or are able to learn an L2 in a better and more effortless way. In addition, it has been shown that aptitude scores can predict the rate of progress for beginner language learners (Doughty, 2019). Another crucial ID to consider in SLA research is age (Singleton, 1989; Birdsong, 1999; Muñoz, 2008a; 2008b; 2010; 2017; Muñoz & Singleton, 2019). The controversy with the Critical Period Hypothesis by Lenneberg (1967) for the L1 and its different interpretations and translations for the L2 have resulted in a range of studies comparing young and adult learners in formal and informal settings.

5.2. Aptitude

5.2.1. Definition and measurement

Aptitude is one of the three cognitive abilities that authors claim are involved in L2 learning, together with intelligence and memory (Ellis, 2008). L2 aptitude is usually described as a specific talent that learners possess to acquire the L2 (Wen et al., 2017) and it is recognised as a good predictor of rate in the early stages of L2 learning (Carroll, 1973). According to Carroll (1989, p.26), aptitude is “the amount of time a student needs to learn a given task, unit of instruction, or curriculum to an acceptable criterion of mastery under optimal conditions of instruction and student motivation”. Carroll proposed a four-component model of aptitude, which is still depicted as the most prominent (Skehan, 2012). The four components included in Carroll’s model (1981) are: phonemic coding ability (the ability of the learner to code and retain unknown sounds and create mental connections with phonetic symbols); grammatical sensitivity (the capacity to identify the functions of words in a sentence); inductive language learning ability (the capacity to infer the linguistic structures from an unknown given corpus) and associative memory (the ability to shape links in memory). Skehan (1989, 1998) proposes a similar framework where he identifies three abilities: phonetic coding ability, language analytic ability, and memory ability. Overall, research coincides that in order to have a good aptitude, learners must be able to learn quickly, identify and memorise new sounds, figure out the function of words in phrases, comprehend grammatical rules and memorise new words.

Several tests have been designed taking into account the four components from Carroll's model to measure L2 aptitude and predict learning rate in formal L2 settings, such as The Modern Language Aptitude Test (MLAT) (1959) and the MLAT-Elementary (1967) by Carroll and Sapon (1959), The Pimsleur Language Aptitude Battery (Pimsleur, 1966) or the LLAMA language aptitude tests (Meara, 2005). High correlations between instructed SLA and the results of these tests have been obtained in several studies (Skehan, 1989, 1998, 2002; Dörnyei & Skehan, 2003; Sparks et al., 2009; Grañena, 2012). In the following sections we will focus on two of the most widely used aptitude tests: the MLAT and the LLAMA tests.

5.2.1.1. The Modern Language Aptitude Test (MLAT)

According to Li (2019), the most widely used aptitude test in the past six decades is the MLAT by Carroll and Sapon (1959). The MLAT measures the four components from Carroll's model through simple language activities. There are five subtests, which cover number learning, phonetic script, spelling cues, words in sentences and paired associates. Each subtest has a different number of questions and a point is awarded for each correct answer (points are not deducted for incorrect answers). These are the maximum scores for each subtest: number learning (43 points), phonetic script (30 points), spelling cues (30 points), words in sentences (50 points) and paired associates (24 points).

Research has proved the MLAT to be a strong predictor of L2 proficiency and more efficient than other tests (Ehrman & Oxford, 1995; Sparks & Ganschow, 2001; Sawyer & Ranta, 2001; Sparks et al, 2009). In their research examining L2 aptitude, Winke

(2005) and Robinson (2005) claim that the components measured in the MLAT tests were correlated with beginner but not with advanced learners. Hummel (2009) obtained completely different results in his study, in which the components were associated with advanced learners, too. In addition to these outcomes, a meta-analysis conducted by Li (2015) with 34 FL aptitude studies showed a statistically significant positive correlation between the MLAT scores and ultimate attainment with either L2 children or adult learners.

Even though there is not much research on L2 aptitude with young low-proficiency learners, there are some tests specifically destined for these ages. Carroll and Sapon also developed an adapted version of this test for primary young learners (Grades 3 to 6), the MLAT-Elementary (1967), which consists of four parts: hidden words, matching words, finding rhymes and number learning. However, few studies have been conducted with the MLAT-Elementary on young learners (e.g., Suárez, 2010; Rosa, 2011; Muñoz, 2014b). Another aptitude test specially designed for children is the Pimsleur Language Aptitude Battery (Pimsleur, 1966), also based on Carroll's model. The latest version of this test comprises six sections, taking into account: grade point average in academic areas other than FLs, interest in learning a FL, vocabulary, language analysis, sound discrimination and sound-symbol association.

5.2.1.2. The LLAMA language aptitude tests

The LLAMA tests (Meara, 2005) were created from a series of projects developed by students of English Language and Linguistics at the University of Wales, Swansea. The design comes from the need to create aptitude tests that were independent of any

L1, taking into account that this form of assessment is used worldwide. The LLAMA tests are a set of four language-neutral sub-tests, based on components from the standardised MLAT tests by Carroll & Sapon (1959), but using the technological advances available at the time to make them more attractive. These four tests are designed to assess L2 aptitude measuring vocabulary learning (LLAMA_B), phonetic (implicit) memory (LLAMA_D), sound-symbol correspondence (LLAMA_E) and grammatical inferencing (LLAMA_F). The LLAMA tests are free and easily available and have become increasingly popular in the past few years, which is shown by more than 700 citations on Google Scholar since 2013 (Rogers et al., 2017). Although Meara warned about their possible lack of validity in 2005, validation studies conducted over the past few years have shown otherwise (Grañena, 2013a; Artieda & Muñoz, 2016; Rogers et al., 2017). For example, a very recent study conducted by Bokander and Bylund (2020) showed that the LLAMA_B test has good internal validity to assess linguistic aptitude for research purposes.

The LLAMA_B test is a vocabulary learning task that measures how good you are at learning words in another language in a very short period of time (two minutes). Based on picture stimuli, it is a simple computer-based task in which students see some visual stimuli. Participants are given two minutes to learn the names assigned to the pictures, which are language-independent. After that, the test sequence starts, the name of an object is displayed on the screen and the test takers have to match the picture with the name. Each correct answer is awarded a point and the final score is displayed on the bottom panel at the end of the test: scores can range from 0 to 20.

5.2.2. The role of aptitude in L2 learning

Carroll (1962) and Skehan (1989) were pioneers in researching the role of language aptitude in L2 learning. Later on, researchers such as Dörnyei and Skehan (2003) claim that language aptitude is one of the IDs which predicts the success of SLA most consistently. L1 and L2 aptitude have been claimed to be related and research has argued that learners who acquired the L1 quickly would also do the same when learning the L2 (Sparks & Ganschow, 1991; Sparks et al., 2006). Skehan (1986) claims that there can be a transfer of L1 skills to L2 aptitude, which has been illustrated in longitudinal studies (Skehan & Ducroquet, 1988; Sparks et al., 2009). Skehan (1988) considered language aptitude a relatively fixed ID, although recent studies have claimed that L2 aptitude is not stable and it changes as children develop cognitively (Milton & Alexiou, 2006; Suárez & Muñoz, 2011).

The relationship between proficiency and aptitude has been examined in the literature. Even though Dekeyser (2000) found no relationship between L2 proficiency and aptitude in young learners, L2 aptitude has strongly correlated with L2 proficiency in some studies, although research on this variable is still scarce (Erham & Oxford, 1995; Hummel, 2009; Abrahamsson & Hyltenstam, 2008; Sparks et al., 2009; Zafar & Meenakshi, 2012). Li (2019) published an overview on the L2 aptitude research performed in the last six decades and he claimed that aptitude was a strong predictor of general L2 proficiency. However, he highlights that “it seems less predictive of L2 writing and vocabulary learning” (p.93). On the other hand, Dahlen and Caldwell-Harris (2013) showed that university students with a higher language aptitude showed greater recall and recognition of unknown TWs than the participants with a low L2

aptitude. It should be noted, though, that the learners of this study had no knowledge of the L2 (Turkish) before performing the experiment. The outcomes of this study are in line with Doughty (2019) who claims that aptitude can predict L2 language learning development at initial stages, but not when the students reach a more advanced level and their L2 proficiency improves.

Research studies on L2 aptitude have recently focused on WM, which “involves the temporary storage and manipulation of information that is assumed to be necessary for a wide range of complex cognitive activities” (Baddeley, 2003, p.189). WM has a significant role in SLA influencing L2 processing (Suárez, 2010) and it has also been proposed as a component of aptitude (Dörnyei, 2005). Wen and Skehan (2011:21) conclude that it can be possible to consider WM as a key component in L2 aptitude as long as there are differences in WM among L2 learners that can be reliably and validly measured, and if it is shown that WM plays a constant and significant role in SLA processes and L2 development.

Hummel (2009, p.243) argues that WM can “predict proficiency in early L2 learning stages in children and adolescents to non-novice adult learners as well”. A study by Muñoz (2014b) analysed the L2 aptitude of 48 young learners attending Grades 5 and 6 (aged 10-11 and 11-12) in a primary school. Learners took the MLAT-Elementary and several language tests to measure their speaking, reading, writing and listening ability. Results indicated that the MLAT-Elementary can predict achievement in beginner young learners. Furthermore, the scores for memory abilities were only slightly

stronger when compared to other aptitude components, suggesting that young learners “rely on memory to a large extent” (p.64) and that high achievers are those who possess superior analytical abilities.

Language aptitude has also been examined in studies in formal and informal settings. Research on L2 aptitude has been conducted in formal L2 settings examining learners’ rate (Ehrman & Oxford, 1995; Harley & Hart, 1997; Muñoz, 2017). Skehan (1989) claims that even though L2 aptitude can predict rate in formal L2 settings, L2 aptitude results will be more relevant in naturalistic environments due to the considerable exposure to huge amounts of input that learners receive and the cognitive strategies they should have to acquire the language implicitly. The studies conducted in naturalistic contexts have analysed the participants’ ultimate attainment and debated on the existence of a Critical Period for L2 learning on children and adults at different proficiency levels (Dekeyser, 2000; Abrahamsson & Hylstenstam, 2008, Dekeyser et al., 2010; Grañena & Long, 2012). Results of the studies performed in both contexts prove that L2 aptitude has positive effects on L2 learning in both instructed and naturalistic settings. As the present study was conducted in a formal setting, the next section focuses on the studies conducted in this context with both young and adult learners.

5.2.3. Aptitude in young and adult learners

Recent research has shown that, unlike adults, children can achieve a native-like proficiency level regardless of their L2 aptitude (Harley & Hart, 1997; Dekeyser, 2000). However, it seems that adult learners need to possess a high L2 aptitude in order to

attain a near-native proficiency in the L2 (Abrahamsson & Hylstenstam, 2008). Robinson (2005), in his study on the role of aptitude in SLA claims that adult L2 learners are much more dependent on the analytic abilities measured by the aptitude tests, which gives them an advantage over young learners. Children, on the contrary, rely more on their memory abilities, as they have not fully developed their analytic abilities. Due to that, young learners may score lower when taking these tests. Research has been performed analysing the L2 aptitude of young and adult learners in formal (Harley & Hart, 1997; Sparks et al., 2009; Suárez, 2010; Rosa, 2011; Muñoz, 2014b) and naturalistic settings (Dekeyser, 2000; Abrahamsson & Hylstenstam, 2008; Grañena & Long, 2012). However, even though there is a recent interest in young learners' language learning aptitude due to the global increase of FL learners, the number of studies with young participants is still scarce (Muñoz, 2014b).

Some studies conducted with adults in naturalistic settings have proved that talented adult L2 learners who possess a high L2 aptitude can attain a near-native level in the L2 (Ioup et al., 1994; Bongaerts et al., 1997). For example, Ioup et al. (1994) analysed the case of Julie, an adult woman considered an exceptionally talented L2 learner due to her accomplishment of acquiring Egyptian Arabic in a natural environment after the puberty stage. The scores proved that Julie was an exceptional learner due to her high L2 aptitude that helped her attain a native-like proficiency level. Subsequently, Bongaerts et al. (1997) also report that some late L2 learners can attain a native-like level. In another study, Dekeyser (2000) analysed the L2 aptitude of 57 Hungarian native speakers living in the USA for a minimum of 10 years with English as their L2. The study showed that some talented late L2 learners with a high level of verbal analytical ability obtained scores within the child acquirers' range.

Another study performed by Abrahamsson and Hylstenstam (2008) analysed L2 aptitude and proficiency on 42 early and late-starter learners of Swedish with Spanish as their L1. The outcomes showed that the late learners' group obtained higher significant results than the early learners' group. However, Abrahamsson and Hylstenstam do not find any relationship between proficiency and aptitude. The authors also confirm that even with a high L2 aptitude level, L2 late learners would not be able to reach a native-like level and they also claim that even for child learners, an early immersion in the L2 may not be enough to become native-like. Finally, Grañena and Long (2012) analysed L2 aptitude, measured with the LLAMA tests. Significant differences were found depending on the age of onset. On the other hand, aptitude and / or length of residence were related to ultimate L2 attainment. Regarding aptitude, they claim that it has an impact on the areas of lexis and collocations.

Rogers et al. (2017) conducted their research on L2 aptitude with adults and children. The authors included the young learners in their study to check the validity of the LLAMA tests for this population due to the lack of reliable testing instruments for children to predict their future L2 learning outcomes. They examined the possible effects of several individual variables such as language spoken, age and formal education for 229 participants of diverse ages (10-75 years old) and different language backgrounds. The scores of the tests indicated no significant differences for gender and language. However, the age variable provided diverse results: the adult participants obtained significantly higher results than young learners (10-11 years old) on the LLAMA_E whereas young learners scored significantly higher than adults on the LLAMA_B. Having received formal education at a postgraduate level also impacted results for LLAMA_B, E and F positively. The authors also recommend the

use of the LLAMA tests with other types of learners, as the learners did not have any difficulties taking the tests, independently of age.

In relation to aptitude and young learners, Harley and Hart (1997) analysed L2 aptitude on 65 Grade 11 early and late L2 French immersion learners in four classes from several schools in Ontario, Canada. The students from two classes (n=36) were part of an immersion program that started in Grade 1, where they received 50% of their lessons in the L2 and the other 50% in English. The students from the other two classes (n=29) had started the immersion program at Grade 7 and they also received the same percentage of lessons in the L2 (50%) and the other 50% in English. This group had also received 40 minutes of French instruction per day from Grades 4 to 6. Participants performed two memory tasks and took an analytical ability test as well as several proficiency tests. Although it was expected that the early immersion students would obtain better L2 aptitude scores than the late immersion learners, this was not the case and aptitude was concluded to be stable and related with L2 proficiency in both early and late immersion groups. However, it was also seen that early immersion was better correlated to memory components (implicit learning) whereas late immersion had a stronger relationship with analytical components (explicit learning).

Suárez (2010) administered the MLAT-Elementary test and she also adapted and translated it into Catalan (MLAT-EC). Both tests were administered to 629 Spanish-Catalan bilingual young learners (8–14 years old) and a series of language tests were also taken by the participants. Results showed higher means between grades 3 and 4 than between grades 4 and 6 and there was a plateau between grades 6 and 7 in aptitude. Regarding aptitude and proficiency, the correlations with the English marks

of the learners were of low to moderate significance. The other proficiency measures correlated with all parts of the MLAT in different ways. However, the author concluded that the total score of the MLAT was the best predictor of proficiency. Rosa (2011) also examined the impact of language aptitude through the MLAT-Elementary on the L2 outcomes of 48 Spanish-Catalan bilingual learners of English attending grade 5 at a primary school in Barcelona (Catalonia). The results showed a strong relationship between language aptitude and the listening, writing and reading scores of the participants. Another study where the MLAT- Elementary was used was performed by Muñoz (2014b). She based her research on the previous study by Rosa (2011) and she intended to explore whether or not language learning aptitude was significantly related to young learners' FL proficiency. Her study comprised 48 young learners (10-12 years old) who were Spanish-Catalan bilingual students of English attending a primary school in Catalonia, Spain. The aptitude scores of the children were compared with their ability in language skills such as speaking, listening, reading and writing and the results showed significant correlations with all language dimensions. Therefore, all these studies show that aptitude is strongly related to the L2 proficiency of young learners in instructional settings.

Just very few studies have explored the potential relationship between aptitude and learning from multimodal input. Teng (2022) is a one-off study on the influence of aptitude on vocabulary learning in adult learners: participants viewed a full-length BBC documentary with or without L2S and completed several vocabulary tests after watching it. The results show that language aptitude significantly influenced the vocabulary scores. To our knowledge, though, only two longitudinal studies have been conducted on TV viewing taking aptitude into account. One is Suárez and Gesa

(2019): EFL learners attending Grade 10 (n=57) and university (n= 60) were pre-taught several TWs during one academic term whereas the experimental group also watched a weekly episode from a TV series that contained the pre-taught TWs. Significant differences were found between the experimental and CGs. A main effect for proficiency was found in form and meaning scores, however, language aptitude was only significant for meaning. The study suggests that proficiency and aptitude, the latter only to a certain extent, influence L2 vocabulary learning through authentic multimodal input. Hence, L2 aptitude impacts L2 learning in adult language learners (Dekeyser, 2000), although the role of L2 aptitude for young learners has also shown to be remarkable in more recent studies (Abrahamson & Hylstenstam, 2008; Sparks et al, 2009; Suárez, 2010; Rosa, 2011; Muñoz, 2014b; Rogers et al., 2017). It remains to be explored if it is actually an influential factor in extensive viewing, especially in young learners. The second study is Gesa and Suárez (2022), examining the role of aptitude in L2 vocabulary learning in Catalan/Spanish primary school learners watching eight episodes from a TV series with L1S during an academic term. The results revealed a significant relationship between language aptitude and word-form learning and indicate that students with higher aptitude showed greater gains in the word-form test.

5.3. Age

Age has been considered a critical issue in SLA and a lot of research has been conducted in this area (Singleton, 1989; Birdsong, 1999; Muñoz, 2008a; 2010; 2017; Muñoz & Singleton, 2019). Studies examining the differences in L2 learning between young and adult learners have shown young learners to have an advantage in

naturalistic settings in the long run (Muñoz 2008a, 2008b). In relation to learning rate of young and adult learners, it has been argued that adult L2 learners will learn in a faster way, both in naturalistic and formal settings, whereas children will eventually catch them up and attain higher proficiency levels in the long run in naturalistic settings (Krashen, et al., 1979). The question regarding the most suitable age to start L2 learning in formal settings has been a crucial issue that has attracted lots of interest in the field of SLA. It is not clear whether children outperform adults in formal settings in the long run, as the input they receive is highly limited (García Mayo & García Lecumberri, 2003; Muñoz, 2006; 2008a). Muñoz and Singleton (2019, p.223) adequately conclude that “the effects of age differ according to the learning environment”. They also claim that learners’ attitude and motivation “are not necessarily connected” (p. 223) with the starting age of learning a language.

5.3.1. Young learners and FL learning

Research has pointed out that children, due to their young age, use implicit learning mechanisms that facilitate language learning and success in naturalistic settings (Dekeyser, 2000; Dekeyser & Larson-Hall, 2005; Ellis, 2005; Paradis, 2009). Bialystok (1997) claims that there are processing differences between children and adults. Children tend to learn implicitly, whereas adults need to make more use of their explicit learning mechanisms. Adult learners also possess some advantages when learning explicitly, and this implies a quicker rate and a shorter time period than when learning implicitly (Dekeyser & Larson-Hall, 2005). Muñoz (2006; 2008b; 2010) points out that even though young learners are able to acquire the L1 implicitly, due to the huge quantities of input they are provided with, they tend to learn the FL explicitly in L2

formal settings. She also argues that in instructional settings young learners are not provided with all of the input that their implicit learning mechanisms require to learn the L2. Singleton (1995) argues that in formal settings, more than eighteen years would be necessary to show benefits comparable to those of starting to learn a language at an early age in a naturalistic context. Therefore, in instructed settings, young students do not show any long-term advantage due to the limited amount and type of input they receive in L2 classrooms (Singleton, 1995; Muñoz, 2006; 2008a; 2010).

Studies conducted in formal settings with a constant amount of exposure and on learners of different ages have proved that late learners show a quicker rate of learning (Muñoz, 2008a), which has been demonstrated in the Barcelona Age Factor project (BAF). The BAF project analysed the effects of age on EFL learners with three different amounts of exposure at three time intervals. The written and oral receptive and productive skills of the participants were examined and the scores indicated that the late starters outperformed the early starters given the same amount of exposure. In addition, the early starters showed a slower rate of acquisition which increased as they grew older, while the late starters, which included adolescent and adult beginners, demonstrated a faster rate. These results are in line with other studies that compared early and late starters in formal settings and the key lies in exposure which, as Muñoz (1997, p.21) claims, “may be as crucial as the age at which initial exposure takes place, that is, the age at which pupils begin their instruction in the foreign language”.

It is evident, then, that findings from L2 studies performed in naturalistic settings with children have mistakenly been generalised to L2 formal settings and the recognised advantages from an early start in naturalistic settings have been erroneously

transferred to instructed L2 learning environments (Dekeyser, 2000; Muñoz, 2010). Furthermore, research has shown that in L2 formal settings, the statement ‘the younger the better’ (Scovel, 2000) will only be true when L2 learners receive a sufficient amount of input, which is not the case in most schools (Cenoz, 2002; García Mayo & García Lecumberri, 2003). Even if European primary schools have introduced FLs earlier in their curricula, children are not provided with enough significant input in this language and they receive it in the form of “isolated words and building sentences” (Muñoz, 2014a, p.37). Thus, this type of learning does not provide them with sufficient exposure to input in the FL classroom (Lightbown & Spada, 2006). Therefore, Muñoz (2008a) suggests that the advantage in ultimate attainment that young learners might have in naturalistic settings will not take place in L2 instructional settings, due to the lack of quality exposure and input that will make it impossible for these students to reach native-like levels. It would be ideal, then, to promote implicit learning for these learners either in or out of the classroom.

According to Pinter (2017) and Bland (2018), the best way to acquire an L2 and infer meaning for children is through fun exercises. FL teaching should provide a wide range of opportunities, such as exposure to authentic texts with pictures and natural language in order to engage children in the process of guessing. Therefore, Bland (2018) claims that FL lessons in primary classrooms should be motivating and supports the use of authentic materials for young learners, although she advises to carefully select adequate ones, implying that they should not be too challenging for their level. As has been previously acknowledged, young low-proficiency learners acquire a FL more easily by concentrating on meaning (Neuman & Koskinen, 1992) rather than on form or grammar, preferably when the input presented to them is just

slightly above their current proficiency level (Krashen, 1985). However, teachers face diverse challenges in the FL classroom, such as the small amount of time (e.g., 2-3 hours in a whole week).

In order to explore how learning can be maximised for young learners in EFL classes, the role of the teacher-researcher is often crucial when conducting classroom research. According to Wood (2008, p.1), teacher research is the “research which is conducted by classroom teachers on their own practice”. Hammersley (1993) claims that there are certain advantages and disadvantages teacher researchers have. On the one hand, outsider researchers are mere observers and only come into the classroom to collect data for a limited amount of time. On the other hand, teachers have a great advantage with the “long-term experience” they acquire with the participants and the settings where research is being carried out. The relationships the teacher builds with the students or other teachers from the same school allow him/her to have important additional data for the study. Although the positive benefits of conducting teacher research are undeniable, there is also one important disadvantage: teacher-researchers have been claimed to be subjective in their decisions (Wood, 2008) and influenced by their already established relationships with the participants. As a result, they need to be as detached as possible when conducting research in their own language classes.

5.3.1.1. Out-of-school exposure (OSE) to English in young learners

Digital media has changed our lives and especially those of children, who nowadays are used to technology from when they are very young onwards (Butler, 2019). These

learners are often exposed to multimodal input in the FL, which implies more opportunities for English language learning outside school (Sundqvist & Wikstrom, 2015). Young learners do sometimes perform out-of-school activities in their L2 (English), such as watching TV series or movies (Prensky, 2001), playing computer games (Aghlara & Hadidi, 2011) or going through social media (De Wilde & Eyckmans, 2017). They are responsible for choosing these activities and, due to that, the motivation to perform them is probably high (Sylvén & Sundqvist, 2012).

English is the dominant language in movies, TV programmes and video games and, because of that, young learners have the unique opportunity to be supplied with great chances of incidental L2 learning without any teacher instruction (Jensen, 2016; De Wilde & Eyckmans, 2017). There has been an increased exposure to FLs, especially English, due to the appearance of several streaming platforms (Muñoz & Cadierno, 2021) where viewers can choose what to watch at any time. Young learners can spend long hours watching TV in their L2 (English) at home (Lindgren & Muñoz, 2013; Sundqvist & Sylvén, 2014). However, even though there are some countries where children get input in English very often, such as Sweden (Sundqvist, 2009; Sylvén & Sundqvist, 2012; Sundqvist & Sylvén, 2014) or Belgium (Kuppens, 2010, Peters, 2018), in other countries like Spain, which also offer the opportunity to watch TV in the L2 (with or without subtitles), this practice is far less popular (e.g., see Muñoz & Cadierno, 2021 for a comparative study in Denmark and Spain).

In addition, most of the videogames that children play tend to be in English. Research has also focused on the specific activity of playing digital games at home and its influence on young learner L2 proficiency (Turgut & Irgin, 2009; Kuppens, 2010;

Sylvén & Sundqvist, 2012; Lindgren & Muñoz, 2013; Sundqvist & Sylvén, 2014; Jensen, 2016; De Wilde & Eyckmans, 2017). For example, Turgut and Irgin (2009) performed their study with 10-14 year olds who played online games in internet cafes in Mersin, Turkey and the results showed that the constant exposure to the same words enhanced the incidental L2 vocabulary learning and translation skills of the learners. Generally speaking, it has been found that a higher engagement in out-of-school activities enhances L2 proficiency whereas a low involvement results in modest gains (Muñoz et al., 2018; Peters et al., 2019; Muñoz & Cadierno, 2021).

According to Webb and Rodgers (2009a, 2009b), L2 learners should be encouraged to watch movies or TV series at home due to the scarce real input they receive at school. The same authors also assert that the presence of high frequency words, quite common in TV series, enhances incidental L2 learning. When they watch authentic videos regularly (extensive viewing), students are provided with the large amounts of multimodal input required to master the 2,000-3,000 WFs needed to be able to have 95% of the lexical coverage for movies or TV programmes (Webb, 2015). In order to boost L2 learning, teachers should try to encourage their students to watch TV in the L2 at home and, if necessary, provide them with tasks that require exposure to multimodal input (Lindgren & Muñoz, 2013).

Research has been conducted with young learners in order to find out whether and how their L2 proficiency is enhanced through OSE to English when performing several activities. There are a few studies that have focused on watching subtitled TV at home (TV series or movies), available on the internet or through the multiple TV channels. These studies have reported positive results for L2 learning (Kuppens, 2010; Sylvén

& Sundqvist, 2012; Lindgren & Muñoz, 2013; Sundqvist & Sylvén, 2014; Jensen, 2016; Muñoz et al., 2018; Muñoz & Cadierno, 2021). However, it should be noted that most of these studies have been performed in non-dubbing European countries such as Sweden (Sylvén & Sundqvist, 2012; Sundqvist & Sylvén, 2014), Denmark (Jensen, 2016) or Belgium (Kuppens, 2010), where large amounts of input are available.

Kuppens (2010) explored the impact of watching subtitled TV series and playing computer games on the incidental L2 learning of 374 primary students (11 years old) in Flanders, Belgium. The participants, who had not received any prior formal instruction in the L2, completed a questionnaire about the out-of-school activities they performed, as well as an English proficiency test. The results showed that watching subtitled TV programmes and movies had a strong significant effect on the translation test scores. In addition, the students who played computer games at home also showed significant but limited effects on the scores, probably due to the fact that the questionnaire did not differentiate between the different types of computer games.

Another study was conducted by Sylvén and Sundqvist (2012) examining the possible relationship between out-of-school activities and L2 English listening comprehension and reading comprehension as well as vocabulary learning in young learners (11-12 years old) in Sweden. It was found that playing digital games was the most common out-of-school activity performed by these participants. A more recent study by the same authors (Sundqvist & Sylvén, 2014) examined the L2 out-of-school activities of 10–11 year-olds in a medium-sized town from Sweden. The students answered a questionnaire and completed a one-week language diary to find out which out-of-school activities the children performed. Seven different options were provided:

reading books, reading newspapers/magazines, watching TV, watching films, using the internet, playing digital games, and listening to music plus an ‘other’ option. Apart from positively self-assessing their English ability, young learners enjoyed watching TV, listening to music and playing digital games. It was also observed that the amount of time spent performing these activities was twice the amount of instruction time spent at school, which indicates the large quantity of input through out-of-school activities these children are exposed to.

More recently, Jensen (2016) performed her study in Denmark and analysed L2 vocabulary knowledge in young learners (8 and 10 years old), who had received English instruction at school for a little longer than a year (two classes per week). They wrote a language diary for one week where they reported the frequency that they performed seven different activities: listening to music, reading books/magazines/webpages, speaking in English, writing, other, watching television/YouTube, Internet and gaming. The results, which indicated that watching TV was among the most popular out-of-school activities, suggested a relationship between watching TV and lexical proficiency, similarly to previous studies (Lindgren & Muñoz, 2013; Sundqvist & Sylvén, 2014).

There is, however, a study that did show more neutral effects for watching TV on language proficiency. De Wilde and Eyckmans (2017) measured incidental L2 learning by testing listening comprehension, reading comprehension, writing ability and speaking through the ‘Flyers’ exam, a Cambridge English Test. The participants were 11 year olds attending the last year of primary school in Ghent, Belgium. They had no prior English instruction at school and they were asked to answer a survey that

collected data on out-of-school activities such as speaking, computer use, gaming, reading, listening to music, watching L1 / L2 subtitled TV or without subtitles. The results indicated that watching subtitled TV did not have an influence on the scores, contrary to previous research (Kuppens, 2010; Lindgren & Muñoz, 2013). However, gaming activity was related to all the tests' results. In addition, computer use was significantly related to receptive VS, speaking ability and reading and writing skills.

Several studies have focused on the out-of-school activities performed by young learners and teenagers in Spain compared to other European countries. Lindgren and Muñoz (2013) conducted their research with students from Croatia, England, Italy, The Netherlands, Poland, Spain and Sweden. The young participants took reading and listening tests and their parents completed a questionnaire about the out-of-school activities performed. The outcomes showed that watching subtitled TV in the L2 explained most of the variance for listening and reading when compared to listening to music or playing computer games. The authors note the relevance of TV watching for improving linguistic activities. It should be noted that some of the countries used dubbing (Spain, Italy and England) so participants were not familiar with subtitling. Another study by Muñoz et al. (2018) investigated the differences in English grammar skills between young learners from two countries (Denmark and Spain) when performing out-of-school activities. Results revealed that the higher exposure the Danish students had to audiovisual programs in a FL made them obtain larger gains. In a more recent study conducted by Muñoz (2020) with teenagers from Spain, the activity of watching TV obtained the second highest correlation with the participants' English grades. The author suggests that these results are a product of the development of new technologies, the easier access to the language and the

globalisation of English. Finally, another study by Muñoz and Cadierno (2021) comparing the correlation between OSE and L2 learning with teenagers from Denmark and Spain found that the Danish students obtained significant differences in all language tests, except the Metalinguistic Knowledge test, probably due to the higher exposure to English outside school of Danish learners and the L2 learning context.

However, the trend is changing in countries such as Spain: the slow process of dubbing and the viewers' impatience to watch their favourite TV programmes forces them to view them in their original L1 (usually English) and with the aid of L1 or L2 subtitles (Muñoz, 2020). Even though some studies have recently been conducted in this area, further research is needed with young participants from countries where dubbing is used over subtitling. Furthermore, due to the growing interest in the field and the many variables that need to be explored, several skills and exposure to the L2 outside school should be considered (Turgut & Irgin, 2009; Kuppens, 2010; Aghlara & Hadidi, 2011; Sylvén & Sundqvist, 2012; Sundqvist & Sylven, 2014; Lindgren & Muñoz, 2013; Jensen, 2016; De Wilde & Eyckmans, 2017; Peters, 2018; Butler, 2019).

5.4. L2 proficiency, vocabulary size and reading speed

5.4.1. L2 proficiency level

Even though a lot of vocabulary can be learned by reading (Laufer, 2001; Webb, 2008), beginner young learners might not have the required reading skills to acquire new words when performing this activity (Coady, 1997; Nation, 1990). In addition, Vanderplank (2016) claims that low-proficiency learners, due to their low language

level and poor reading skills, might not benefit from multimodal input as much as adult learners.

Research has observed that proficiency level can be a decisive variable when choosing the best type of subtitles for L2 learners with different proficiency levels as we have pointed out (d'Ydewalle & Pavakanun, 1995; Danan, 2004; Vanderplank, 2010; 2016). So far, not many studies have been conducted with low-level learners (Montero-Pérez et al., 2013; Muñoz, 2017).

The use of multimodal input with low-proficiency learners has been discussed in several studies. Zanon (2007) and Martínez (2002) claim that authentic videos are too difficult for low-level learners, as they present real language at a normal speed, containing mixed structures and difficult vocabulary. However, even though it might be challenging for low-level learners to watch TV series or films, these students might enjoy being exposed to videos in the L2 due to their engaging characteristics such as sound effects and images (Lin & Siyanova, 2015). It has been argued that authentic input should not be modified, but audiovisual materials carry some important challenges for low-proficiency learners (Gilmore, 2007). Therefore, real unmodified input for beginners may be hard to find. The use of cartoons, which require lighter cognitive processing than other types of TV programme, can be suitable for children who begin to learn a FL (Bahrani & Soltani, 2011; Bahrani & Sim, 2012). This type of programme is claimed to facilitate the task for low-level students with a limited knowledge of the language. In addition, the inclusion of subtitles could be helpful for L2 comprehension (Rodgers & Webb, 2017) as even if their reading ability may not be

excellent, the combination with aural input will help in enhancing their L2 comprehension and at the same time will lessen their anxiety (Gowhary et al., 2015).

Other studies on the influence of multimodal input on L2 comprehension with participants at the intermediate (Huang & Eskey, 1999; Markham et al., 2001; Markham & Peter, 2003; Hayati & Mohmedi, 2011; Latifi et al., 2011; Başaran & Köse, 2012; Rodgers, 2013) or advanced proficiency level (Etemadi, 2012; Ebrahimi and Bazaee, 2016) have provided mostly positive results. Studies analysing the role of multimodal input on L2 word-form recognition with intermediate (Montero Pérez et al., 2014) and advanced L2 learners (Markham, 1999) have also obtained favourable outcomes. However, more studies on L2 comprehension at different proficiency levels with audiovisual materials need to be performed (Vandergrift, 2007).

Research has been conducted on the effects of multimodal input on low-proficiency learners, yielding positive results for L2 comprehension (Guillory, 1998; Baltova, 1999; Winke et al., 2010; Gowhary et al., 2015; Galimberti & Miralpeix, 2018) and vocabulary acquisition (Neuman & Koskinen, 1992; Koolstra & Beentjes, 1999; Hui, 2007; Sydorenko, 2010; Nagira, 2011; Lekkai, 2014). Nevertheless, other studies have concluded that the presence of subtitles is negative and did not enhance comprehension for beginner learners (Taylor, 2005; Lavaur & Bairstow, 2011). Even though it has been proved that low-proficiency learners can improve their L2 learning through multimodal input, more research in this area is needed considering the different results obtained (Matielo et al., 2015; Peters et al., 2016), probably due to the wide range of measures and materials used, which make the generalisability of the results extremely difficult (Montero Pérez et al., 2013).

Previous research has also suggested that beginner learners should start with L1S and move to L2S when their proficiency improves (Markham et al., 2001; Markham & Peter, 2003; Danan, 2004; Lin & Siyanova, 2015). The processing of L2S might be less demanding for advanced learners than for beginners and this is the reason why the former could obtain better scores than the latter in this subtitling condition (Van der Zee et al., 2017). Pujolà (2002) claimed that L2S functioned simply “as a backup to their listening activity” (p.254) for advanced learners whereas they were fundamental for the low-proficiency subjects to understand the content.

There are few studies comparing the effects of multimodal input on L2 learners with different proficiency levels for comprehension (Bianchi & Ciabattini, 2008; Lavaur & Bairstow, 2011; Gowhary et al., 2015) and vocabulary learning (Hui, 2007; Frumuselu et al., 2015; Peters et al., 2016) and varied results have been obtained regarding the best type of subtitles for beginner learners. Bianchi and Ciabattini (2008) conclude that L1S are better suited for low-level learners whereas L2S should be used with high-proficiency learners only. However, Gowhary et al., (2015) obtained significant differences for L2S on comprehension at all proficiency levels, although the advanced learners group obtained the highest scores with L2S. Hui (2007) conducted his study on L2 vocabulary learning with participants at high and low proficiency levels. The scores of the vocabulary test were significantly different, with the L2S group scoring higher than the others, followed by the L1S group and then the no subtitles group. Similar results were obtained by Frumuselu et al. (2015): the scores obtained by the L2S group were significantly higher than the L1S group and L2S were not a barrier for language learning when watching audiovisual materials for low-proficiency learners.

Another experiment was performed by Peters et al. (2016), who analysed the effects of multimodal input on L2 lexical acquisition with beginner, pre-intermediate and intermediate participants. The EFL learners from the first study performed similarly and no significant differences were found. On the other hand, the beginner students in the second study obtained lower gains due to their low proficiency level. The authors conclude that the authentic materials used for this study were too challenging and did not allow the low-proficiency learners to acquire new knowledge. Furthermore, they suggest, in line with Danan (2004) and Vanderplank (2010), that the audiovisual materials need to be selected carefully and should only be slightly above the proficiency level of the learners.

There is another study on the role of age and proficiency on subtitle reading performed by Muñoz (2017) that is in line with the results obtained by Peters et al. (2016). The participants (low and high-proficiency learners at different ages) watched two short videos from two episodes of the TV cartoon series 'The Simpsons' either with L1S or L2S. Results indicated that children (beginners), due to their lower L2 level compared to the adults, required more time to process L2S than L1S, even if both types of subtitles were attended to. The low L2 level of the students made it harder for them to read the L2S, whereas the higher level of the adult learners (some were already advanced learners) indicated that they did not pay much attention to L1S. Regarding proficiency, beginner learners were the ones that needed to focus more on both L1S and L2S, due to their low proficiency level, although L1S received less fixations. Intermediate learners resembled beginners and focused for longer on L2S than on L1S, although the numbers were lower compared to the beginners group. The results also indicated that L1S are more appropriate for children with low reading skills in the

L2, confirming previous research (Markham et al., 2001; Markham & Peter, 2003; Danan, 2004; Lin & Siyanova, 2015), whereas L2S might be more suitable for adolescents and adult learners with higher reading skills and proficiency levels (Vanderplank, 2010). However, further research on subtitle type in audiovisual materials is recommended.

To sum up, studies seem to favour L1S for beginners and L2S for more proficient learners (Borrás & Lafayette, 1994; Danan, 2004; Taylor, 2005; Bianchi & Ciabaton, 2008; Lin & Siyanova, 2015). On the other hand, research has also suggested that L2S might be more beneficial for low-proficiency learners (Neuman & Koskinen, 1992), and other experiments show that L2S can yield positive results to any proficiency levels (Bird & Williams, 2002; Winke et al., 2010; Frumuselu et al., 2015; Peters et al., 2016). The evidence provided for the most beneficial subtitle condition (L1 / L2) in low proficiency learners is still inconclusive, especially because there are very few longitudinal studies (d'Ydewalle & Van de Poel, 1999; Danan, 2015).

5.4.2. L2 vocabulary size (VS)

VS is the number of words that learners are able to recognise and understand in an L2. According to Meara (1996, p.3), “the basic dimension of lexical competence is vocabulary size”: learners with a large vocabulary knowledge are more proficient than learners with smaller vocabularies. It has been shown that vocabulary is a reliable indicator for language proficiency (Staer, 2009). In their study on the relationship between VS and EFL language skills, Miralpeix and Muñoz (2018) claim that “vocabulary is determinant at low proficient levels” (p.21). There are well-known tests that measure receptive and productive VS, for example, The Vocabulary Levels Test

(VLT) (Nation, 1983; Schmitt et al, 2001), the Eurocentres Vocabulary Size test (Meara & Jones, 1990) and the Vocabulary Size Test (VST) (Nation & Beglar, 2007). The VST by Nation and Beglar (2007) and the VLT (Nation, 1983; Schmitt, et al., 2001) are MC tests, whereas the Eurocentres Vocabulary Size test (Meara & Jones, 1990) is a yes / no test. The results provided by these VS tests provide important data concerning the development of the vocabulary language of the participants (Cameron, 2002).

The VST by Nation and Beglar (2007) measures the vocabulary knowledge reading proficiency levels of L1 and L2 learners. There are two parallel versions of the VST (A and B). The VST consists of MC questions in which learners are provided with the word form in an undetermined context and have to mark the right meaning from the four options available. The three distractors included in each question usually belong to word meanings of the same 1,000-word frequency level. This test was designed “to provide a reliable, accurate and comprehensive measure of the learner’s VS from the first 1,000 to the fourteenth 1,000-WFs of English” (Nation & Beglar, 2007, p.9), although beginner or intermediate learners do not need to take all the 140 questions, due to their low proficiency level.

The VS of the students is important to help determine the coverage they have of the input they are exposed to. Nation and Beglar (2007) argue that, when interpreting the scores of the VST, it is estimated that in order to get the required 98% of text coverage necessary to understand a text, 9,000 WFs are needed for novels, 8,000 for newspapers and 6,000 for children’s movies (Nation & Beglar, 2007; Nation, 2012). Research has been conducted with different VS tests trying to establish the VS necessary for text comprehension (Hirsch & Nation, 1992; Cobb & Horst, 1999; Nation, 2006; Jiménez Catalán & Terrazas Gallego, 2008). Hirsch and Nation (1992) analysed

the VS of teenage English natives who read three short novels written for their age. The results of the VS test indicated that, in order to be able to fully understand the written texts, the participants needed to be at the 5,000-word level.

Nation (2006) analysed the VS needed for reading authentic materials such as novels, newspapers, watching a children's movie and understanding spoken English. The amount of VS needed for reading differed depending on the text genre. In order to comprehend 98% of authentic novels and newspapers, 8,000 words were needed, whereas only 4,000 were required to obtain a 95% coverage. On the other hand, the graded reader analysed in the study, which belonged to level 3 of the Oxford Bookworm Series, required the learners to be at the 3,000-word level. Nation also focused on audiovisual materials and claimed that in order to understand the animated film *Shrek*, 4,000 words were needed to understand a 96.74% of the text, whereas to comprehend the 87.91%, the learners only needed to be at the 2,000-word level. The reoccurrence of some low-frequency words in the film might have helped to lower the lexical demands, but the presence of spoken language and the lack of chances to go back to the input did not facilitate language acquisition.

Jiménez Catalán and Terrazas Gallego (2008) examined the VS of 270 young learners (taking into account that previous research had mainly focused on adult learners). The participants were 10-year-old students attending Grade 4 at a primary school in Spain. They took different tests, including the VLT 1,000 and 2,000-word level frequency band, which is adequate for beginner learners (Schmitt, 2000). The results indicated that the VS of the young learners was within the 1,000-word frequency level, although half of the students identified less than two-thirds of the words. Regarding the 2,000-word frequency band, the scores reported even worse results and only a few words

were recognised by the children. However, the authors claim that the results are not surprising due to the young age of the students and the little formal instruction in the L2 they had received. When analysing the results, a relationship between VS and L2 proficiency was found, although further research with children in this area is recommended.

Even though language learning is enhanced by reading, children and adults spend many hours a day watching TV and this practice may be useful for L2 lexical acquisition (Webb & Rodgers, 2009a). Most research on lexical coverage has been performed in relation to reading (Hirsch & Nation, 1992; Cobb & Horst, 1999; Nation, 2006; Jiménez Catalán & Terrazas Gallego, 2008), and some recent research has concentrated on multimodal input. TV programmes offer a great source of input for L2 learners but these students need a minimum VS in order to understand and learn from them. This can be challenging for low-proficiency learners (Webb & Rodgers, 2009b). Several studies have analysed L2 vocabulary learning through multimodal input, but the studies controlling the VS of L2 learners are scarce and diverse results have been found: Montero Pérez et al. (2014) and Peters et al. (2016) obtained a significant correlation between the VS of learners and the vocabulary learned, whereas Rodgers (2013) did not find any.

In a study focusing on the VS needed to understand TV programmes, Webb and Rodgers (2009b, p.420) claim that in order to obtain a 95% of coverage, “the minimum vocabulary size” is 3,000 to 4,000-WFs, which might be an attainable goal in an instructional setting. On the other hand, 7,000-WFs are required to obtain a 98% of coverage, which might be too challenging and improbable to get. The study mentions,

though, that language learning is also possible with a 90% coverage. They also analysed the number of WFs needed in order to understand the audiovisual input provided with a 95% coverage for different TV genres and concluded that, due to the potential benefits audiovisual materials offer for vocabulary learning, the 3,000-word level should be enough to comprehend TV programmes (in line with a later study by Rodgers and Webb, 2011). However, this figure decreases to 2,000-WFs for children's programmes, which is in line with Nation (2006).

Webb and Rodgers (2009b) also argued in their study that a less demanding text coverage could be explained through narrow reading (i.e., when students read a set of materials that is similar, such as a series of books on the same topic or by the same author), which has proved to be an effective way of learning new words. Narrow reading implies that the number of infrequent words can decrease, as the same words tend to appear recurrently (Hwang & Nation, 1989). The concept of narrow reading can also be applied to TV programmes when watching related episodes of the same TV series (Webb & Rodgers, 2009a), which can imply a lower VS demand.

Nation (2006) argued that the VS needed to understand an animated film was lower than reading different types of texts. As has been previously stated, cartoons have distinguishable qualities, such as the attractiveness of the images provided in the visual input and a slower pace than other TV genres. Cartoons require less cognitive processing than other TV genres, which aids understanding (Bravo, 2008; Bahrani & Sim, 2012). This type of audiovisual material may not require a big VS for low-proficiency learners to understand and they could "play a facilitative role" for L2 learning (Peters et al., 2016, p.135). They can also be a bit challenging for them and

present some unknown words, which would fulfil the criteria for Krashen's (1985) Input Hypothesis (i+1).

5.4.3. Reading Speed (RS)

RS can be defined as “the rate at which something is read, often expressed in terms of words per minute” (Collins English Dictionary). There has been a quest for the ‘ideal’ reading rate in several studies, although there is no agreement on the speed required for optimal reading in the L1. Higgins and Wallace (1989) establish a minimum of 180 words whereas Jensen (1986) and Nuttall (1996) argue that 300 words per minute (WPM) is the reading rate that a L1 native speaker should have in order to read properly. When focusing on young learners, Shroyer and Birch (1980) conclude that the RS for 8-13 year olds is 116 WPM whereas Karamitroglou (1998) established the RS for learners aged 6-14 years old at 90-120 WPM. However, when these young learners need to read in their L2, a lower RS can be expected from them. Whitford and Joannis (2018) conducted a study where they analysed the online reading performance of monolingual and bilingual children (L1 vs. L2). The results showed a slower RS in the L2 for bilingual children, which increased their reading time. They argue that the bilingual young learners showed a “slower processing of lower-frequency L2 words” (p.333) due to a lower exposure to these L2 words, whereas the outcomes for higher-frequency words showed similarities in the levels of exposure, which implied that “they were comparably processed in both languages” (p.334).

Roberts and Felser (2011) claim that reading in the L2 can be a slow process for low-proficiency learners who do not have a great command of the language and tend to read in a slower and less automatic way. In addition, Muñoz (2017) argues that

children, due to their young age and low proficiency level, will face more challenges when reading in a FL because of their insufficient vocabulary knowledge. Subtitles may be too hard to read for low-proficiency learners due to a demanding RS, and might be inadequate for them (Zárate, 2008). However, various experiments performed with subtitled audiovisual materials have argued that reading L1S or L2S is an automatic behavior (d'Ydewalle & Gielen, 1992; Pavakanun & d'Ydewalle, 1992) and it has been proved that when young learners are presented with multimodal input, the presence of images does not distract them from reading the text (d'Ydewalle & Vanrensbergen, 1989; d'Ydewalle & Bruycker, 2007; Tragant & Pellicer-Sánchez, 2019). Vanderplank (2013) argues that RS is a key variable that correlates with proficiency and he suggests that low-proficiency learners might lack the necessary reading abilities to read the subtitles (especially L2S) included in TV programmes or films, which will limit their improvement in L2 comprehension.

The importance of reading and understanding the written input is also stressed by Tamayo (2016), who claims that “the time a subtitle remains on screen should be enough to allow [...] an adequate reading pace [...] and the cognitive processing of the information to understand it” (p.276). Marzá and Torralba (2015) claim that young and less-skilled learners will not be able to follow the regular six-second rule of subtitles (i.e., they only appear for six seconds on the screen) and they will only be able to attend to them adequately once they have started Grade 5 (10-11 years old) in primary school. However, there is not a definite conclusion on the exact RS required for children's programmes, especially in the L2, where more difficulties will be encountered by FL learners. Sometimes subtitled programmes, especially those designed for children, will be modified and an eight-second rule will be applied (i.e.,

they will appear on the screen for eight seconds) in order to facilitate reading in the FL (Koolstra et al., 2002).

Koolstra et al. (1999) investigated how children in Grades 2, 4 and 6 of primary school gained from lengthening the presentation time of the subtitles, which were shown following a six, eight and ten-second rule. The young learners watched 9-minute videos, with L1S from two American TV series: *Full House* and *Knightrider*. The scores indicated that the older children (Grades 4 and 6) used the same amount of time to read the L1S. However, Grade 2 learners did not pay much attention to the L1S in the six and eight-second rule, due to their speed. The outcomes indicated that these second graders spent more time reading the L1S when they were presented with the ten-second rule, implying that L1S in the six-second rule were not followed by the learners due to a lack of time. Furthermore, the results showed significant differences between the L1S on the ten-second rule and the other two (six and eight-second rule). The authors suggest that extending the display time of L1S is a good option to aid language learning in children and conclude that the reading ability of the learners can alter the results.

However, there is no standard RS established for programmes addressed to children, although these tend to present a lower reading rate (Zárate, 2010). A RS lower than 70 - 80 WPM has been recommended for these type of programmes, even though Zárate claims that high RSs are needed in order to be able to read the closed captions that appear on the screen. She argues that subtitles should be presented at a rate that could be read and understood by all audiences and she encourages further research on the inclusion of slower subtitles for children. She proposes the use of edited

subtitles in order to reduce the reading pace, which would provide young learners with more time to devote to the images. She also recommends consideration of the user's experience when reading subtitles, especially in the case of young audiences. Due to the inclusion of technology in our daily lives, children might be regular users of subtitles through watching TV or playing videogames at home, and these habits should be taken into account in further research exposing learners to multimodal input (Zárate, 2010).

On the other hand, RS affects processing, either in the L1 or the L2 (Kaan et al., 2015) and research has focused on comparing the effect of this variable on L1 and L2 learners and on the relationship between RS and proficiency with inconclusive results (Roberts & Felser, 2011; Kaan et al., 2015). In Roberts and Felser (2011), 24 L1 English native speakers and 24 Greek L2 English learners had to read ten syntactically ambiguous sentences (10 words each) with “optionally transitive verbs followed by a finite complement clause” (p. 306). The results indicated that, as had been expected, RS was related to proficiency as the native speakers read faster and obtained more favourable results than the L2 readers. Moreover, both groups obtained high comprehension accuracy scores. The authors encourage further research that takes learners' IDs into account when analysing RS and language processing.

The study performed by Kaan et al. (2015) obtained different results from Roberts and Felser (2011). They examined the effects of RS on L2 sentence processing. The participants, 39 L1 English speakers and 71 advanced L2 English learners, had to read and complete five paragraphs (20 words each) where the lexical items were truncated after the first few letters. The students also performed a naming task and a

paper-and-pencil repair task where they had to correct grammatical errors in 14 different sentences. Surprisingly, the results showed similar online reading patterns for L1 and L2 participants, although the L2 learners obtained better reading scores than the native speakers. The authors also claim that differences in RS did not correlate with differences in English proficiency.

The studies mentioned above have measured RS using computer tests. However, another more traditional method to measure RS is used in many schools across Catalonia (Spain), where this variable is now being controlled at different stages of primary education. Teachers perform a very simple test where the learners are asked to read aloud a pre-selected text for one minute on a one-to-one basis (the instructor tells the students when to start and stop reading). This variable is examined in order to check for any issues that may arise in the reading patterns of young learners while they are developing their reading skills. Online tests, however, can easily be performed online with instant feedback provided, facilitating the teachers' job. Most of the tests available measure RS in English (e.g., Free reading test: <http://www.freereadingtest.com>, Reading soft: <http://www.readingsoft.com/>, My read Speed: <http://www.myreadspeed.com/>). Tests in Spanish can also be found online (Lectura àgil: <https://vip.lecturaagil.com/test-lectura-rapida/>, Lapicero màgico: <http://lapiceromagico.blogspot.com.es/>). In these tests, several comprehension questions are also asked at the end in order to make sure the student understands what s/he is reading. That is why the term “reading efficiency” is used as well to refer to RS (i.e. comprehension is taken into account).

In sum, watching TV programmes in a FL can be challenging especially for low-proficiency learners, due to the subtitle time on screen, students' RS and the unknown words that appear in the videos (Sydorenko, 2010; Webb, 2011). It has been argued that the ability of the learners to understand the subtitles will vary depending on their reading ability (Marzà & Torralba, 2015), which is an important issue for L2 learners (Vanderplank, 2016) and may not be fully developed if learners are still young. Further research taking into account the RS and the receptive VS of the children who watch a subtitled TV programme needs to be carried out. In addition, special care should be taken when choosing materials with an adequate coverage and when choosing the type of subtitles for optimum language learning.

CHAPTER 6 – RESEARCH QUESTIONS AND METHOD

This chapter presents the RQs as well as the methodology of the study. It starts by putting forward the four RQs this thesis will answer (6.1). They have arisen after careful reflection on the previous research presented and the need to explore how we can help young learners acquiring English through multimodal input. Next, in the methodology there is a description of the participants (6.2.1) followed by the instruments (6.2.2), which include the TV series selected for this research and the questionnaires and tests administered to the learners before and during the treatment. The chapter continues with a report of the procedure followed before, during and after the treatment (6.2.3), providing all the necessary information about the pedagogical intervention. Finally, the data analysis performed to answer the four RQs is presented (6.2.4).

6.1. Research Questions

The RQs of this thesis are grounded on what we have presented in the literature review sections, so they focus on: L2 comprehension and word-form recognition and the possible influence of other variables such as L2 aptitude and L2 proficiency, RS and OSE in extensive cartoon viewing by young learners. The fact that the study is longitudinal also allows for the exploration of whether extensive viewing changes the way in which children watch TV episodes throughout the treatment. Therefore, the present dissertation aims to answer four RQs:

1) Is there any difference between extensive viewing of animated TV series subtitled either in the L1 or in the L2...

1a. in episode comprehension?

1b. in written recognition of known and new vocabulary?

2) Are episode comprehension and written vocabulary recognition scores related to linguistic aptitude and proficiency variables (i.e., L1/L2 reading speed, L2 vocabulary size and school English marks)?

3) Does extensive viewing over a period of five months influence the way in which L2 learners watch subtitled TV series?

4) Does out-of-school-exposure (OSE) to multimodal input have an effect on episode comprehension and written vocabulary recognition scores when watching subtitled TV series in class?

6.2. Methodology

6.2.1. Participants

The participants of this study ($n= 92$) were Spanish/Catalan bilingual young learners of English. They were all 11 and 12 year-olds ($M= 11.55$), and were in Grades 5 and 6 (second cycle of primary education) at a semi-private primary school in Catalonia (Spain). 75% of the bilingual students usually used Spanish in their everyday lives and 25% used Catalan (this information was obtained from the OSE questionnaire that participants answered before the treatment, see 'Instruments' section).

The students, originally distributed into four school classroom groups, two in Grade 5 (lines A and B) and two in Grade 6 (lines A and B), were assembled into two groups for the purpose of this study [L1S ($n=47$) and L2S ($n=45$)]. Each group consisted of one Grade 5 class and one Grade 6 class. They watched the TV series *Curious George* either with Spanish subtitles (L1S) or English subtitles (L2S). No differences in English proficiency level and receptive VS were observed between groups and that is why they could be selected for the purpose of this study. The results of the ANOVA conducted on these scores show no significant differences between them, as reported in full in the results section.

All the participants had been receiving English lessons since pre-primary school (three years old) and had been exposed to English for three hours a week since then. They were instructed through the 'Artigal' methodology, which mainly focuses on oral communication and repetition of words and phrases from the stories the learners were

told. At the beginning of primary school, in Grades 1 and 2, the students attended Science classes (1.5 hours/week) and Arts and Crafts (1 hour and 45 minutes/week) in English, in addition to the 2 hours and 15 minutes of English every week. From Grade 4 and until the end of primary school, all the participants received three hours of English per week, plus 2.5 hours more of CLIL instruction in English (1.5 hours of Science and 1 hour of Arts and Crafts). This makes a total of 735 hours of English from Grade 1 to Grade 4, 892.5 from Grade 1 to Grade 5 and 1050 from Grade 1 to Grade 6. The English teacher of all the participants is the researcher who conducted the study. She has had considerable experience teaching these students since the beginning of primary school and has been their only teacher since then.

Initially, the two intact groups had 50 (L1S group) and 51 (L2S group) students respectively, although the final sample was made up of 92 students (47 in the L1S group and 45 in the L2S group). The final sample was also quite balanced in terms of gender (see Table 6.1).

There were two main reasons for excluding some of the students (three in the L1S group and six in the L2S group) from intact groups in the final sample of participants: (1) two were excluded because they were 'special education' learners with severe reading problems, (2) the rest (seven) were excluded because they were not always in class and had missed more than one episode of *Curious George* during the treatment. Furthermore, the sessions they had missed were contiguous or close to one another (e.g., a participant who had not been in class for two weeks). The students that missed two episodes from one part of the treatment (episodes between 1 - 10 or between 11 - 20) were definitely not included in the final sample for the study, as it

was desirable that all participants had had comparable amounts of exposure to the TV series throughout the treatment weeks.

There were other participants who missed one or two sessions, but they were included in the study: eighteen participants were absent in just one of the viewing sessions (e.g., it may be the case that somebody missed one day in a 20-week pedagogical intervention), but this fact was not thought to severely affect their performance and they were kept in the study. There were also five students that missed two viewing sessions (one from the first half – episodes 1 to 10 - and one from the other half – episodes 11 to 20) but they were not excluded because they were following the treatment regularly over the weeks. The final number of participants in the study is presented in Table 6.1.

Table 6.1 – Participants in the study.

Group	Males (n)	Females (n)	Total (n)
L1S (Spanish subtitles)	24	23	47
L2S (English subtitles)	20	25	45

6.2.2. Instruments

6.2.2.1. *Curious George* TV series

Curious George is a North-American animated TV series produced by Universal 1440 Entertainment (i.e., 'Universal Studios Family Productions' before 2013) and Imagine Entertainment, WGBH Boston. It is animated by Toon City. All the episodes (10 minutes long each) are closed captioned by the Caption Center at WGBH for deaf or

hard-of-hearing viewers. The show debuted in 2006 and ended in 2015 after its ninth season, although PBS Kids is still broadcasting daily re-runs in the USA, where it received renowned awards such as the Daytime Emmy (for animated series). Research has been conducted on the *Curious George* TV series in the USA in order to evaluate its educational impact, obtaining very favourable outcomes. The study evaluating the viewing of the *Curious George* episodes by the Concord Evaluation Group (2012) showed that the young learners' knowledge of Science and Maths significantly improved after watching the videos compared to the groups that did not view them. They claimed that "watching *Curious George* episodes resulted in a knowledge boost for the children in the television Group." (p. 16). The TV series was also shown to be positively evaluated by parents in that study. They also recommend it to other parents who want to "encourage their children to explore their surroundings" (p. 27). In spite of the success in the US, this cartoon is not very popular in Spain. At the time the study was conducted, it was not being broadcast on any TV channel. The participants were asked if they knew about the TV series after watching an extract of the episode selected for the video-watching questionnaire (see below), and they confirmed that they had not seen any previous episode and claimed not to know anything about *Curious George*.

The series is about a curious monkey who has different adventures and is aimed at young viewers. It is based on a children's book series of the same name. The series wants to motivate children to explore key concepts of various disciplines and help them develop their Science, Engineering and Maths literacy. All the episodes focus on George, a curious little monkey who is the protagonist of the show: he has the brain and the imagination of a child and his curiosity causes unexpected problems. In each

episode, George explores, observes and discovers new things in a funny way. All his adventures lead to new discoveries and unexpected endings. He solves any trouble he gets into by himself or with the help of 'The Man with the Yellow Hat', the man George lives with. He acts as his mentor/tutor/father and takes care of him. As 'The Man with the Yellow Hat' was a very long name to use in the tests of the present study, his name was changed to 'Ted'. It was found out that the name 'Ted' was used in the *Curious George* film to refer to 'The Man with the Yellow Hat', too. That is why this particular name was adopted in this study. The episodes of this TV series have two main settings: the city and the countryside, which allow George to experience adventures in both urban and rural environments. In the city (the audience is never told its name) George and Ted live in a spacious, high-income apartment building near Endless Park, a zoo and a museum. When they are on holiday in the countryside, they stay at a small country house located near Lake Wanasinklake.

The complete first season of *Curious George* consisted of sixty episodes (also of approximately ten minutes each) that were broadcast in pairs, although they were plot independent. They aired on American TV for the first time from September 2006 until February 2007. For the treatment, twenty episodes from season one were used. In the episodes selected, George faces many different adventures such as making a home for his pigeon friend, playing mini-golf, working as a doorman in the building where he lives and learning about the tadpole's life cycle.

This particular series was chosen because it is entertaining and aimed at children, but with a richer dialogue compared to other cartoons, which usually include characters who just make signs or produce only sounds or one-word utterances. Furthermore, in

Curious George, its characters have conversations that cannot be understood by only watching the images. The students watching the videos also needed to understand the text (i.e., what the characters were saying) to understand what was going on. The two main characters (George and Ted) appear in each episode and their recurring presence helps the students follow the storyline, as has been pointed out in the previous literature on TV series. Episode length was also considered adequate for the level and age of the participants (longer episodes in a language that they have a limited proficiency in may be boring or tiring and make them lose interest in the activity). Therefore, although there are many excellent TV series aimed at children with longer episodes, these were discarded: young learners watching cartoons in a FL need to be focused on the activity and we considered that episodes should not last more than 15 minutes for them to be attentive.

The 20 *Curious George* episodes selected for this study had an average running time of 12 minutes and 26 seconds (the shortest was 12:15 and the longest 12:49) including approximately 48 seconds of the TV series theme song at the beginning and 50 seconds of closing credits each (see Appendix H for the running time for each episode). Furthermore, the 20 episodes were selected following specific criteria: the researcher watched all the episodes from the first season and wrote down the ones with unknown vocabulary for the participants. The episodes that were considered too challenging for the participants, mainly due to the vocabulary demands or sometimes because of the content, were discarded, as it was important that the episodes selected were entertaining for learners to engage in the activity.

The lexical demands of the episodes were carefully examined. The tool used to conduct the analysis of the lexical coverage from each episode was VocabProfile v.2 (Cobb, ongoing). In addition, Laufer and Nation's original four-way frequency sorter for texts (Laufer & Nation, 1995) based on the General Service List and Academic Word List (GSL-1k, GSL-2k, AWL, and Offlist) was used to assess word difficulty and classify words depending on whether they belonged to the most common 1,000 words (1k), between 1,000-2,000 (2k), or whether they were academic words (AWL). Those not belonging to any of these lists were classified as 'offlist'.

The transcripts of the episodes were downloaded from two different websites *Beano Wikia* (<http://beano.wikia.com>) or *TV Ark* (<http://tv.ark.com>), which offered access to the transcripts from the TV series for free (see Appendix I for the transcripts of episodes 9 and 19 [hereafter E9 and E19]). The transcripts were pruned in order to be analysed with the computer tools mentioned above. All the contractions that appeared in the transcripts were turned into two separate words (e.g., 'I'm' → 'I am', 'You'll' → 'You will') to facilitate its categorisation. Finally, onomatopoeias (e.g., 'wham', 'phew'), interjections and marginal words (e.g., 'ah', 'oh', 'hmm', 'huh') were deleted due to their similarity in the participants' L1 and because they do not convey much meaning.

After the transcripts from the twenty episodes were pruned, they were introduced into the VocabProfile processor in order to analyse their lexical coverage and to find out if the *Curious George* episodes were appropriate for the participants. Research has proved that a 95% lexical coverage is needed in order to understand a written text (Laufer, 1989). However, when watching audiovisual programmes such as the TV series of the current study, the VSs needed for the learners to understand the input

are lower, due to the support offered by the images and the written text in the form of subtitles, as pointed out in the previous chapters. It has been argued that in order to understand children's programmes, viewers need to be at the 2,000-word (2k) level (Nation, 2006). However, Webb and Rodgers (2009a) suggest that a less demanding lexical coverage for cartoons is required for narrow viewing, implying a lower VS demand for successful comprehension.

In this study the viewers watched several episodes of the *Curious George* TV series, which include many repeated encounters of recurring characters and similar situations. Due to that, this type of audiovisual material does not require big VSs for learners to understand what is happening and have enough opportunities to learn new words. The lexical coverage of the twenty episodes selected ranges from 90.2% to 97.1% and it is appropriate for the VS accounted for the participants in the study, who achieved an average VS score of around 12 in the VST (see results chapter). This coverage level was reached in many episodes (E1, E5, E6, E7, E9, E10, E13, E14, E15, E17 and E20) and the rest of the episodes show a coverage over 90%. The lexical coverage of each episode included is listed in Table 8.1 (see Appendix H).

6.2.2.2. Vocabulary size test

All the participants were tested using the *Vocabulary Size Test, version A* (VST) (Nation & Beglar, 2007). This test measures the receptive VS in English up to 10,000 words and is divided into ten frequency bands. The L2 learners in this study only took the first three bands (there are ten words per band) due to their low proficiency level. The test uses a MC format and presents each word to be assessed in a short non-

defining context (Figure 6.1 shows an example of a VST item). Learners have to choose the right meaning from one of the four options provided (one is the correct answer while the three distractors share some semantic features with the TW). The distractors from the test usually belong to word meanings from the same frequency band of the TW. Each question was awarded one point and the wrong answers were not penalised. The maximum score that could be obtained was 30 and the minimum was 0. The following MC question corresponds to item number one in the test:

1. see: They <saw it>.
 - a. closed it tightly
 - b. waited for it
 - c. looked at it
 - d. started it up

Figure 6.1 – Example of a VST Item (from Nation & Beglar, 2007)

This test had already been used for research purposes in primary school classes in a similar school context in Barcelona by Fusté (2013) and Casulleras (2014), obtaining comparable results in the same grades to those found in other studies in similar contexts (e.g., EFL learners in Spain) with other receptive VSTs like the *Vocabulary Levels Test* (VLT) (Nation, 1990); see for instance Agustín Llach and Terrazas (2009) or Jiménez Catalán and Ruiz de Zarobe (2009). The test can be found in Appendix A.

6.2.2.3. Reading speed tests

Two different RS tests were used in order to measure the RS of the participants in Spanish (L1) and English (L2). Both were online tests and had the same format, also

asking four MC comprehension questions at the end so as to check that participants were understanding what they were reading.

6.2.2.3.1. Spanish reading speed test

The Spanish RS test included in *Lectura Ágil* (<https://lecturaagil.com/>) was used to measure the RS of the participants in this study. *Lectura Ágil* is the largest online platform across Spain and South America that provides tools to measure and help improving comprehension, memory and speed while reading. In order to start working on these aspects, the website offers a free online Spanish RS test. *Lectura Ágil*, which is settled in Barcelona (Spain), aids people from all over the world in reading, understanding and memorizing input faster. Resources from *Lectura Ágil* are recommended by the Spanish Ministry and have been advertised on the main Spanish radio and TV channels.

The online free RS test from *Lectura Ágil* includes an excerpt from *The Little Prince*, which contains 414 words, and can be read and understood by young learners. This test can be found and also performed online for free (the only action requested is to introduce a valid email address) at <https://vip.lecturaagil.com/test-lectura-rapida/>. It was selected over other online RS tests because the latter included more challenging texts with difficult vocabulary that might have been too complex for the students participating in this study. Learners are asked to read a text and press a stop button when they finish. Then, participants have to answer four MC comprehension questions appearing on the screen (one at a time). After answering the last question, three

scores appear on the screen: the RS per minute, the amount of time used to read the whole text in seconds and the total comprehension scores.



Figure 6.2 – Screenshot from the results of the reading speed Spanish test.

The total comprehension score of the test is shown in percentages (Figure 6.2) and each correct question is awarded the same amount (25%). The test takers can obtain four possible scores: 25%, 50%, 75% and 100%, which is the maximum, and learners are not penalised for the wrong answers. For the purpose of this study, two scores were kept: RS (WPM) and the comprehension percentage of each participant (see Appendix B for the Spanish RS test).

6.2.2.3.2. English reading speed test

An online test was used to assess the English RS of the L2S group. The RS test is available from *Free Reader Digest* (<http://www.freereadingtest.com>), which is a free service from *Ace Reader*, some award-winning North American software created to provide educational programmes to schools that would improve skills such as RS,

comprehension and fluency. The Free Reading Test website (<http://www.freereadingtest.com>) offers free RS and comprehension tests through themed and leveled texts. For this study, a text including verbs in the present simple tense was selected due to the low proficiency level of the students and their lack of familiarity with the other verb tenses. The chosen text was on *Fun Facts* (five possible themes are listed on the test website). Level 1 was selected (thirteen levels are available) and Story 1 was chosen from the three options (see also Appendix B for the English RS test).

The assessment is very similar to the RS test in Spanish. It consists of a very simple text (which has been selected taking into account the low proficiency level of the students in English) that learners had to read. After reading, the student clicks the 'done reading' button and four comprehension questions appear on the screen. The test taker has to mark one answer from each MC question and then press the 'done' button. These questions supplied reliable data on the good performance of the test as its answers corroborated if the learners had read and understood the content of the excerpt.

Immediately after that, the learners press the button 'view results' and the number of words read per minute, as well as the percentage obtained in the comprehension test (25% per question and a maximum of 100%) are shown on the screen (Figure 6.3). Students are also able to see the feedback for their responses (see Appendix B).

Free Reading Speed Test

Test your reading speed and comprehension

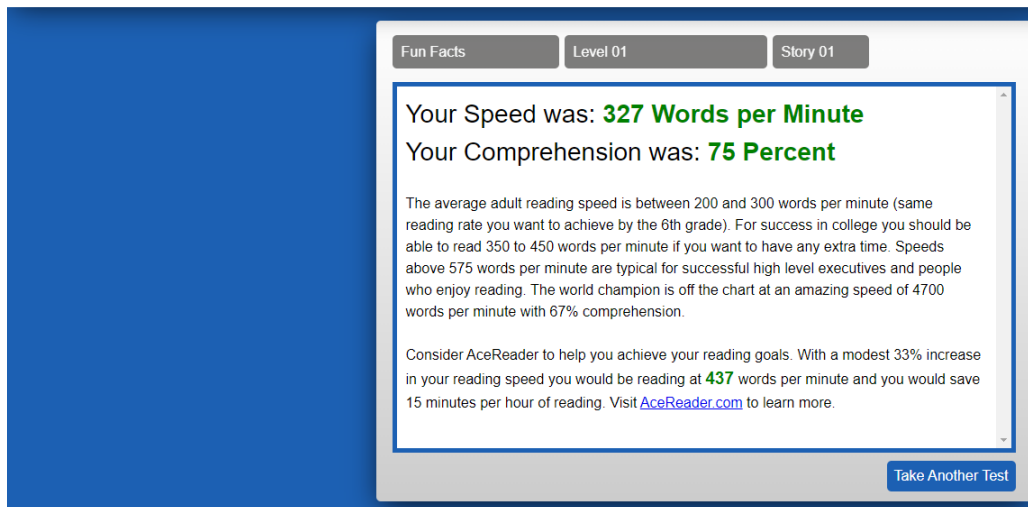


Figure 6.3 – Screenshot from the results of the English reading speed test.

6.2.2.4. Aptitude test: LLAMA_B

The LLAMA tests (Meara, 2005) are a set of four language-neutral sub-tests created to measure language aptitude for FL learning. They are based on components from the standardised MLAT tests by Carroll and Sapon (1959), but are computer-based. They test language aptitude by measuring vocabulary learning (LLAMA_B), phonetic (implicit) memory (LLAMA_D), sound-symbol correspondence (LLAMA_E) and grammatical inferencing (LLAMA_F). They have been widely used in SLA research (Rogers et al., 2017) and have proved to work reliably in different studies (Grañena 2013a; 2013b; Rogers et al., 2017). Bokander and Bylund (2020) also showed that the test had good internal validity for research purposes. This test can be downloaded from Lognostics (<http://www.lognostics.co.uk/tools/llama/>) (accessed on 02-11-2018).

Llama_B measures rote learning and associative memory for vocabulary learning. It is quite similar to the vocabulary learning section of the MLAT, but with a more modern

interface. Based on picture stimuli, it is a simple computer-based task made up of two phases. It starts with a learning phase where the students see 20 pictures of unusual objects displayed together on the screen and they are given two minutes to click on the pictures to see their names (see Figure 6.4). The learners taking this test are required to learn the name of each drawing. In the next step (the testing phase) all the pictures remain on the screen whereas the names of the items appear one by one. The test taker needs to click on the picture that they think matches with the name that is shown. After that, another name is disclosed until the 20 names have been assigned to one picture. There is no time limit and five points are awarded for each correct answer, with a maximum score of 100. Students are not penalised for giving the wrong answer and the final result (in percentages) appears and is saved at the end. This test takes approximately 5 - 10 minutes to complete.



Figure 6.4 – LLAMA_B screenshot.

6.2.2.5. Out-of-school exposure (OSE) questionnaire

The OSE questionnaire (in Catalan), created by the teacher-researcher of the study, consisted of two parts (see Appendix C). The first part contained five questions in total. Questions 1 and 2 (hereafter Q1 and Q2) elicited some biodata and necessary linguistic information for the study (e.g., languages spoken at home and with friends). The rest of the questions (3, 4 and 5) inquired about participants' exposure to the English language. In Q3, they were asked if they received extra lessons after school, while Q4 was related to possible stays abroad. Q5 presented three grids where the students had to indicate whether they did several activities in English, including listening to music, reading, playing videogames or watching TV series or films, which was relevant for the purpose of the study. The grids on playing videogames and watching TV series or films in English also asked about the use of subtitles. In affirmative responses, students also had to indicate how often they carried out these activities. Those who watched TV series or films in English were asked to fill in the second part of the questionnaire with the help of the teacher, i.e. the answers provided in Q5 of the first part determined whether participants continued answering the second part with the individual help of the researcher.

The second part contained open and closed questions. Q6 inquired about the specific titles of films and TV series the participants watched, who they watched them with (if it was the case) and if they enjoyed performing this activity. Q7 was divided into three parts. In Q7a, the students were asked to think about how they watched the audiovisual materials with subtitles: they were given different options to choose from, which included listening to the soundtrack and reading the subtitles at the same time,

listening to the soundtrack and sometimes reading the subtitles, only listening to the soundtrack or only reading the subtitles without paying attention to the soundtrack. In Q7b, the students were asked if they thought the subtitles appeared too fast on the screen while in Q7c they were asked about possible reasons they had for reading the subtitles. For this last question (7c), they were also given some space to provide further information if their answer was not any of the ones included.

6.2.2.6. Video-watching questionnaire

The video-watching questionnaire was based on a two-minute excerpt from an episode of the animated TV series selected for the study. It was designed by the teacher-researcher and it aimed to know more about the students' experiences when watching the video excerpt. For example, it asked whether participants understood the video, had read the subtitles and learnt new words from it.

In order to answer the questions, the students were asked to watch first a short (2-minute) video from an episode not included in the treatment. The video excerpt was the same for both groups and it was shown with L1 or L2 subtitles depending on the group the students had been assigned to.

The title of this short video is *Charkie Escapes* and it starts with George and Ted playing football in the park. Suddenly, Ted reminds George he has to go and take care of Charkie (Aunt Margaret's dog). George meets Aunt Margaret and takes Charkie, but immediately after she leaves, Charkie frees himself from his lead and escapes.

The first part of the questionnaire (front page) included closed questions in Catalan, (Likert scales from 0 to 5 and MC questions). Question 1 asked the students if they had understood the video and what had happened. In addition, the subsequent questions enquired about the viewing and reading behavior of the students while watching this excerpt, for example, describing different ways of viewing the video (e.g., with/without reading the subtitles). Question 3 and subsequent questions enquired more specifically about subtitles, so only those participants saying they had read them answered the questions in this section.

The second part (on the other side of the same sheet of paper) included five comprehension questions in Spanish (the language of L1S) about the excerpt. Therefore, the minimum comprehension score was zero and the maximum 5. The comprehension questions were open and each correct answer was awarded one point. It also included a word recognition exercise with twelve words: seven were taken from the subtitles, while the other five were distractors. These distractors did not appear in the video but were semantically related to the main plot of the video segment the participants had just watched. Participants had to circle those appearing in the video. One point was awarded for each correct answer (the minimum score was zero and the maximum 7). The video-watching questionnaire can be found in Appendix D.

6.2.2.7. Episode-based comprehension tests

In this study, content comprehension was measured following Rodgers (2013) and the items proposed in each of the comprehension tests implied “understanding the topics contained in relatively lengthy viewing sections and in the episodes as a whole” (p.

30). The twenty episode-based comprehension tests included in the treatment (one per episode) included three exercises each: a T / F exercise, a MC exercise and an item sequencing exercise (see Appendix E). They aimed at measuring the extent to which participants in both groups understood what had happened in the video they had just seen. In order to make sure that all participants understood the questions (and that their English language proficiency was not hindering their ability to understand the questions), the tests were administered in Spanish. All comprehension tests were exactly the same for both groups¹.

The three exercises included five items each and each exercise explained in detail what students needed to do. Exercises one (T / F) and two (MC) were on the front page of the paper whereas exercise 3 (sequencing) was on the back. Each exercise included items from the beginning, middle and end of each episode and the items in each exercise were presented in the same order, following the storyline of the video. Careful attention was paid to the formulation of all items so that all statements were clear and concise and avoided misunderstandings. The maximum total score of the comprehension test was 14 points, which was the sum of the three exercises that the test comprised (5 points for exercise one, 5 for number two and 4 for exercise three). The students were not penalised for incorrect answers in any of the exercises.

Exercise 1 contained five statements where learners had to mark if they were True (T) or False (F). Exercise two contained five MC questions, each with three possible

¹ There were two episodes, *Muddy Monkey* and *Curious George Bunny Hunt* (corresponding to week 6 and 14, respectively), in which a frog and a bunny had different names in the subtitles depending on whether they were in the L1 or the L2. In this case, the names given to the animals were taken from the episode that each group watched.

alternatives (A, B and C). The three options included one correct answer, an answer related to the correct response but with incorrect information and a very different option totally unrelated to the episode. Finally, exercise three was a sequencing exercise. Students were asked to order the five events in the same chronological order in which they had occurred in the episode. The five events (statements) presented in each test were usually short and concise and dealt with some of the main events that occurred in the storyline of each video episode, from the beginning to its end. There was a low hyphen before each sentence where the participants had to write the correct number. Event number 3 of each test was always numbered so that they could have an indication of which was the event in the middle and to facilitate the ordering. The rest of the sentences were mixed up. Each correct item was awarded one point. The maximum score for this exercise was 4.

In order to create the comprehension questions, the content from each episode was first divided into idea units, which were extracted from the transcripts of each episode. When designing the item types, we also adopted Rodgers' approach and we included items that reflected the use of bottom-up and top-down processing. Topic comprehension questions indicated the use of top-down processing strategies, whereas the items inquiring about more specific details were intended to assess the students' decoding of lexical units or sentences and use of bottom-up processing.

The teacher-researcher extracted approximately 20 idea units per episode (the amount varied depending on the content of each video) and selected 15 for each test. The idea units from each episode were not repeated in any of the three exercises from the same test in order to avoid guessing from similar answers when doing the tests.

Each episode contained sufficient idea units in the storyline to create enough distinct items to be tested. It was important to formulate questions that could not be answered just by inferring from images and special attention was paid to ask for those issues that also required an understanding of linguistic input (oral or written). Items included checked understanding main/topic ideas (51.4%) or specific details (37.3%) and inferring (11.3%). Appendix J presents the item type for the idea units included in the 20 *Curious George* episodes selected for the treatment. Different idea units were used in the comprehension questions of each test: depending on their suitability, they were turned into T / F, MC or item sequencing, for example, see Appendix L to see all the idea units and item types of E16). The format of the exercises included in the test was very familiar to the participants, as these types of activity were regularly included in the course book used in their English lessons at school.

An example of an idea unit from E16 *Curious George and the Dam Builders* was when Mrs. Renkins tells George that she will have the boat fixed tomorrow first thing in the morning: *“I can have it fixed for you first thing tomorrow morning, Captain”*. This idea unit was converted into a T/ F item, number 1 in the first exercise from E16 test (see Figure 6.5).

- La Sra. Renkins le dice a Jorge que tendrá el barco arreglado esa misma tarde. V / F

Figure 6.5 – Item nº 1 of the T / F task in the comprehension test (E16).

Another idea unit included in the same episode was identified when it is raining and George is at home: *“George didn’t mind rain. It means his pond will be bigger*

tomorrow.” This was turned into an item included in the MC exercise (Item nº 4) (see Figure 6.6).

- A Jorge no le importa la lluvia, ya que piensa que así...
 - A) Los castores se irán a otro sitio.
 - B) Su estanque será más grande.
 - C) Tendrá agua aunque su presa se rompa.

Figure 6.6 – Item nº 4 of the MC task in the comprehension test (E16).

Finally, the idea unit in which: *“George and Bill watched the beavers work together to defend their home against the raging waters”* became a sequencing item (nº 5 in the sequencing task) (see figure 6.7).

- ___ Jorge y Bill observan como los castores trabajan en equipo para defender su hogar.
- ___ Jorge se ha olvidado el barco y vuelve a casa a buscarlo.
- _3_ La Sra. Renkins le dice a Jorge que coja la madera que necesite.
- ___ Bill dice a Jorge que ha sacado fotos a los castores.
- ___ Jorge decide que hará una presa para él solo.

Figure 6.7 – Sequencing task in the comprehension test (E16).

6.2.2.7.1. Item analysis

In relation to the item analysis conducted (see ‘Statistical Analyses’ section), the results for the Difficulty Index reveal that 96% of the T / F items for both groups ranged

from 'very easy' to 'medium'. Very similar results were obtained for both groups (see Appendix K for the complete results of the Difficulty Index). When examining the results for the Difficulty Index in the MC items, 92.9% (L1S) and 87.2% (L2S) are placed in the 'very easy', 'easy' or 'medium' difficulty category. However, L1S has a higher number of 'very easy' items whereas the L2S group has more 'easy' than 'very easy' items. These results show that the L1S group found the tests easier than the L2S group. On the other hand, the outcomes for the Discrimination Index in the T / F and MC items show that 50% of those were between the 'very good' and 'medium' ranges whereas the other 50% were poorer discriminators. It should be noted, though, that the learners in this study had a homogeneous proficiency level, which also implies a high percentage of items not having a high Discrimination Index.

6.2.2.8. Episode-based vocabulary tests

The twenty episode-based vocabulary tests were created for the purpose of the study and aimed at measuring L2 vocabulary learning from subtitled video watching. More specifically, they assessed participants' written word recognition of vocabulary from the video they had watched. All the tests included the same type of word recognition exercise. The instructions, written in Spanish, asked the participants to circle from a list only the words and expressions that appeared in the video they had just watched. They were also explicitly told that not all of them appeared in the video. The instruction did not indicate the exact number of words that should be circled, i.e. students could circle as many words as they wanted. There were twelve lexical items in total, either single words or MWUs (eight selected from each episode and four distractors). Of the eight TWs, four were known and four new. They were presented in four columns of

three words each. When selecting the vocabulary for this exercise, single words and also MWUs were chosen. In the vocabulary tests, each correct answer was awarded one point and wrong answers were not penalised, so the maximum score was 8 and the minimum 0.

The single words included nouns, verbs, adjectives and adverbs (approximately 70% of the items). The MWUs in the test included collocations (13.3%), compounds (7.1%), phrasal verbs (2.9%), linking adverbials (1.7%) and speech act formulae (0.8%). In order to be considered a collocation, a two-word combination needed to be found in either one of the following dictionaries: *The English Collocations Dictionary Online*, *The Online Oxford Collocations Dictionary*, *The Longman Collocations Dictionary and Thesaurus* or *The Free Online Collocations Dictionary*. The compounds included in the tests could be spelled with only one word (e.g., 'snowshoes'), two words (e.g., 'mobile phone') or separated by a hyphen (e.g., 'bubble-maker'). The linking adverbials selected included prepositional (e.g., 'over there') and adverb phrases (e.g., 'too fast'), whereas the speech act formulae consisted mainly of greetings (e.g., 'Merry Christmas'). The complete inventory and categorisation of single words and MWU included in the tests can be found in Appendix F.

The transcripts of all the episodes from the treatment were downloaded from two websites *Beano Wikia* (<http://beano.wikia.com>) or *TV Ark* (<http://tv.ark.com>). To obtain the list of the most frequent words from each *Curious George* TV episode, all the transcripts were processed through V_Words (Meara, 2010), a utility program that outputs type and token counts for short texts by Lognostics (see examples of transcripts from two episodes in Appendix I). The selection of the TWs was carried out

taking into account the frequency of the items in each episode (the most frequent items were selected over those that did not appear as often).

The most frequent lexical items (either single words or MWUs) were extracted from the episode and classified into ‘known’ and ‘new’ by the participants, depending on whether they had been taught to the students in regular classes. The inclusion of four TWs already known in each vocabulary test served two purposes: (1) to find out whether the learners recognised them more often than the new ones that had never appeared in class before when watching the video episodes and (2) to comfort the students who were not able to recognise any of the new TWs (e.g., ‘door’ vs. ‘shown live’ in E1, ‘frog’ vs. ‘muddy’ in E6, ‘easy’ vs. ‘hutch’ in E14 or ‘lake’ vs. ‘tadpoles’ in E18). Table 6.2 shows the type and number of lexical items included in each test.

Table 6.2 –Target words (TWs) and distractors in each vocabulary test.

Lexical items			
TWs	New		4
	Known		4
Distractors	Semantically related to the content of the episode	New	1
		Known	1
	Not semantically related to the content of the episode	New	1
		Known	1

Knowledge about the students’ English proficiency by the teacher-researcher was determinant in selecting those items. The teacher-researcher had been the only EFL teacher of the participants since they started learning English so she had the ability to recognise, from the most frequent lexical items, which words would probably be known

by the students. These known words had already been introduced to the students in previous courses through the students' textbook, "*The Incredible English Kit*" by Oxford University Press. In addition, other language learning resources such as graded reader books were employed in their English lessons, and the teacher could easily identify the words students would know. The distribution of words and MWUs into different categories appears in Table 6.3. The complete list of single words and MWUs per episode with its frequency of occurrence is detailed in Appendix G.

Table 6.3 – Word type figures for known and new TWs and distractors.

Single Words				MWUs
TWs New				
Nouns	Verbs	Adjectives	Adverbs	
30	12	4	1	Referential Function: 28 Textual Function: 3 Free Word Combinations: 2 Total: 33
TWs Known				
51	6	3	0	Referential Function: 16 Textual Function: 1 Communicative Function: 1 Free Word combinations: 2 Total: 20
Distractors				
49	7	4	4	Referential Function: 12 Communicative Function: 1 Free Word Combinations: 3 Total: 16

It should be noted that, whenever possible, a combination of salient vs. non-salient words was chosen in each test. Several studies have tried to identify what saliency means in SLA. Goldschneider and DeKeyser (2001, p.22) claim that saliency has been defined as referring to "how easy it is to hear or perceive a given structure". Recent studies in language acquisition have broadened the notion of salience. Ellis (2016,

p.342) claims that it can also be used to refer to items that “stand out from the rest”. In this study, salience was considered by including those words that were key in the episode or were crucial to understand the story. For example, the words ‘water’ (known) or ‘dam’ (unknown) when a barrier was constructed so as to keep the pond full in episode 16.

The four distractors included in each test were words and MWU not appearing in the video. The distractors were selected taking into account their semantic relationship to the content of the episode and the previous knowledge of the students (half of them were semantically related to the content in the episode and the other half were not) and could be either single words or MWUs. Also, following the same criteria for TWs, half of them were known and half were new.

6.2.2.9. School English mark

A global English proficiency mark was awarded by the school as part of students’ regular assessment for the English subject. This mark was obtained each trimester from the unit tests the students took (70%), their daily work (20%) and their attitude (10%) in their regular English lessons. Regarding the unit tests, they evaluated the four basic skills of language learning: reading, listening, writing and speaking. In each test there were grammar and vocabulary tasks, a written text with some comprehension questions and one listening exercise. The global English proficiency mark used in the present study was obtained from the average mark of the second and third trimesters of the 2015-16 school year. The grades from the first trimester were not taken into account because the study comprised the second and third

trimester, so it was believed those marks would better reflect the proficiency level of the students at the time of data collection.

6.2.3. Procedure

In this study, two groups of primary school young learners of English watched one episode a week of the animated TV series *Curious George* -with subtitles either in English (N=47) or in Spanish (N=45)- over a period of 5 months (20 episodes). These beginner students were tested on comprehension and vocabulary recognition (i) immediately after watching each episode and (ii) in two special episodes without subtitles (middle and end of treatment, episodes 10 and 20). In order to make sure that the instruments were appropriate for the students at this level, a pilot study was conducted in a different semi-private school from the one in the main study in advance.

6.2.3.1. Pilot study

Some instruments of the study required piloting because they were designed for the purpose of the present research and had not been used on similar populations previously. The appropriateness of the TV series chosen also needed to be checked in terms of comprehension and lexical demands (at the same time, students should find it a motivating experience). The OSE questionnaire was first piloted with Grade 4 students at the same school where the study was performed. Even if these students were a bit younger than the participants in the main study, it was hypothesised that if these students had difficulties answering any questions, the participants of the main study would as well, due to the fact that the students' profiles were similar. 22 students

took the questionnaire in November 2015, one and a half months before starting the study. The piloting showed that the students had difficulty understanding some questions because their answers were not fully appropriate and left some information gaps. Therefore, it was decided that the first part of the questionnaire would be revised and rewritten and it would be read aloud by the teacher-researcher instead of just leaving the students to do everything on their own. She would read the question and provide the students with the necessary time to answer each question in turn, before moving on to the next one. Another change introduced after the piloting was that only the students who had answered that they watched TV series or movies in English would be asked the remaining questions individually by the teacher-researcher. She would be noting down the answers of these questions in the questionnaire, as some students did not provide the necessary information.

Another semi-private primary school in a nearby area was contacted and visited so as to pilot the updated versions. The school approved the piloting with their Grade 5 and 6 students (same age as the participants in the main study). One intact group in Grade 5 ($n=26$) and one in Grade 6 ($n=27$) answered the video-watching questionnaire and the episode-based comprehension and vocabulary test from week 1 one month before starting the study (December, 2015). The students participating were very cooperative and their English teacher, who was present during the piloting, also encouraged them to do the tasks following the indications provided. The participants started watching the first two minutes of the episode chosen for the questionnaire: *Charkie escapes*. In order to avoid disclosing important information that could have helped them to figure out the main ideas of the storyline, the title of the episode was not shown.

After viewing the video, the teacher-researcher distributed the video-watching questionnaire and the learners were told they were going to answer only the first part (front page). The teacher-researcher read the questions aloud and waited for the students to finish answering each question until moving on to the next. When all the students were finished with this part, they were told to continue with the second part of the questionnaire, where they could find the comprehension and vocabulary questions on the video segment they had just watched. The teacher-researcher read all the questions aloud and, after that, advised the students to re-read the questions carefully before answering. The students were told they had no time limit to complete the questionnaire and when they finished, they raised their hands and the teacher-researcher collected it. It took them about 30-35 minutes to complete both parts (10-15 minutes for the first one and 15-20 minutes for the second one). No modifications were made to this questionnaire, as the participants showed no problems doing it and they could also answer the comprehension and vocabulary questions without any ceiling effect.

After completing the video-watching questionnaire, the same students watched the first episode selected for the study and completed its corresponding episode-based comprehension and vocabulary tests during the same session (December 2015). Participants watched *A Zoo Night* and immediately after it, the researcher distributed the tests. The students remained quiet while they were watching the video and also when they were taking the test. When they had finished, they raised their hands and the researcher collected the tests. There was no time limit and they were advised not to rush and take their time. It took them about 10-15 minutes to complete both tests. Additionally, they all replied positively when asked if they had enjoyed watching the

TV series. Even their English teacher mentioned she would buy the DVDs for the class, as she had noticed they had really enjoyed watching the video. Students also confirmed they had not watched the series before, which confirmed our prediction.

6.3.2.2. Main study

The longitudinal study for this PhD thesis was conducted with the two groups in a classroom setting: they both watched 18 episodes of the animated TV series *Curious George*, but one group watched the series with Spanish subtitles (L1S) and the other with English (L2S) subtitles. Furthermore, the participants watched two episodes of the same TV series without subtitles, one in the middle and one at the end of the treatment. The study took place in a semi-private school in Catalonia (Spain), and it consisted of multiple data collection times throughout a 5-month period (the second and third terms of the 2015-16 academic year). The teacher-researcher designed the tests on the TV series and collected all the data herself. She was responsible for giving clear instructions on how to do the tests, which were pen-and-paper, except for the Spanish and English RS tests (which were taken on the school iPads in class). Table 6.4 provides the calendar of the study, which comprised pre-treatment and treatment sessions, as detailed below.

Table 6.4 – Calendar of the study

Pre - treatment sessions	Date
RS test – Spanish	26/11/2015 – 27/11/2015
RS test – English	02/12/2015 – 03/12/2015
Vocabulary Size Test	09/12/2015 – 10/12/2015
Out-of-school exposure questionnaire	15/12/2015 – 17/12/2015
Video-watching questionnaire	11/01/2016 – 12/01/2016
Week 1 - A Zoo Night	14/01/2016 – 15/01/2016
Week 2 - Curious George Rocket Ride	18/01/2016 – 19/01/2016
Week 3 - Curious George the Architect	25/01/2016 – 26/01/2016
Week 4 - Animal Magnetism	01/02/2016 – 02/02/2016
Week 5 - Curious George Station Master	09/02/2016 – 11/02/2016
Week 6 - Curious George Muddy Monkey	15/02/2016 – 16/02/2016
Week 7 - Curious George Makes a Stand	22/02/2016 – 23/02/2016
Week 8 - Curious George Door Monkey	29/02/2016 – 01/03/2016
Week 9 - Curious George Ski Monkey	07/03/2016 – 08/03/2016
Week 10 - Curious George Home for Pigeons (Test) (No subtitles)	15/03/2016 – 17/03/2016
EASTER BREAK	22/03/2016 – 24/03/2016
Week 11 - Curious George the Grocer	31/03/2016 – 01/04/2016
Week 12 - Curious George Bee is for Bear	04/04/2016 – 05/04/2016
Week 13 - Curious George the All Animal Recycled band	11/04/2016 – 12/04/2016
Week 14 - Curious George Bunny Hunt	18/04/2016 – 19/04/2016
Week 15 - Curious George Buoy Wonder	26/04/2016 – 28/04/2016
Week 16 - Curious George and the Dam Builders	03/05/2016 – 05/05/2016
Week 17 - Curious George Roller Monkey	09/05/2016 – 10/05/2016
Week 18 - Curious George discovers the Poles	13/05/2016 – 17/05/2016
Week 19 - Curious George Low High Score	24/05/2016 – 26/05/2016
Video-watching questionnaire	24/05/2016 – 26/05/2016
Week 20 - Curious George versus Winter (Test) (No subtitles)	31/05/2016 – 02/06/2016

6.2.3.2.1. Pre-treatment sessions

Before starting the treatment, the participants of the study took different tests. All these tests were administered in the regular English classes the learners received at school.

Participants took RS tests (approximately 10-15 minutes each) in two languages:

Spanish and English (the teacher-researcher in the primary English classroom of the school was in charge of administering these tests. The RS in Spanish is not usually measured in schools in Catalonia, as Catalan is the vehicular language in which instruction is provided. However, given the fact that most students were Spanish-dominant bilinguals and that the L1S group would watch the *Curious George* video episodes with Spanish subtitles, students also took a RS test in this language. The Spanish and English RS tests were taken online so all the students could work on an individual basis with the iPads in the English classroom at the same time. The instructions to take these two tests were explained by the teacher-researcher in Catalan to avoid any confusion and to make sure the participants took the tests correctly.

The procedure to take the Spanish RS test was the following: the teacher distributed the iPads to the students and wrote the website of the test on the white board, she told the students to access it and wait for more instructions. The students accessed the site and they were instructed to begin the test. They were also told that, after finishing reading, they would answer some questions about it. When they pressed the 'start' button, the text appeared on the screen. After reading it, they pressed the 'stop' button and they answered the four comprehension questions. Finally, after completing all the questions, they pressed the 'finished' button to obtain their results. Then, they raised their hands and waited for the teacher to come and write down their RS and comprehension scores in her notebook.

In the case of the English test, also performed on the schools' iPads, all of the students were required to choose the same text in order to obtain valid and reliable results. As

reading in the L2 was more difficult for students due to their young age and their low proficiency level, one sample test was administered as a trial so as to make them feel comfortable before administering the real one. Accessing the text selected to test the RS on the website required following several instructions. The teacher-researcher chose a random text from the same website and showed the students how to perform the whole process on the board, so that the participants could clearly see the required steps to follow when doing the test. She also informed them that when they finished reading the text some questions would appear on the screen. After this information was provided, students took the real test and answered the four comprehension questions following the instructions provided. When they finished, the learners obtained their RS and the results on the comprehension test and raised their hands so that the teacher-researcher could write down the RS scores in her notebook.

The VST was taken in class the week after taking the RS tests. The teacher explained how to complete the test in Catalan so as to avoid any confusion. As the participants in the study were beginner learners, they only took the test up to the 3rd frequency band (i.e. from 0 to 3,000 words), which includes 30 questions (10 words per band). The teacher handed out the test to the students and she read aloud each item with its four possible answers once, then she waited for them to answer. After completing the 30 questions, the test was collected and corrected by the teacher-researcher.

In the next two sessions, the students answered two questionnaires before starting the treatment. First, they were given the OSE questionnaire. All the instructions on how to fill in the questionnaire were given in Catalan. The teacher-researcher distributed the questionnaires and told the students to write their name, date and

grade. She explained to the learners that she would be reading the questions aloud and that she would wait for them to answer. They were also allowed to ask questions if they had any doubts after the teacher had read each question. The participants were given unlimited time to answer and the teacher waited until everyone had answered each question before reading the next one. The same process was followed for the first five questions. When everyone had finished answering question 5, all the questionnaires were collected. The teacher-researcher checked all the questionnaires after class and in the next session she called the students who said they watched films, TV series or cartoons in English in order to ask them individually questions six and seven of the questionnaire: one student at a time was interviewed while the rest of the group performed some other tasks. The student sat down next to the teacher and the teacher read the question to them and wrote down the answer. There was no time limit to answer the questions and the teacher waited for the student to answer one question until moving on to the next one. Taking this part of the questionnaire individually took much more time than letting the participants do the questionnaire on their own, but considering the young age of the learners and the responses received during the pilot study (sometimes incomplete), it was concluded that the best way to get reliable answers from the students was to perform this part of the test individually.

After completing the OSE questionnaire, the two groups answered the video-watching questionnaire during the second week of January 2016, before they started viewing the *Curious George* episodes included in the treatment. The video excerpt the participants watched to answer this questionnaire contained L1 or L2 subtitles (depending on the group). Immediately after watching the video, the teacher-researcher distributed the questionnaire to them and provided the basic instructions in

Catalan before they started answering it. She told them that there were two parts and that they would start with the first one. The students were not given a time limit to perform it (it took them approximately 20 - 25 minutes to complete it). In the first part, the teacher-researcher read the reading behaviour questions aloud one-by-one. She paused after each question until all the students had finished answering it. When students finished answering the first part, she read all the questions in the second part aloud and gave the students unlimited time to answer them. All questionnaires were collected as soon as the students had finished.

At the end of this session, the teacher-researcher asked the participants if they had liked the video excerpt so as to make sure they had enjoyed it. The students answered very positively and confirmed to the teacher-researcher that they had not watched *Curious George* before. They also requested her to let them watch the rest of the episode but their demand could not be fulfilled as exposure to videos was highly controlled in both groups. They were told they would repeat the experience soon.

6.2.3.2.2. Treatment sessions

The study treatment sessions included 20 episodes in total and they started after the Christmas holidays (January, 2016). Participants watched eighteen episodes with either L1S or L2S and two episodes without subtitles (E10 –in the middle- and E20 at the end of the treatment). The soundtrack of all these episodes was always in English. Learners viewed the episodes at the beginning of their English lessons at school. The Grade 5 students usually watched the episodes and took the tests on Tuesdays, whereas the Grade 6 pupils did that on Mondays. In the case there was a bank holiday

on the day they usually watched the episode, the activity was carried out in their next English class.

In each session, participants took an episode-based comprehension test and the vocabulary test immediately after watching the *Curious George* episode (one per session). These tests were administered in order to assess comprehension and vocabulary learning from watching the videos. The scores from each participant were introduced in a SPSS matrix in order to perform the corresponding analyses. Before watching the first episode of the treatment, the teacher-researcher asked the students to pay attention and told them that they were going to do some exercises related to the video after watching it. They were advised that subtitles would also appear in the cartoons (which they could read or not, as they wished). Watching TV with subtitles is not a common activity among learners in schools across Catalonia. As a result, the teacher-researcher took special care to tell them before starting the activity. In the following sessions, it was not necessary to repeat the instructions, as the students clearly understood the procedure to follow. Table 6.4 shows the calendar of the whole study and the viewing order of the twenty episodes included in the treatment.

In each session the same procedure was followed at the beginning of the class: participants watched the video and took the episode-based comprehension and vocabulary tests. Students were very quiet and watched the video attentively. They were also instructed to read the questions carefully as there was no time limit to perform the test. If they had problems, they were allowed to ask the teacher-researcher questions by raising their hands and the teacher approached each student individually, although many of the questions could not be responded because

participants were requesting the answers to the questions. When they finished, they raised their hands and the teacher collected the tests. The learners were in complete silence while performing the tests and they were asked to remain quiet until the last student had finished. It took them approximately 10 - 15 minutes to complete the comprehension and vocabulary questions. Before watching episodes 10 and 20 (no subtitles), the teacher warned the students that the videos did not have any subtitles. They were told it was a problem with the computer. Even though they complained, they were told that the procedure would be the same as in the other sessions anyway.

The treatment was interrupted for one week during the Easter holidays (March 2016). In the first session after this break (E11), the teacher-researcher told the students they were going to continue with the same practice in the English classroom and that they would continue watching *Curious George* weekly. She also asked them if they remembered what they were required to do. Although they responded affirmatively, she reminded them of the steps to follow and how the session was going to unfold. The students also inquired about the presence of subtitles: they were told that the problem in E10 had been fixed and that they would be able to watch the episode with subtitles (they looked relieved when they knew they would be watching the video with subtitles). The rest of the treatment took place as planned and the students watched the weekly episode without any other difficulty until its end (May 2016). The TV series was very successful among the students and they were highly motivated and eager to watch it every week (they even asked the teacher-researcher if they could watch more episodes when the study had finished).

Nearly at the end of the treatment (week 19), immediately after the participants had taken the corresponding episode-based comprehension and vocabulary tests, they were asked to fill in the first part of the same video-watching questionnaire they had filled in at the beginning of the study (the comprehension and vocabulary questions from the video segment of *Charkie Escapes* were removed from the test) as we were actually interested in their viewing behavior, so they answered the questions in relation to the last episode they saw. Furthermore, we had already tested comprehension and vocabulary recognition throughout the whole study and there was no need to check it again. The students only answered the questions on how they were performing the watching activity. This self-reported information on their behaviour would help to throw light on how they approached this practice before and after sustained exposure to multimodal input (e.g., to see if the reading behaviour of the participants in both groups had changed or if it was unaltered). According to the teacher-researcher's knowledge, at the time of planning the study of this dissertation, there was no data available in our context related to viewing behavior before and after sustained exposure. Finally, the last episode of the treatment (E20) was watched without subtitles, following the same procedure as in episode 10. The teacher-researcher told the students the computer did not work properly, so they watched the video without subtitles. The learners complained about the absence of written input and their faces reflected disappointment, but they did the activity as planned.

6.2.3.3. The teacher-researcher

It has been claimed that educational research should be conducted by teachers and integrated as a part of their jobs (Hammersley, 1993). According to McNiff (2013),

Action Research is “a method of systematic enquiry that teachers undertake as researchers of their own practice” (p. 90). In this way, the teacher can understand, observe, interact, gain confidence and find out where learners are in terms of proficiency and understanding before the study and, due to that, know them better than a researcher (Kosnik & Beck, 2000). Moreover, it has been suggested that when conducting research with young learners, teachers develop a deeper understanding of their students’ learning process and show a stronger commitment to the development of a study they have planned and designed (Borgia & Schuler, 1996).

Studies with young learners are scarce, probably due to the extra difficulty of conducting research with students at these ages: children may find it hard to understand why an outsider (researcher) enters the classroom and tells them what to do. This lack of confidence may lead to misunderstandings and data that is not accurate, which is why their own teacher can be a better researcher in this context. On the other hand, there is some concern about teacher-researchers and their ability to “ensure that selected modes of inquiry are appropriate to the research problem” and if they “can assess whether they have, in fact, achieved the desired outcomes for the research process” (Newton & Burgess, 2008, p.24).

As has been mentioned previously, the author of the present study was the English teacher of all the groups throughout the whole treatment. This teacher-researcher had known the participants for more than four years and had been their only English teacher at primary school. Participants were not told about the study because it could have influenced their performance while viewing the episodes and also the results of the tests. The study was conducted in a non-disruptive way for the participants, and in

similar conditions to those usually found in EFL primary school classes. The only information the participants were given was that the teacher wanted to check if they had understood the videos they watched in English. The activities they performed in the English class were different from the ones they regularly did for the study. However, they felt comfortable and did not ask many questions about the practice of watching subtitled cartoons in class. They explicitly said they missed doing that when the treatment had finished. It should also be mentioned that, when taking the tests, all the students were in complete silence until they had all finished. Students' behavior was excellent in all sessions and it was clear they felt comfortable in class with their usual teacher.

6.2.4. Data analysis

The tests administered to the participants were never corrected in class. They were marked by the teacher-researcher following the criteria presented below. In this section we also explain the analyses performed to answer the RQs proposed.

6.2.4.1. Test scoring

6.2.4.1.1. Vocabulary size test

The *Vocabulary Size Test, version A* (Nation & Beglar, 2007) was scored by the teacher-researcher dichotomously. One point was awarded for each correct answer and students were not penalised for wrong answers. The maximum score the learners

could obtain was 30 points, as only the first three bands of this test were completed by the students.

6.2.4.2.2. Reading speed tests

The Spanish and English RS tests were taken online and their scores (WPM in Spanish or English) were provided by the computer software of the two programs selected to do the tests. In addition to these RS tests, the students also took a comprehension test after reading the extract. The scores for these two tests were expressed in percentages: the minimum score that could be obtained was 0% and the maximum 100%. It should be noted that participants not obtaining 75% correct answers in the RS test were not included in the analysis involving this measure. This was considered to be an adequate level of understanding, as 50% (i.e., answering 2 out of 4 questions correctly) could be reached by chance. The final number of participants for the RS analyses were 32 for L1S and 23 for L2S.

6.2.4.2.3. LLAMA_B test

In the LLAMA_B test, the results were provided on the screen by the programme's computer software when the students finished the test. The maximum score that could be attained for this test was 100 points (each correct answer was awarded 5 points) and the minimum 0 points.

6.2.4.2.4. Episode-based comprehension and vocabulary tests

The episode-based comprehension and vocabulary tests were scored dichotomously. Regarding the comprehension test, each of the T / F, MC and item sequencing questions were awarded one point for each correct answer. There were five questions in the T/F and the MC exercise, which amounted to a maximum total of 10 points. Incorrect answers were not awarded any points and carried no penalties. Regarding the sequencing exercise, it contained five sentences that had to be ordered chronologically but sentence #3 was already provided in all tests in order to aid participants in performing this task. Due to that, the maximum score that learners could obtain in this exercise was 4 points. Hence, the highest score that learners could achieve in the whole comprehension test was 14 points (5+5+4).

On the other hand, the vocabulary test included twelve words and four of them were distractors. Our learners were awarded one point for each word-form they recognised correctly and they were not penalised if they selected any of the four words not included in the episode watched. The maximum score the learners could get in this test was 8 points.

6.2.4.2. Statistical analyses

In order to provide an answer to the four RQs in our study, several statistical analyses were conducted, as presented below. Additionally, we carried out some reliability analyses on the comprehension tests. All statistical data analyses were performed using SPSS (version 20). For all tests, unless otherwise indicated, the alpha level was

set at .05. Prior to performing each of the statistical analyses, data were assessed for normality and when this assumption was violated, scores were treated non-parametrically. Specific decisions that were taken when needed in each case are also indicated. To guide this section, we reproduce the RQs in each section followed by the results and the findings.

6.2.4.2.1. RQ1: Effects of watching L1/L2 subtitled TV series on L2 comprehension and vocabulary learning

First of all, in order to analyse if both groups experienced an increase in their L2 comprehension and vocabulary recognition scores throughout the longitudinal treatment, and if there were significant differences between groups, Repeated Measures (RM) ANOVAs with the L2 comprehension and vocabulary scores as the dependent variables and time and group as factors were performed. Then, Mann-Whitney *U*-tests were performed in order to more specifically check whether the L1/L2 subtitled input provided played a significant role in L2 comprehension and word-form recognition throughout the intervention. Furthermore, Mann Whitney *U*-tests were also conducted with E10 and E20 scores, which did not include any subtitles. In this way, we could assess possible differences between the two groups, when subtitles were not provided (in the middle and at the end of the treatment). In order to obtain more accurate results regarding L2 word-form recognition, Mann Whitney *U*-tests were performed to examine whether L1S or L2S were more convenient for recognition of new and known words and whether there were significant differences between recognising these lexical items after video watching.

6.2.4.2.2. RQ2: Relationship between L2 comprehension and vocabulary scores with aptitude and proficiency variables

To answer the second RQ, (Are episode comprehension and written vocabulary recognition scores related to linguistic aptitude and proficiency variables (i.e., L1/L2 reading speed, L2 vocabulary size and school English marks)?), we examined the relationship between the scores on comprehension and vocabulary with the LLAMA_B test, as well as RS, VS and English scores. Spearman Rho product-moment correlation analyses were performed between the aptitude/proficiency variable and the results from comprehension and vocabulary tests. Moreover, we also made a difference here between new and known words.

Finally, regression analyses were also performed to determine the possible impact of the variables on comprehension and vocabulary scores more specifically. As VS and class grade were correlated, we just entered VS into the regression to avoid collinearity problems. The ANOVAs with the predictor and dependent variables will also be presented to provide a more complete picture of the significant regression results.

6.2.4.2.3. RQ3: Relationship between extensive viewing and viewing behaviour

To answer the third RQ, we explored whether repeated viewing affected the way our students watched the animated TV series. The data provided by the participants' responses in the video-watching questionnaire, at the beginning (Time 1 [T1]) and at end of the treatment (Time 2 [T2]), was taken into consideration. A descriptive analysis

of the answers in the questionnaire is provided through several tables and diagrams comparing the two groups. Finally, in order to statistically compare the results before and after the treatment in each group, McNemar's tests were carried out, as well as an item analysis for the comprehension tests.

As these tests were designed by the researcher, item analyses were conducted to assess whether test items from the T /F and MC exercises produced reliable results. The Difficulty and the Discrimination Indexes were computed through the Prof Testing website (http://www.proftesting.com/test_topics/steps_9.php) (retrieved November 8, 2018). According to McCowan and McCowan (1999, p.18) the Difficulty Index is to do with “the percentage of people who answer an item correctly”. It can range from 0 to 1 and the higher the Difficulty Index is, the easier the item will be to answer. On the other hand, “item discrimination compares the number of high scorers and low scorers who answer an item correctly” (p.20). It is believed that good students are more likely to answer the item correctly, whereas poor students are more likely to answer it incorrectly. The Discrimination Index ranges from 1 to -1. Usually, positive figures show that the learners who tend to score higher also obtain high results on the item, whereas a negative figure reveals that the low-achievers who usually score lower on the test are also more likely to score lower on an item.

6.2.4.2.4. RQ4: Effects of OSE on L2 comprehension and vocabulary learning from extensive viewing

To answer the fourth RQ, (Does out-of-school-exposure (OSE) to multimodal input have an effect on episode comprehension and vocabulary scores when watching

subtitled TV series?), we examined the results from Q5 in the OSE questionnaire through several *t*-tests. Learners in each group were categorised according to whether they had received OSE or not: *t*-tests were run comparing the results for those who had received OSE to English and those who had not.

In order to make groups depending on the OSE learners had received, we took into account Q5 in the questionnaire, which was composed of three tables with several statements each. To answer this question, the students had to mark the answer that suited them better with a cross, depending on the OSE they were receiving (type of activity and frequency). Students in each of our groups (L1S and L2S) were divided into two groups, according to their OSE. The 'No OSE' group was composed of students who did not receive any exposure apart from formal instruction at school, whereas the 'OSE' group was formed by students who had been exposed to multimodal input outside the English classroom.

CHAPTER 7- RESULTS

In this chapter, the results of the analyses performed for the present study are presented and organised according to the four main RQs. Section 7.1 addresses RQ1 of the study and presents the descriptive statistics of the comprehension and vocabulary scores. Then, the results obtained by the L1S and L2S participants on the episode-based comprehension and vocabulary tests are statistically compared. Section 7.2 reports results for RQ2, which investigates the possible relationship between aptitude and comprehension and vocabulary scores on the one hand, and on the other it examines whether the proficiency variables of VS, L1 / L2 RS and the English class grade are related to the students' scores when learning from multimodal input. Section 7.3 focuses on the video-watching questionnaire and analyses students' answers before and after the intervention. Finally, RQ4 assesses whether previous exposure to multimodal input out-of-school may influence learning from subtitled TV series (7.4).

7.1. RQ1: Comprehension and vocabulary learning from extensive viewing

In this section we examined the first RQ of the study, analysing and comparing the results obtained by both groups on the 18 episode-based comprehension and vocabulary tests, as well as those obtained in the tests for the two episodes watched without subtitles (E10 and E20). We will also analyse whether there is a difference in word-form recognition between known and new lexical items in both groups.

7.1.1. L2 comprehension

In order to find out the effect of time and condition on comprehension, a between-within RM ANOVA was conducted. The results of the within-subject effects tests showed (using the Greenhouse-Geisser correction) a significant interaction between time and comprehension, $[F(12.43, 857.84) = 11.84, p = .000, \text{partial eta squared} = .146]$. Moreover, the between-subject effects test, also using the Greenhouse-Geisser correction, indicated significant differences between the participants who watched the videos with L1S or L2S and a significant interaction between groups: $[F(12.43, 857.84) = 2.29, p = .007, \text{partial eta squared} = .032]$, favouring the L1S learners.

The results of the Mann-Whitney *U*-tests revealed that the L1S learners scored significantly higher in most of the episodes (E2, E4, E5, E6, E8, E9, E11, E12, E13, E14, E15, E16 and E18 – see Table 7.1 for the results of the Mann-Whitney *U*-tests conducted with the comprehension scores from the 18 episode-based comprehension exercises performed after watching the 18 *Curious George* episodes).

Table 7.1 – Descriptive statistics and Mann-Whitney results for comprehension scores.

Episode	L2 comprehension						Mann-Whitney results
	L1S (n=47)			L2S (n=45)			
	Mean (sd)	Min.	Max.	Mean (sd)	Min.	Max.	
E1	10.62 (2.20)	5.00	14.00	9.89 (1.97)	5.00	14.00	$U(92) = 856.00, z = -1.591, p = .112$
E2	10.55 (2.23)	4.00	14.00	9.84 (2.04)	3.00	13.00	$U(91) = 786.50, z = -1.990, p = .047$
E3	10.79 (1.72)	8.00	14.00	10.00 (2.26)	6.00	14.00	$U(87) = 749.50, z = -1.689, p = .091$
E4	11.94 (1.79)	7.00	14.00	11.09 (1.99)	3.00	14.00	$U(92) = 781.50, z = -2.192, p = .028$
E5	11.93 (2.30)	6.00	14.00	9.58 (2.75)	1.00	13.00	$U(89) = 465.50, z = -4.341, p = .000$
E6	11.40 (2.26)	4.00	14.00	9.71 (2.55)	3.00	14.00	$U(92) = 620.00, z = -3.451, p = .001$
E7	12.04 (1.56)	6.00	14.00	11.27 (2.05)	7.00	14.00	$U(92) = 832.50, z = -1.794, p = .073$
E8	10.21 (2.72)	2.00	14.00	8.78 (2.23)	3.00	13.00	$U(91) = 690.00, z = -2.754, p = .006$
E9	11.34 (2.43)	5.00	14.00	10.64 (1.72)	5.00	14.00	$U(91) = 737.50, z = -2.392, p = .017$
E11	11.07 (2.56)	4.00	14.00	9.86 (2.15)	3.00	13.00	$U(89) = 649.50, z = -2.822, p = .005$
E12	12.04 (2.60)	3.00	14.00	10.41 (2.46)	6.00	14.00	$U(89) = 556.50, z = -3.605, p = .000$
E13	10.27 (2.59)	4.00	14.00	9.09 (2.42)	4.00	14.00	$U(92) = 771.00, z = -2.254, p = .024$
E14	12.58 (2.43)	3.00	14.00	11.89 (2.51)	5.00	14.00	$U(91) = 796.50, z = -1.981, p = .048$
E15	10.87 (2.54)	5.00	14.00	8.93 (2.14)	5.00	13.00	$U(91) = 568.50, z = -3.735, p = .000$
E16	11.11 (2.60)	2.00	14.00	9.44 (2.30)	5.00	14.00	$U(91) = 587.00, z = -3.585, p = .000$
E17	10.26 (2.10)	3.00	14.00	10.57 (2.31)	6.00	14.00	$U(90) = 928.00, z = -.685, p = .493$
E18	12.33 (2.20)	5.00	14.00	10.32 (2.55)	2.00	14.00	$U(89) = 499.00, z = -4.089, p = .000$
E19	10.94 (2.48)	5.00	14.00	10.36 (2.61)	3.00	14.00	$U(92) = 914.50, z = -1.127, p = .260$

Finally, the comparisons conducted between groups for E10 and E20 on comprehension showed no significant differences between the two groups (see Table 7.2). It should be noted that in E10 and E20 comprehension means are the lowest in the L1S group.

Table 7.2 – Descriptive statistics and Mann-Whitney results for comprehension scores (E10 and E20).

Episode	L2 comprehension		
	L1S (n= 47)	L2S (n= 45)	Mann-Whitney results
E10	8.69 (2.19)	8.58 (2.21)	$U(88) = 967.50, z = .000, p = 1.000$
E20	9.38 (2.69)	9.31 (2.33)	$U(92) = 1023.50, z = -.268, p = .789$

Figure 7.1 shows the mean scores obtained by both groups in the twenty comprehension tests of the treatment. The thirteen episodes for which students from the L1S group obtained significant differences in comprehension are circled.

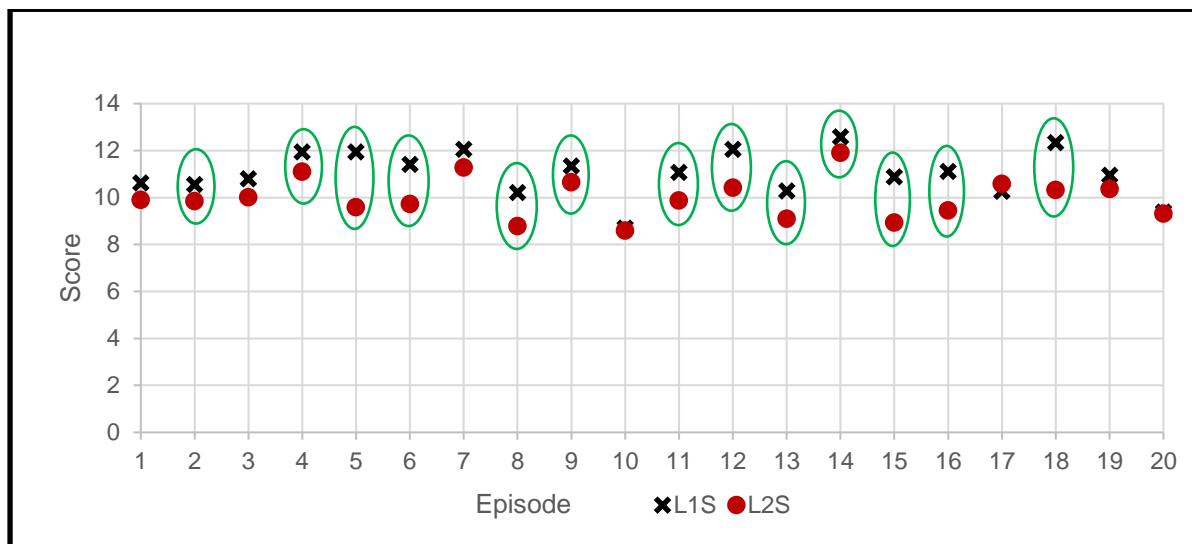


Figure 7.1 – Mean comprehension scores for all episodes. Ovals show the significant differences between groups.

7.1.2. Vocabulary learning

A between-within RM ANOVA was performed to test the significance of time and condition for vocabulary learning. The results of the within-subject effects indicated a significant effect between time and vocabulary using the Greenhouse-Geisser correction [$F(11.9, 809.06) = 19.23, p = .000, \text{partial } \eta^2 = .220$]. Furthermore, the between-subject effects test also showed a statistically significant main effect between the L1S and L2S groups, [$F(1.68) = 12.17, p = .001, \text{partial } \eta^2 = .152$]. The interaction between time and group also revealed a significant effect: [$F(11.90, 809.06) = 2.81, p = .001, \text{partial } \eta^2 = .040$].

More specifically, the scores from the Mann-Whitney *U*-tests indicated that there were significant differences favouring the L2S group in eight episodes (E1, E2, E7, E11, E15, E16, E18, and E19). Even though in three episodes the mean was higher for the L1S students (E4, E5 and E8), no significant differences were found between groups in the outcomes from these episodes (see Table 7.3 for the results of the Mann-Whitney *U*-tests conducted with the results of the word-recognition tests).

Table 7.3 – Descriptive statistics and Mann-Whitney results for vocabulary scores.

Episode	L2 vocabulary						Mann-Whitney results
	L1S (n= 47)			L2S (n= 45)			
	Mean (sd)	Min	Max	Mean (sd)	Min	Max	
E1	5.32 (1.04)	2.00	7.00	5.82 (.66)	4.00	7.00	<i>U</i> (91) = 724.00, <i>z</i> = -2.677, <i>p</i> = .007
E2	4.61 (1.36)	1.00	7.00	5.16 (1.14)	2.00	8.00	<i>U</i> (91) = 785.00, <i>z</i> = -2.058, <i>p</i> = .040
E3	4.00 (1.66)	.00	7.00	4.51 (1.21)	2.00	7.00	<i>U</i> (87) = 787.00, <i>z</i> = -1.387, <i>p</i> = .166
E4	4.28 (1.16)	2.00	7.00	4.22 (1.02)	2.00	6.00	<i>U</i> (92) = 1022.00, <i>z</i> = -2.88, <i>p</i> = .774
E5	4.66 (1.55)	1.00	7.00	4.62 (1.28)	3.00	7.00	<i>U</i> (89) = 940.50, <i>z</i> = -.415, <i>p</i> = .678
E6	4.83 (1.34)	1.00	8.00	5.31 (1.24)	2.00	8.00	<i>U</i> (92) = 847.50, <i>z</i> = -1.692, <i>p</i> = .091
E7	3.81 (1.24)	1.00	7.00	4.40 (1.18)	2.00	7.00	<i>U</i> (92) = 787.50, <i>z</i> = -2.175, <i>p</i> = .030
E8	4.72 (1.47)	1.00	8.00	4.66 (1.06)	3.00	7.00	<i>U</i> (91) = 1007.50, <i>z</i> = -.219, <i>p</i> = .826
E9	4.54 (1.15)	1.00	6.00	5.00 (1.24)	3.00	8.00	<i>U</i> (91) = 863.50, <i>z</i> = -1.408, <i>p</i> = .159
E11	4.69 (1.36)	2.00	7.00	5.70 (1.27)	1.00	8.00	<i>U</i> (89) = 576.50, <i>z</i> = -3.486, <i>p</i> = .000
E12	4.62 (1.17)	2.00	7.00	4.86 (1.07)	3.00	8.00	<i>U</i> (89) = 900.00, <i>z</i> = -.768, <i>p</i> = .443
E13	4.70 (1.32)	1.00	8.00	5.13 (1.20)	3.00	7.00	<i>U</i> (92) = 864.00, <i>z</i> = -1.554, <i>p</i> = .120
E14	5.24 (1.34)	1.00	8.00	5.33 (1.22)	2.00	8.00	<i>U</i> (91) = 1003.50, <i>z</i> = -.258, <i>p</i> = .796
E15	5.26 (1.10)	3.00	7.00	5.82 (1.50)	.00	8.00	<i>U</i> (91) = 746.50, <i>z</i> = -2.356, <i>p</i> = .018
E16	4.78 (1.21)	2.00	7.00	5.58 (1.22)	3.00	8.00	<i>U</i> (91) = 693.00, <i>z</i> = -2.789, <i>p</i> = .005
E17	5.52 (1.38)	2.00	8.00	5.93 (1.19)	4.00	8.00	<i>U</i> (90) = 863.50, <i>z</i> = -1.234, <i>p</i> = .217
E18	5.20 (1.20)	3.00	8.00	6.18 (1.35)	3.00	8.00	<i>U</i> (89) = 595.00, <i>z</i> = -3.320, <i>p</i> = .001
E19	4.95 (1.53)	.00	8.00	5.89 (1.17)	2.00	8.00	<i>U</i> (92) = 652.00, <i>z</i> = -3.261, <i>p</i> = .001

Figure 7.2 shows the mean scores obtained by both groups in the twenty vocabulary tests realised throughout the treatment. The eight episodes for which students from the L2S group obtained significant differences in word-form recognition are circled.

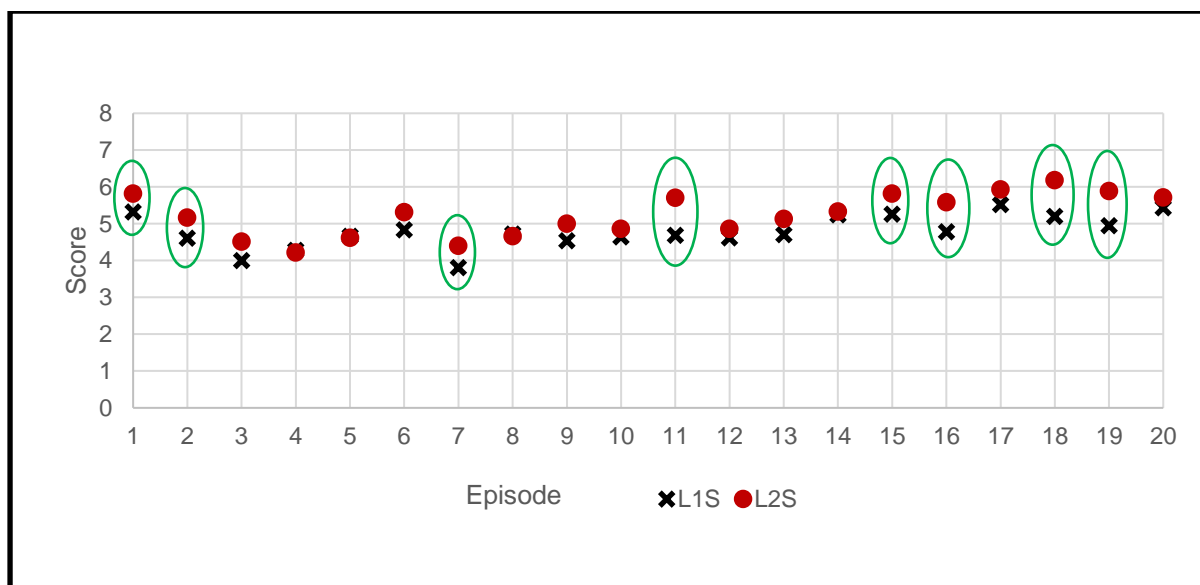


Figure 7.2 – Mean total vocabulary scores across all 20 episodes. Ovals show significant differences between groups.

In order to obtain more accurate results from the L2 word-form recognition task, Mann-Whitney U -tests were also conducted with the L1S and L2S scores for the L2 word recognition of items already known. The findings indicated that there were significant differences for three episodes, favouring the L2S group: E3 ($U = 677.50$, $z = -2.524$, $p = .012$), E7 ($U = 621.00$, $z = -3.760$, $p = .000$) and E19 ($U = 789.50$, $z = -2.455$, $p = .014$). However, in E10 ($U = 938.00$, $z = -0.67$, $p = .946$) and E20 ($U = 894.00$, $z = -1.579$, $p = .114$) no significant differences were found. The results for the Mann-Whitney U -tests conducted for all episodes are shown in Table 7.3.

Table 7.4 – Descriptive statistics and Mann-Whitney results for known words scores.

Episode	Known words		
	L1S (n= 47) Mean (sd)	L2S (n= 45) Mean (sd)	Mann-Whitney results
E1	3.81 (.50)	3.93 (.25)	$U(92) = 967.50, z = -1.302, p = .193$
E2	3.53 (.74)	3.55 (.63)	$U(91) = 1011.50, z = -.210, p = .834$
E3	3.00 (1.08)	3.53 (.77)	$U(87) = 677.50, z = -2.524, p = .012$
E4	2.55 (.88)	2.67 (.71)	$U(92) = 1029.00, z = -.243, p = .808$
E5	2.82 (.76)	2.76 (.71)	$U(89) = 918.50, z = -.640, p = .522$
E6	3.02 (.94)	3.04 (.82)	$U(92) = 1045.50, z = -.099, p = .921$
E7	2.98 (.74)	3.53 (.59)	$U(92) = 621.00, z = -3.760, p = .000$
E8	3.49 (.66)	3.57 (.59)	$U(91) = 979.50, z = -.499, p = .618$
E9	3.37 (.74)	3.53 (.79)	$U(91) = 873.00, z = -1.469, p = .142$
E11	3.22 (.77)	3.27 (.69)	$U(89) = 942.50, z = -.425, p = .671$
E12	3.64 (.65)	3.86 (.35)	$U(89) = 833.00, z = -1.810, p = .070$
E13	3.04 (.83)	3.22 (.93)	$U(92) = 923.00, z = -1.118, p = .264$
E14	3.15 (.82)	2.98 (.78)	$U(91) = 883.00, z = -1.297, p = .195$
E15	3.30 (.69)	3.47 (.75)	$U(91) = 879.00, z = -1.379, p = .168$
E16	3.61 (.61)	3.53 (.92)	$U(91) = 894.50, z = -1.290, p = .197$
E17	3.28 (.72)	3.25 (.75)	$U(90) = 992.50, z = -.173, p = .863$
E18	2.98 (.78)	3.23 (.80)	$U(89) = 808.50, z = -1.588, p = .112$
E19	3.34 (.81)	3.69 (.63)	$U(92) = 789.50, z = -2.455, p = .014$

On the other hand, the Mann-Whitney U -tests conducted with the new words scores from both groups obtained significant differences favouring the L2S group in 7 out of 18 episodes from the treatment: E2 ($U = 764.50, z = -2.237, p = .025$), E6 ($U = 749.50, z = -2.603, p = .009$), E11 ($U = 495.50, z = -4.239, p = .000$), E16 ($U = 511.00, z = -4.307, p = .000$), E17 ($U = 727.00, z = -2.450, p = .014$), E18 ($U = 596.00, z = -3.390, p = .001$) and E19 ($U = 774.50, z = -2.307, p = .021$). These results indicate that watching the videos with L2S significantly helped the identification of new words by the L2S participants in nearly half of the episodes included in the treatment. See Table 7.5 for the results of the Mann-Whitney analyses conducted with the new words variable.

Table 7.5 – Descriptive statistics and Mann-Whitney results for new words scores.

New words			
Episode	L1S (n= 47) Mean (sd)	L2S (n= 45) Mean (sd)	Mann-Whitney results
E1	1.59 (.74)	1.84 (.60)	$U(92) = 836.00, z = -1.937, p = .053$
E2	1.19 (.99)	1.66 (.91)	$U(91) = 764.50, z = -2.237, p = .025$
E3	1.04 (1.03)	.88 (1.05)	$U(87) = 856.00, z = -.812, p = .417$
E4	1.66 (.96)	1.73 (.78)	$U(92) = 993.50, z = -.536, p = .592$
E5	1.93 (1.17)	1.84 (1.13)	$U(89) = 947.50, z = -.360, p = .719$
E6	1.80 (.90)	2.24 (.71)	$U(92) = 749.50, z = -2.603, p = .009$
E7	.83 (.84)	.84 (.88)	$U(92) = 1051.00, z = -.054, p = .957$
E8	1.04 (.86)	1.11 (.97)	$U(91) = 1004.50, z = -.247, p = .805$
E9	1.17 (.93)	1.47 (1.01)	$U(91) = 878.00, z = -1.310, p = .190$
E11	1.47 (1.01)	2.43 (.99)	$U(89) = 495.50, z = -4.239, p = .000$
E12	.96 (.88)	1.05 (.96)	$U(89) = 953.00, z = -.321, p = .749$
E13	1.66 (.98)	1.91 (.82)	$U(92) = 841.00, z = -1.805, p = .071$
E14	2.09 (.96)	2.36 (.91)	$U(91) = 869.00, z = -1.399, p = .162$
E15	1.96 (.92)	2.33 (1.04)	$U(91) = 830.00, z = -1.705, p = .088$
E16	1.17 (.95)	2.22 (1.28)	$U(91) = 511.00, z = -4.307, p = .000$
E17	2.24 (.95)	2.72 (.75)	$U(90) = 727.00, z = -2.450, p = .014$
E18	2.22 (.88)	2.91 (.96)	$U(89) = 596.00, z = -3.390, p = .001$
E19	1.81 (1.08)	2.27 (.96)	$U(92) = 774.50, z = -2.307, p = .021$

Tables 7.4 and 7.5 present the results according to whether words were known or new to students. It can be observed that, when there are significant differences, these are in favour of the L2S group, and that we find it more often for the new words than for the known words. On the other hand, we do not find significant differences between groups in episodes 10 and 20 when no subtitles are added to the video (see Table 7.6).

Table 7.6 – Descriptive statistics and Mann-Whitney results for the vocabulary scores (E10 and E20)

Vocabulary scores	Episode	L1 (n= 47) Mean (sd)	L2 (n= 45) Mean (sd)	Mann-Whitney results
Total Vocabulary	E10	4.64 (1.28)	4.86 (1.19)	$U(88) = 886.50, z = -.695, p = .487$
	E20	5.43 (1.57)	5.51 (1.14)	$U(92) = 954.00, z = -.828, p = .407$
Known Words	E10	3.38 (.86)	3.40 (.77)	$U(87) = 938.00, z = -0.67, p = .946$
	E20	3.51 (.75)	3.73 (.62)	$U(92) = 894.00, z = -1.579, p = .114$
New Words	E10	1.20 (.97)	1.43 (.97)	$U(87) = 835.00, z = -.979, p = .328$
	E20	1.94 (1.24)	1.91 (1.08)	$U(92) = 1056.50, z = -.0058, p = .994$

Finally, the differences between the known and new words scores from each group were also analysed through Mann-Whitney *U*-tests (due to the violation of the assumption of normality of the variables). First, two Mann-Whitney analyses were conducted comparing groups. Significant statistical differences were found showing that the L2S group consistently recognised more known and new words than the L1S group. However, the difference is borderline in the case of the known vocabulary. Then, a Wilcoxon Signed Ranks test showed that both groups significantly recognised more known than new words (see Table 7.7).

Table 7.7 – Mann-Whitney results for known and new words scores and the known vs. new words scores.

Group	Mean (sd)		Wilcoxon Signed Ranks test results
	Known (out of 72)	New (out of 72)	
L1S	57.0 (6.95)	27,28 (9.48)	$z = 5.97, p < .001$
L2S	60.0 (5.72)	33.53 (8.43)	$z = 5.84, p < .001$
Mann-Whitney test results	$U(92) = 788.500, z = -2.104, p = .035$	$U(92) = 683, z = -2.928, p = .003$	

7.1.3. Summary of findings

In conclusion, the RM ANOVAs performed for both comprehension and vocabulary scores indicated significant effects for the within and between-subject tests. When Mann-Whitney analyses were conducted, the results revealed that the L1S group scored consistently higher in comprehension, whereas the L2S group was better at L2 word-form recognition (even if significant differences were not found in all episodes). The tests conducted to assess recognition of known and new words showed that both groups tended to recognise words that were known significantly more often and that,

when significant differences were shown, these also favoured the L2S group with regards to new word recognition.

7.2. RQ2: Aptitude and proficiency effects on learning from extensive viewing

In this section we examined the second RQ of the study. Spearman Rho product-moment correlations were conducted with the average score of the 18 episode-based comprehension and vocabulary tests as well as with the LLAMA_B (aptitude), L1 / L2 RS and VST scores and the learners' English class grade. Correlation analyses were also conducted between aptitude and proficiency variables and the scores obtained by students in E10 and E20 (watched without subtitles).

Table 7.8 presents the descriptive results in each group for the variables under study: aptitude, L1 / L2 RS (expressed in WPM), VST and the English class grade (in percentages). The results of the correlations performed between the mean comprehension and vocabulary scores with aptitude and proficiency variables are reported in Tables 7.9 and 7.10.

Table 7.8 – Descriptive statistics for aptitude, L1/L2 RS, VS and English class grade scores.

	L1S				CI	L2S				CI
	Mean	SD	Min.	Max.		Mean	SD	Min.	Max.	
Aptitude	4.98	3.09	0.00	14.00	5.06 – 8.35	5.1	2.7	1.00	16.00	4.40 – 6.41
Reading Speed (RS)	146.94	57.85	51.00	317.00	121.8 – 173.1	95.03	40.44	32.00	171.00	79.93 – 110.13
Vocabulary Size (VS)	12.69	4.08	0.00	22.00	13.02 – 15.56	11.76	4.98	0.00	22.00	10.45 – 14.49
English class grade	71.70	17.58	40.00	95.00	71.65 – 77.17	71.56	17.80	35.00	95.00	70.07 – 81.93

Table 7.9 – Spearman Rho correlations between the comprehension scores and aptitude, L1/L2 RS, VS and English class grade scores.

	Comprehension	
	L1S	L2S
Aptitude	Rho=.510** , n = 47, p <.01	Rho=.523** , n = 44, p <.01
Reading Speed (RS)	Rho=.371* , n = 47, p <.05	Rho=-.091, n = 30, p = .631
Vocabulary Size (VS)	Rho=.570** , n = 47, p <.01	Rho=.474** , n = 45, p <.01
English Class grade	Rho=.507** , n = 47, p <.01	Rho=.754** , n = 45, p <.01

Table 7.10 – Spearman Rho correlations between the vocabulary scores and aptitude, L1/L2 RS, VS and English class grade scores.

	Vocabulary	
	L1S	L2S
Aptitude	Rho=.330* , n = 47, p <.05	Rho=.076, n = 44, p = .625
Reading Speed (RS)	Rho=.207, n = 32, p = .256	Rho=-.132, n = 30, p = .487
Vocabulary Size (VS)	Rho=.445** , n = 47, p <.01	Rho=.035, n = 45, p = .819
English Class grade	Rho=.297* , n = 47, p <.05	Rho=.283, n = 45, p = .060

7.2.1. Aptitude in relation to comprehension and vocabulary scores: correlations

The analyses revealed strong significant correlations between aptitude scores and comprehension for both groups: L1S (Rho = .510**, n = 47, p < .01) and L2S (Rho = .523**, n = 44, p < .01). Furthermore, there were significant correlations in the episodes watched without subtitles, although these were just moderate and less strong (and not consistent for both groups): L1S showed a medium significant correlation in E10 (Rho = .320*, n = 45, p < .05), whereas L2S showed one for E20 (Rho = .344*, n = 44, p < .05). The language aptitude of the participants from both groups was moderately related to their comprehension of the TV episodes included in the treatment, watched either with or without subtitles.

The analyses showed a medium significant correlation between learners' aptitude and the total vocabulary scores for the L1S group ($Rho = .330^*$, $n = 47$, $p < .05$), whereas no significant results were obtained for the L2S learners. In addition, the episodes without subtitles also showed one medium significant correlation for L1S in E10 ($Rho = .363^*$, $n = 45$, $p < .05$) and a stronger one for E20 ($Rho = .498^{**}$, $n = 47$, $p < .01$), which also indicated a positive relationship between aptitude and written lexical recognition when subtitles were not included in the videos in the L1S group.

When we make a distinction between known and new words, the Spearman Rho product-moment correlations conducted between the aptitude scores and the known words variable means revealed a medium significant correlation for the L1S group ($Rho = .346^*$, $n = 47$, $p < .05$), but none for the L2S one. In addition, it should be noted that medium significant correlations were obtained for L1S on E10 ($Rho = .416^{**}$, $n = 45$, $p < .01$) and E20 ($Rho = .369^*$, $n = 47$, $p < .05$). These findings are parallel to those obtained for overall lexical gains in both groups. In relation to the new words scores, in line with the previous vocabulary outcomes, the analyses showed a medium significant correlation for the L1S group ($Rho = .303^*$, $n = 47$, $p < .05$) but none for the L2S one. When analysing the episodes watched with no subtitles, there was a medium significant correlation for the L1S group on E20 ($Rho = .401^{**}$, $n = 47$, $p = < .005$). Again, these results indicate that aptitude played a favourable role in new word recognition (also in one of the episodes without subtitles) but only for the L1S participants.

In conclusion, comprehension of the TV series was found to be quite strongly related to the learners' aptitude in both groups. On the other hand, the correlation performed

between aptitude and lexical recognition scores was significant (albeit not as strong as that of comprehension) for the L1S group but not for L2S.

7.2.2. Proficiency variables in relation to comprehension and vocabulary scores: correlations

7.2.2.1. L1 / L2 reading speed

Spearman Rho product-moment correlations were conducted between the L1 / L2 RS test results and the comprehension and vocabulary mean scores from the 18 *Curious George* episode-based comprehension and vocabulary tests. The scores from E10 and E20 were not included in the analyses, as these two episodes were watched without subtitles. It should be noted that in order to perform more accurate analyses, participants not obtaining at least 75% correct answers in the RS test were not included in the analysis. That is, our RS score implies that, consequently, students obtained 75% or 100% in the comprehension questions from the L1 / L2 RS tests students took.

In relation to comprehension, the analyses revealed a medium significant correlation between comprehension and the L1 RS of the L1S group ($Rho = .371^*$, $n = 47$, $p < .05$). However, there was not a significant correlation between comprehension and the L2 RS for the L2S group.

In relation to vocabulary, the analyses revealed no significant correlations in any group, indicating that the RS of the L1S and L2S participants was not associated with

their written lexical recognition. In addition, to investigate the relationship between known word identification and RS from the L1S and L2S groups, Spearman Rho product-moment correlations were also conducted: no significant correlations were found in any groups. Finally, the possible connection between the L1 / L2 RS of the participants and their recognition of new words was also explored with Spearman Rho product-moment correlations. The results obtained were in line with the previous vocabulary analyses performed and no significant correlations were found for any group.

7.2.2.2. L2 vocabulary size

Spearman Rho product-moment correlations were conducted between the results of the VST test and those obtained in the episode-based comprehension and vocabulary tests. Apart from giving an indication of the receptive VS of the students, this test was used to make sure there were no significant differences in terms of VS between the groups in the study, which was confirmed (see 'Participants' section). In addition, it should be noted that the results were correlated with the English class grade of the participants at the .01 level ($r = .449^{**}$, $n = 95$, $p = .000$), which indicated that this test was a very good indicator of proficiency.

The analyses for L2 VS revealed a large significant correlation in the L1S group ($Rho = .570^{**}$, $n = 47$, $p < .01$) and a medium one for the L2S group ($Rho = .474^{**}$, $n = 45$, $p < .01$). Furthermore, the outcomes from the analyses of the episodes without subtitles showed significant correlations for the L1S group in E10 ($Rho = .501^{**}$, $n = 45$, $p < .01$) and E20 ($Rho = .418^{**}$, $n = 47$, $p < .01$). These results indicated that previous

lexical knowledge was a relevant factor for understanding the TV series in both groups, independently of whether the subtitles were in the L1 or the L2.

The findings for the correlations performed between the VST results and the total vocabulary scores indicated a medium significant correlation for the L1S group ($Rho = .445^{**}$, $n = 47$, $p < .01$). It should be noted that the correlations performed for the L1S group with the non-subtitled episodes were also significant for E10 ($Rho = .361^*$, $n = 45$, $p < .05$) and E20 ($Rho = .347^*$, $n = 47$, $p < .05$). On the other hand, the results for the L2S group revealed no significant correlations. This shows that students with larger VSs learned more vocabulary than those with small VSs, but only in the L1S group.

The analyses conducted with the known words variable revealed a medium significant correlation for the L1S group ($Rho = .428^{**}$, $n = 47$, $p < .01$), whereas no significant outcomes were obtained for the L2S group. Furthermore, the L1S outcomes showed a significant correlation for E10 ($Rho = .349^*$, $n = 45$, $p < .05$). The analyses for the L1S group with the new words scores showed a medium significant correlation ($Rho = .372^{**}$, $n = 47$, $p < .01$), but there were no significant outcomes for the L2S group. These findings mirror those in the previous lexical analyses, which show that the L1S learners' VS scores were strongly related to their written vocabulary recognition results in the study, but this is not the case for the learners in the L2S group.

7.2.2.3. English class grade

Regarding English class grade and comprehension, the analyses showed quite strong correlations for both groups: L1S ($Rho = .507^{**}$, $n = 47$, $p < .01$) and L2S ($Rho = .754^{**}$,

$n = 47, p < .01$). The tests conducted with the scores from the episodes which included no written input also revealed medium significant correlations for the L1S groups in E20 ($Rho = .480^{**}, n = 47, p < .01$) and L2S (E20 ($r = .396^{**}, n = 45, p < .01$). Therefore, when watching TV episodes, either with L1, L2 or no subtitles, participants' English proficiency was a significant factor in understanding what was happening.

Regarding vocabulary scores, a low significant correlation was found for the L1S group ($Rho = .297^*, n = 47, p < .05$), whereas no significant correlations were obtained for the L2S students. In the analyses performed with the scores from the episodes viewed without subtitles, medium and large significant correlations were obtained for L1S in E10 ($Rho = .329^*, n = 45, p < .05$) and E20 ($Rho = .419^{**}, n = 47, p < .01$) and for L2S in E20 ($Rho = .338^*, n = 45, p < .05$). These analyses indicated that for L1S learners, their English class grade positively influenced their vocabulary identification after viewing the TV episodes, but this was not so for L2S students.

A medium significant correlation was also found between English grades and known words identification in the L1S group ($Rho = .350^*, n = 47, p < .05$). Moreover, the analyses performed with the scores for E10 and E20 showed significant correlations as well: E10 ($Rho = .445^{**}, n = 45, p < .01$) and E20 ($Rho = .322^*, n = 47, p < .05$). These outcomes indicate that, as above, English proficiency level was significant for the identification of known words for L1S learners but not for the L2S group.

Further analyses with the new words mean scores revealed no significant outcomes for any group. However, the analyses conducted with the results from the episodes with no subtitles showed a medium significant correlation for L1S on E20 ($Rho = .339^*,$

$n = 47$, $p < .05$). These outcomes evidenced that English class grade was not determinant for learners when learning new word forms from L1 and L2 subtitled episodes.

7.2.3. Proficiency variables in relation to comprehension and vocabulary scores: regression analyses

Standard regression analyses were also conducted for both groups in order to assess to what extent the results of the episode-based comprehension and vocabulary tests could be influenced by the variables under study. There were no collinearity problems between the variables, except for VST and English class grade. Therefore, the English class grade (which was not measured in the study but provided by the school) was not included in the analysis. Table 7.11 presents the results of the standard regression analyses for L1S and L2S and the corresponding ANOVA Tables can be found below (Table 7.12, Table 7.13 and Table 7.14).

Table 7.11 – Standard regression analyses for L1S and L2S.
Predictor variables: LLAMA_B, RS L1 / L2 and VS.
Dependent variables: comprehension and vocabulary scores.

L1S				
Predictor / independent variable	Dependent variable	R Square	Adjusted R Square	Std. Error of the Estimate
LLAMA_B	Mean	.260	.244	1.42
Reading Speed (RS) L1	Comprehension	.060	.028	1.61
Vocabulary Size (VS)	all episodes	.258	.242	1.42
LLAMA_B	Mean	.120	.100	.828
Reading Speed (RS) L2	Vocabulary all	.028	-.004	.875
Vocabulary Size (VS)	episodes	.103	.084	.836
L2S				
Predictor / independent variable	Dependent variable	R Square	Adjusted R Square	Std. Error of the Estimate
LLAMA_B	Mean	.266	.249	1.22
Reading Speed (RS) L2	Comprehension	.003	-.033	1.43
Vocabulary Size (VS)	all episodes	.128	.108	1.33
LLAMA_B	Mean	.024	.001	.71
Reading Speed (RS) L2	Vocabulary all	.015	-.021	.72
Vocabulary Size (VS)	episodes	.000	-.023	.716

As can be observed, LLAMA_B explains 24.4% of the variance in L1S learners' comprehension scores and 24.9% for those of L2S learners. This is the most significant variable influencing the participants' comprehension scores from both groups. It also explains 10% of the vocabulary scores in the L1S group. Another variable that influences comprehension scores is VS, especially for the L1S group (24.2%). It also accounts for a 10% of the variance in the case of the L2S group. However, it explains just 8% of the vocabulary scores in the L1S group.

**Table 7.12 – ANOVA results. Predictor variable: LLAMA_B.
Dependent variables: comprehension and vocabulary scores.**

Predictor variable: LLAMA_B scores. Dependent variable: comprehension scores					
L1S	Sum of squares	Df	Mean Square	F	Sig.
Regression	31.869	1	31.869	15.839	.000
Residual	90.542	45	2.012		
Total	122.411	46			
L2S	Sum of squares	df	Mean Square	F	Sig.
Regression	22.719	1	22.719	15.240	.000
Residual	62.611	42	1.491		
Total	85.331	43			
Predictor variable: LLAMA_B scores. Dependent variable: vocabulary scores					
L1S	Sum of squares	df	Mean Square	F	Sig.
Regression	4.207	1	4.207	6.129	.017
Residual	30.885	45	.686		
Total	35.092	46			
L2S	Sum of squares	df	Mean Square	F	Sig.
Regression	.521	1	.521	1040	.314
Residual	21.042	42	.501		
Total	21.563	43			

**Table 7.13 – ANOVA results. Predictor variable: RS (L1RS for L1S and L2RS for L2S).
Dependent variables: comprehension and vocabulary scores.**

Predictor variable: RS Dependent variable: comprehension scores					
L1S	Sum of squares	df	Mean Square	F	Sig.
Regression	4.935	1	4.935	1.909	.177
Residual	77.560	30	2.585		
Total	82.494	31			
L2S	Sum of squares	df	Mean Square	F	Sig.
Regression	.153	1	.153	.075	.787
Residual	57.396	28	2050		
Total	57.549	29			
Predictor variable: RS Dependent variable: vocabulary scores					
L1S	Sum of squares	df	Mean Square	F	Sig.
Regression	.666	1	.666	.869	.359
Residual	22.983	30	.766		
Total	23.649	31			
L2S	Sum of squares	df	Mean Square	F	Sig.
Regression	.212	1	.212	.414	.525
Residual	14.331	28	.512		
Total	14.542	29			

**Table 7.14 – ANOVA results. Predictor variable: VS.
Dependent variables: comprehension and vocabulary scores.**

Predictor variable: VS scores. Dependent variable: comprehension scores					
L1S	Sum of squares	df	Mean Square	F	Sig.
Regression	31.593	1	31.593	15.654	.000
Residual	90.818	45	2.018		
Total	122.411	46			
L2S	Sum of squares	df	Mean Square	F	Sig.
Regression	11.189	1	11.189	6.320	.016
Residual	76.127	43	1.770		
Total	87.315	44			
Predictor variable: VST scores. Dependent variable: vocabulary scores					
L1S	Sum of squares	df	Mean Square	F	Sig.
Regression	3.613	1	3.631	5.194	.027
Residual	31.461	45	.699		
Total	35.092	46			
L2S	Sum of squares	df	Mean Square	F	Sig.
Regression	.000	1	.000	.000	.996
Residual	22.064	43	.513		
Total	2.064	44			

7.2.4. Summary of Findings

In relation to our second RQ, language aptitude was found to have the largest effects on the results of both groups regarding comprehension and also regarding vocabulary scores for the L1S group (but not for the L2S group). It should be noted that the analyses performed with the L1S total vocabulary and known words scores showed correlations for the non-subtitled episodes (E10 and E20) too, whereas the outcomes from the new words scores showed one medium significant correlation for E20. The results of the regression analyses conducted between this ID and comprehension for both groups confirmed that aptitude strongly influenced the scores (up to 24.4% for L1S and 24.9% for L2S). It influenced vocabulary scores up to a 10%, but just in the L1S group.

The relationship of the participants' L1 / L2 RS with their comprehension and vocabulary mean scores was examined. Comprehension was found to be minimally related with the L1 RS of the learners, although just with a very moderate correlation, and the variable was not significant in the regression analysis. No effects of L2 RS were observed on any variable for the L2S group.

VS was found to have a large significant effect on comprehension scores for the L1S group (it explained up to 24.2% of the variance) and for the L2S group (with a 10.8%). L2 VS also had an effect on vocabulary learning in the L1S group (there was a significant correlation between the two variables and it explained 8.41% of the word recognition scores). No relationship was found in the case of the L2S group.

Finally, the English class grade variable was found to be strongly related with the comprehension mean scores and the analyses showed large significant outcomes for both groups. However, they were just correlated to vocabulary scores in the L1S group (low correlation). Overall, English class grade gave similar results to those of VS (and, as the two variables were correlated, class grade was not entered into the regression).

7.3. RQ3: Extensive viewing and viewing behaviour

In this section we present the results for the third RQ of the study, examining and comparing the outcomes of the video-watching questionnaire, which was answered before starting the treatment (T1) and at the end (T2) (after watching E19 from the *Curious George* animated TV series). The questionnaire asked the participants from both groups about their watching and reading behaviour while viewing the L1 / L2-subtitled *Curious George* TV episodes.

As mentioned in the instruments section, the first part of the questionnaire consisted of three main questions (in the form of Likert scales and MC types). The results from each question at the two testing times and the McNemar's analysis are presented below. It should be noted that the total amount of answers for each question might not match with the exact number of participants in the study in some cases, as some were not present on the day that the questionnaire was administered (or because a particular question was left blank).

7.3.1. Self-reported level of understanding

Question 1 (Q1) asked participants if they had understood what had happened in the video. Six possible options were provided on a Likert scale. Results showed that very few participants in both groups chose 'Not at all' (0) and 'Not much' (1). Most L1S students selected answer 3 'Quite', 4 'A lot' or 5 'All of it' at T1. Furthermore, options 4 and 5 increased considerably at T2.

On the other hand, the L2S group's answers were distributed differently at T1: most of the L2S participants chose options 2 'A little', 3 'Quite' and 4 'A lot'. Option 5 'All of it' showed a low percentage (11.35%), although the results at T2 were different and nearly half of the L2S learners selected this option (47.73%). Table 7.15 shows the raw numbers and percentages of students from both groups for each option included in this question.

Table 7.15 – Raw numbers and percentages for ‘self-reported level of understanding’ at T1 / T2.

Date	T1				T2			
Group	L1S		L2S		L1S		L2S	
0-Not at all	1	2.13%	0	0%	0	0%	0	0%
1-Not much	0	0%	2	4.55%	1	2.13%	0	0%
2-A little	4	8.51%	10	22.73%	1	2.13%	1	2.27%
3-Quite	13	27.66%	17	38.64%	4	8.51%	9	20.45%
4-A lot	13	27.66%	10	22.73%	15	31.91%	13	29.55%
5-All of it	16	34.04%	5	11.35%	26	55.32%	21	47.73%

The results of the McNemar’s test show a significant increase for both groups between the first and the second time the questionnaire was administered. Both groups said they were understanding more at T2. It should be noted that, in order to be able to conduct the test, the categories ‘Not at all’, ‘Not much’ and ‘A little’ were grouped into one category as few participants selected these options. Table 7.16 shows the descriptive statistics at T1 and T2 and the McNemar’s results. Figure 7.3 presents the percentages of each option for L1S and L2S at both testing times.

Table 7.16 – McNemar’s results for ‘self-reported level of understanding’.

Q1 – Did you understand what happened in the video?		T1	T2	p -global value	p -group value
L1S	Not at all / Not much / A little	5 (10.6 %)	2 (4.3 %)	0.001	0.038
	Quite	13 (27.7 %)	4 (8.5 %)		
	A lot	13 (27.7 %)	15 (31.9 %)		
	All of it	16 (34.0 %)	26 (55.3 %)		
L2S	Not at all / Not much / A little	13 (28.9 %)	2 (4.4 %)	0.001	0.001
	Quite	17 (37.8 %)	9 (20.0 %)		
	A lot	10 (22.2 %)	13 (28.9 %)		
	All of it	5 (11.1 %)	21 (46.7 %)		

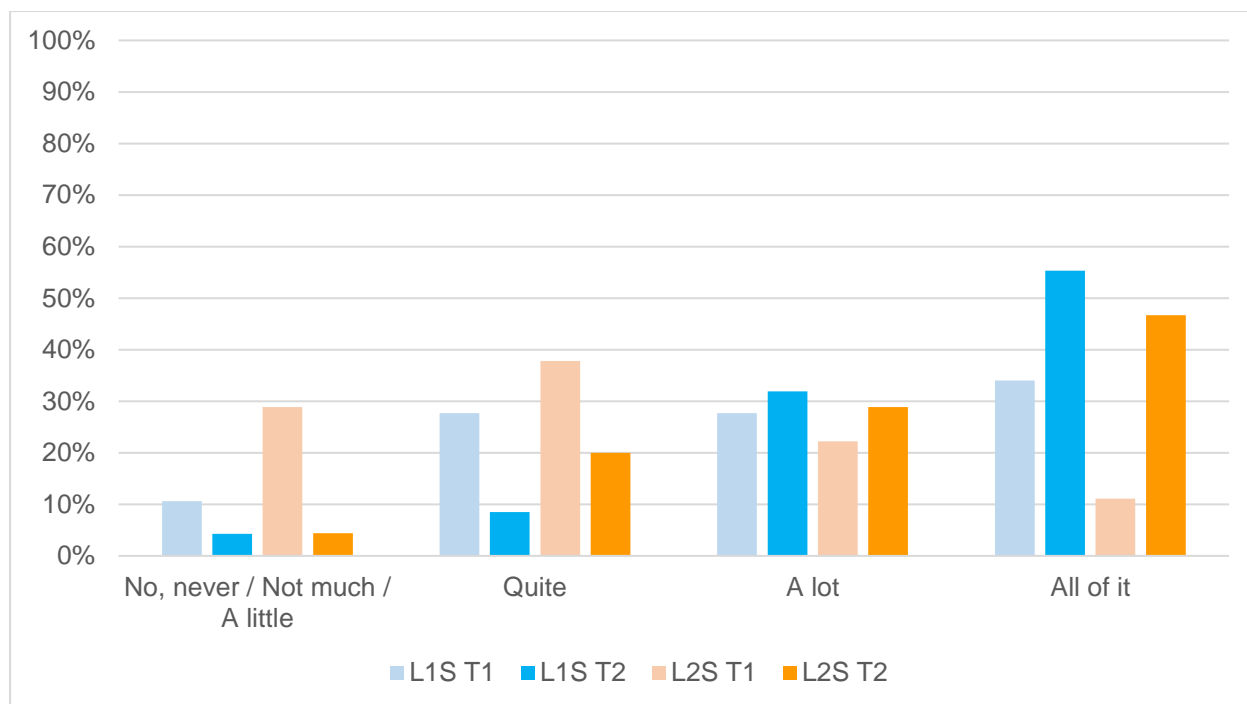


Figure 7.3 – Descriptive data for ‘self-reported level of understanding’ at T1 / T2 in both groups.

7.3.2. Viewing behaviour

Question 2 (Q2) inquired about the participants’ TV viewing behaviour for the *Curious George* videos. It was presented in the form of MC and introduced by a question asking students about how they had just viewed the subtitled episode: (1) ‘Listening and reading the subtitles at the same time’, (2) ‘Listening and reading a few times’, (3) ‘Only listening and I never read the subtitles’ and (4) ‘Only reading and I did not pay attention to the audio’. Results indicated that most participants from both groups selected answers 1 or 2 at the two data collection times. In addition, the number of students who gave the same answer at both testing times was very similar. The low percentages obtained in options 3 and 4 in both groups show that the majority of the students were reading the subtitles and listening to the soundtrack at the same time

as they were watching the videos since the beginning of the intervention. Tables 7.17 and 7.18 present the descriptive results for this question.

Table 7.17 – Raw numbers and percentages for ‘viewing behaviour’ (L1S).

Date	T1		T2	
Group	L1S			
Listening and reading the subtitles at the same time	24	51.1%	29	61.7%
Listening and sometimes reading	18	38.3%	16	34.04%
Listening only. Never read the subtitles	0	0%	1	2.13%
Reading only. Not paying attention to the audio	5	10.6%	1	2.13%

Table 7.18 – Raw numbers and percentages for ‘viewing behaviour’ (L2S).

Date	T1		T2	
Group	L2S			
Listening and reading the subtitles at the same time	14	31.8%	21	47.7%
Listening and sometimes reading	27	61.4%	21	47.7%
Listening only. Never read the subtitles	3	6.8%	2	4.6%
Reading only. Not paying attention to the audio	0	0%	0	0%

The McNemar’s test results for this question show no significant differences between T1 and T2 for either group. These findings confirmed what the descriptive data was already showing (no significant changes from the beginning to the end of the intervention). In order to be able to conduct the test, the options ‘Listening only. Never read the subtitles’ and ‘Reading only. Not paying attention to the audio’ were grouped into one category as they were selected by few participants. Table 7.19 presents the descriptives and the results of the McNemar’s test for ‘viewing behaviour’. Figure 7.4 shows the descriptive data in percentages for each option.

Table 7.19 – McNemar’s results for ‘viewing behaviour’.

Q2 – This video had subtitles: Reflect on how you watched the video...		T1	T2	p-global value	p-group value
L1S	Listening and reading the subtitles at the same time.	24 (51.1 %)	29 (61.7 %)	0.166	0.215
	Listening and sometimes reading.	18 (38.3 %)	16 (34.0 %)		
	Listening only. Never read the subtitles. / Reading only. Not paying attention to the audio.	5 (10.6 %)	2 (4.3 %)		
L2S	Listening and reading the subtitles at the same time.	14 (31.8 %)	21 (47.7 %)		0.392
	Listening and reading a few times.	27 (61.4 %)	21 (47.7 %)		
	Listening only. Never read the subtitles. / Reading only. Not paying attention to the audio.	3 (6.8 %)	2 (4.6 %)		

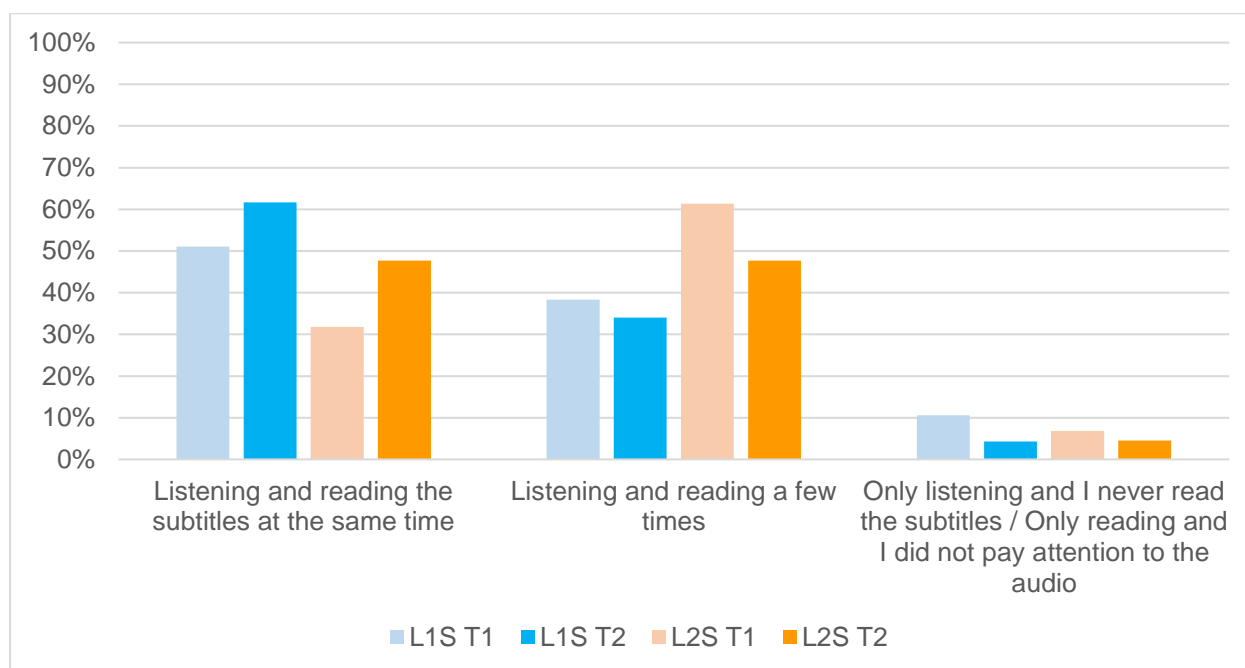


Figure 7.4 – Descriptive data for ‘viewing behaviour’ at T1 / T2 in both groups.

7.3.3. Reading and learning from subtitles

Question 3 (Q3) consisted of three sub-questions. These questions asked: (Q3a) if the participants had had the necessary time to read all the subtitles, (Q3b) why they felt the need to read them and (Q3c) if they thought they had learnt new words while

reading them. Each of these sub-questions was followed by several alternatives, presented through Likert scales (Q3a, Q3c) or MC question types (Q3b). Students had to select only one option in the Likert scales but they were given the possibility to choose more than one answer in the MC question. McNemar's tests were conducted for each sub-question included in Q3.

7.3.3.1. Time taken to read subtitles

In Q3a, the students needed to answer whether they had had enough time to read the subtitles through a Likert scale with six options. Many L1S students marked that they could read the subtitles 'very often' (34.04%) or 'always' (25.53%) at T1. At the second data collection time, these same answers were usually selected by this group, with a slightly higher percentage for 'very often' (38.30%) but lower for 'always' (23.40%). On the other hand, many L2S learners expressed that they required more time to read the L2S by selecting 'a little' (38.64%), whereas the second most popular option for this group was 'very often' (18.18% at T1 and 20% at T2). It should be noted that these options were less frequently selected than 'very often' (the most popular answer) at both testing times (38.64% and 42.23%, respectively). The number of students and percentages for each option in Q3a are presented in Table 7.20.

Table 7.20 – Raw numbers and percentages for 'time taken to read subtitles'.

Date	T1				T2			
	L1S (n= 47)		L2S (n= 45)		L1S (n= 47)		L2S (n= 45)	
0-No, never	0	0%	4	9%	0	0%	2	4.3%
1-Not much	4	8.5%	5	11.1%	3	6.4%	4	8.9%
2-A little	5	10.6 %	19	42.3%	3	6.4%	7	15.6%
3-Quite often	9	19.1 %	2	4.4%	10	21.3%	11	24.4%
4-Very often	18	38.4 %	9	20.0%	13	27.6%	16	35.7%
5-Always	11	23.4 %	6	13.3%	18	38.3%	5	11.1%

In order to be able to conduct the McNemar’s tests, we needed to group together the two categories ‘No, never’ and ‘Not much’, as some of these categories were not selected by any of the participants. The McNemar’s test conducted to compare the outcomes of each group at the beginning and at the end of the intervention indicates that no significant differences were found for either group, even if results for the L2S group nearly reach significance (see Table 7.21 and Figure 7.5).

Table 7.21 – McNemar’s results for ‘time taken to read subtitles’.

Q3a – Time to read subtitles		T1	T2	ρ -global value	ρ -group value
L1S	No, never / Not much	4 (8.5 %)	3 (6.4 %)	0.070	0.276
	A little	5 (10.6 %)	3 (6.4 %)		
	Quite often	9 (19.1 %)	10 (21.3 %)		
	Very often	18 (38.4 %)	13 (27.6 %)		
	Always	11 (23.4 %)	18 (38.3 %)		
L2S	No, never / Not much	9 (20.0 %)	6 (13.3 %)	0.057	0.057
	A little	19 (42.3 %)	7 (15.6 %)		
	Quite often	2 (4.4 %)	11 (24.4 %)		
	Very often	9 (20.0 %)	16 (35.7 %)		
	Always	6 (13.3 %)	5 (11.1 %)		

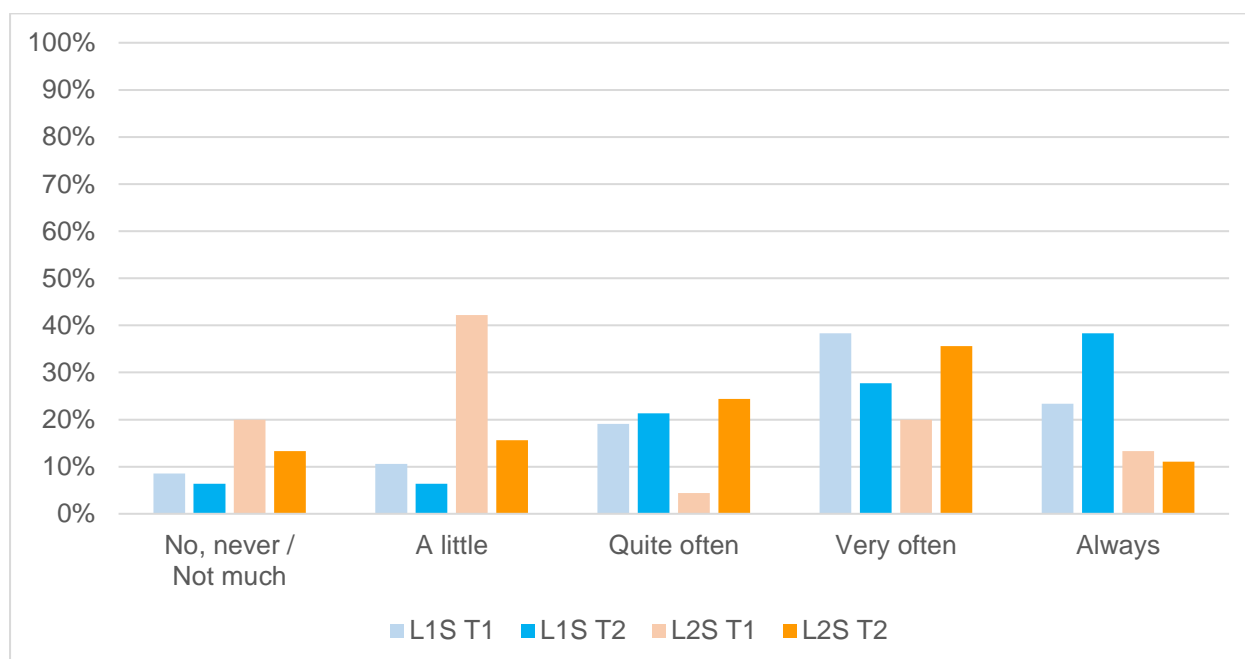


Figure 7.5 – Descriptive data for ‘time taken to read subtitles’ at T1 / T2 in both groups.

7.3.3.2. Reasons for reading subtitles

Q3b inquired about the participants' reasons for reading the subtitles. In this question more than one answer was accepted (some of the students chose two of them). The possible options, listed in the form of a MC question type, included: (1) Because if I don't read them, I don't understand what is being said, (2) Because they help me to understand what is happening, (3) Because they appear on the screen and I cannot avoid reading them and (4) Because I think I learn more. The results from both groups showed that option 2 was the most popular, obtaining similar percentages at both data collection times. Tables 7.22 and 7.23 show the number of students from each group that selected each option for this question.

Table 7.22 – Raw numbers and percentages for 'reasons for reading subtitles' (L1S).

Date	T1		T2	
Group	L1S			
(1) Because if I don't read them, I don't understand what is being said.	11	23.4%	7	14.9%
(2) Because they help me to understand what is happening.	29	61.7%	30	63.9%
(3) Because they appear on the screen and I cannot avoid reading them.	1	2.1%	9	19.1%
(4) Because I think I learn more.	22	46.8%	24	51.1%

Table 7.23 – Raw numbers and percentages for 'reasons for reading subtitles' (L2S).

Date	T1		T2	
Group	L2S			
(1) Because if I don't read them, I don't understand what is being said.	8	17.8%	7	15.6%
(2) Because they help me to understand what is happening.	29	64.4%	19	42.2%
(3) Because they appear on the screen and I cannot avoid reading them.	9	20.0%	8	17.8%
(4) Because I think I learn more.	17	37.8%	25	55.6%

A McNemar's tests was also conducted in order to compare the answers at both testing times. The results show no significant differences for options 1, 2 and 4, whereas option 3 was chosen significantly more often at T2 in the L1S group. Table

7.24 presents the results of the McNemar’s tests conducted and Figure 7.6 shows the descriptive statistics for ‘Reasons for Reading Subtitles’ at T1 and T2 for L1S and L2S.

Table 7.24 – McNemar’s results on ‘reasons for reading subtitles’.

0= option not selected, 1=option selected

Q3b – Reasons for reading subtitles		T1		T2	ρ -global value	ρ -group value
L1S	Because if I don’t read them, I don’t understand what is being said.	0	36 (76.6 %)	40 (85.1 %)	0,332	0.289
		1	11 (23.4 %)	7 (14.9 %)		
L2S		0	37 (82.2 %)	38 (84.4 %)		1.000
		1	8 (17.8 %)	7 (15.6 %)		
L1S	Because they help me to understand what is happening.	0	18 (38.3 %)	17 (36.2 %)	1,000	1.000
		1	29 (61.7 %)	30 (63.9 %)		
L2S		0	16 (35.6 %)	10 (22.2 %)		1.000
		1	29 (64.4 %)	19 (42.2 %)		
L1S	Because they appear on the screen and I cannot avoid reading them.	0	46 (97.9 %)	38 (80.9 %)	0,189	0.021
		1	1 (2.1 %)	9 (19.1 %)		
L2S		0	36 (80.0 %)	37 (82.2%)		1.000
		1	9 (20.0 %)	8 (17.8 %)		
L1S	Because I think I learn more	0	25 (53.24 %)	23 (48.9 %)	0,132	0.824
		1	22 (46.8 %)	24 (51.1 %)		
L2S		0	28 (62.2 %)	20 (44.4%)		0.077
		1	17 (37.8 %)	25 (55.6%)		

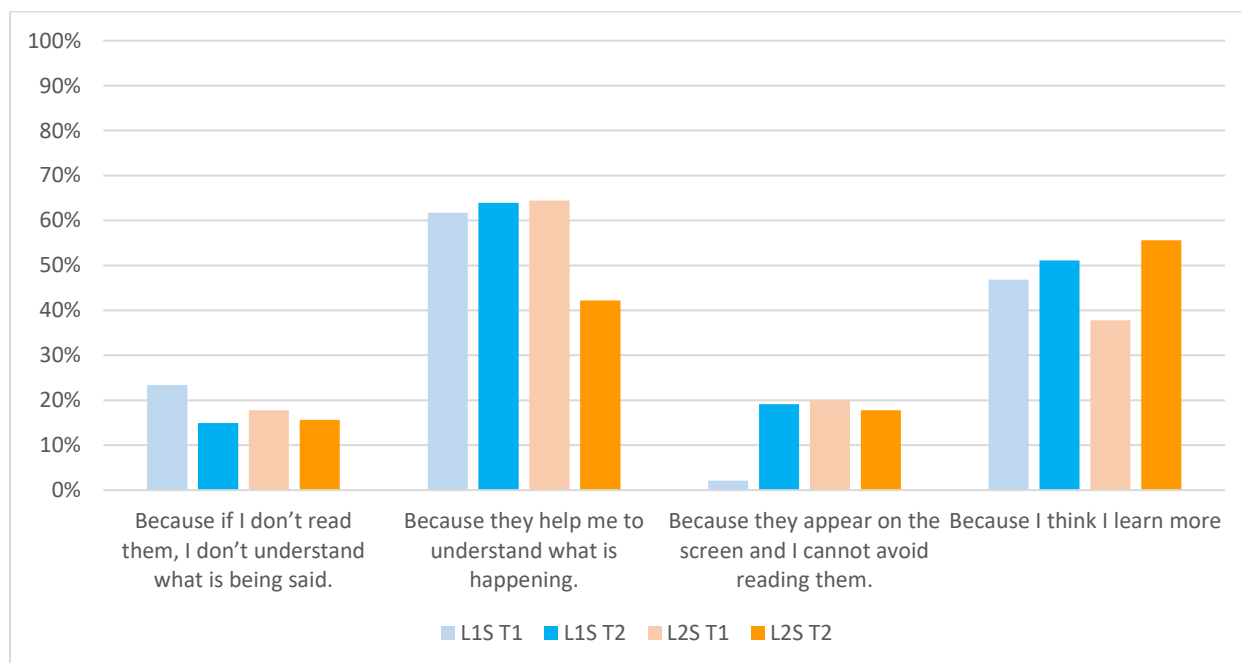


Figure 7.6 – Descriptive data for ‘reasons for reading subtitles’ at T1 and T2 in both groups.

7.3.3.3. Subtitles and vocabulary learning

Q3c asked participants whether they thought that subtitles helped them in learning new words. The learners could choose between three possible answers and, to investigate which were the most preferred replies, the results from the L1S and L2S students at the two data collection times were compared. Learners from both groups perceived an improvement in their word acquisition since the beginning of the study. The most widely chosen option at both testing times was number 2, which implied that the students felt they had learnt some new words after watching the subtitled episodes. However, the L2S percentages were higher, especially for answer 1 (4.26% to 25%), indicating that many L2S students realised that they had learnt a lot of new words throughout the whole treatment.

Results from the McNemar's test conducted with the responses from Q3c at the two testing times showed significant differences for the L2S group between T1 and T2. On the other hand, no significant differences were observed in the answers of the L1S group. Table 7.25 reports the descriptives in raw numbers and percentages at T1 and T2 and the results of the McNemar's test for Q3c.

Table 7.25 – Raw numbers and percentages for 'subtitles and vocabulary learning'.

Subtitles and vocabulary learning		T1	T2	p -global value	p -group value
L1S	Yes, a lot	2 (4.3 %)	4 (8.7 %)	0.001	0.144
	Yes, some	27 (58.7 %)	32 (69.6 %)		
	No, none	17 (37.0 %)	10 (21.7 %)		
L2S	Yes, a lot	1 (2.4 %)	9 (22.0 %)		0.002
	Yes, some	35 (85.4 %)	31 (75.6 %)		
	No, none	5 (12.2 %)	1 (2.4 %)		

Therefore, participants' responses indicate that the presence of either L1S or L2S in the video episodes implied lexical learning for the participants. However, the L2S group perceives the help of the L2 subtitles more significantly, while the L1S learners report 'some help' from the subtitles or even none, and no significant differences are observed between data collection times.

7.3.4. Summary of findings

RQ3 analysed the participants' responses from the video-watching questionnaire answered twice: once at the beginning and once at the end of the treatment. The aim of the analysis was to find out if the repeated viewing of several L1 or L2 subtitled videos affected their way of watching animated TV series. Overall, the outcomes suggest that the presence of L1 or L2 subtitles in the TV series was very helpful for the young low-proficiency L2 learners in the study (both groups indicated the same reasons for reading them and percentages were similar).

Results on self-reported level of understanding showed significant differences between the first and the second testing time for both groups, both reporting higher levels of understanding at T2. Regarding Q2, no differences were observed for their viewing behaviour in any group: they tended to pay attention to subtitles and audio at the same time.

In relation to reading and learning from subtitles, in Q3a, inquiring about whether learners had time to read the subtitles, most L1S students answered in the same way at both testing times (most had enough time 'always' or 'often'). In the L2S group, there was significant change during the treatment. In the end, those who had enough time

to follow the subtitles had significantly increased. On the other hand, answers to Q3b about reasons for reading the subtitles were basically the same for each group at T1 and T2. The only notable difference was that at T2 a greater significant number of L1S group members affirmed that they had been reading the subtitles because 'they were on the screen'. Finally, Q3c asked participants if they thought they had learnt new words with the L1 or L2 subtitles. The McNemar's results showed significant differences for the L2S learners, who by the end of the treatment considered the subtitles more helpful.

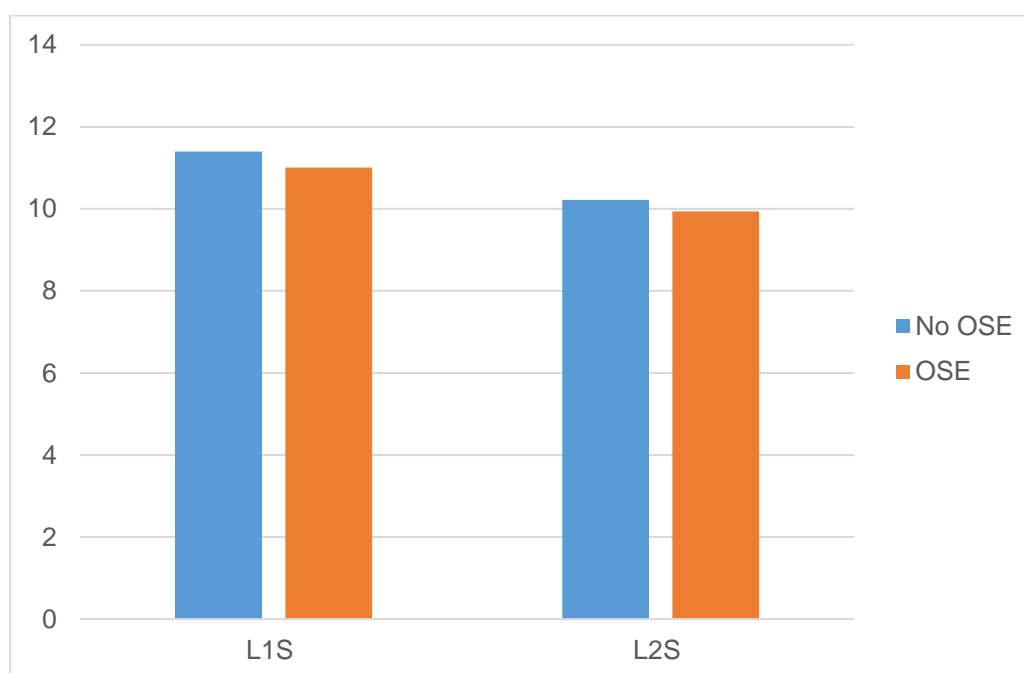
7.4. RQ4: Effects of OSE on L2 comprehension and vocabulary learning from extensive viewing

In this section, we present the results for RQ4 of the study, i.e., "Does out-of-school-exposure (OSE) to multimodal input have an effect on episode comprehension and written vocabulary recognition scores when watching subtitled TV series in class?". We wanted to find out if learners' habits related to multimodal input (e.g., watching TV series, cartoons or films in English at home with L1S, L2S or without subtitles) had any influence on the results we obtained for comprehension and vocabulary. Table 7.26 presents the groups, the mean and standard deviation, as well as the results of the *t*-tests conducted with the comprehension and vocabulary scores.

Table 7.26 – T-tests results comparing students with and without OSE in terms of their comprehension and vocabulary scores.

Group	Variable	G1	Mean (sd)	G2	Mean (sd)	Results
L1S	Comprehension mean all episodes	n = 27	11.40 (1.44)	n = 20	11.01 (1.89)	$t(45) = -.805, p = .425$
	Vocabulary mean all episodes		4.90 (.79)		4.62 (.947)	$t(45) = -.111, p = .272$
L2S	Comprehension mean all episodes	n = 25	10.22 (1.37)	n = 20	9.94 (1.48)	$t(43) = -.653, p = .517$
	Vocabulary mean all episodes		5.26 (.84)		5.23 (.52)	$t(43) = -.128, p = .899$

The results obtained from the *t*-tests conducted showed playing videogames and / or watching films, cartoons or TV series in English at home with L1S, L2S or without subtitles did not influence the comprehension results of the learners in the L1S [$t(45) = -.805, p = .425$] or the L2S group [$t(43) = -.653, p = .517$] as no significant differences were found. Neither did the analysis give any significant results for L2 word-form recognition when comparing students who received OSE to multimodal input from those who did not (see Figures 7.7 and 7.8 for the comparison of the comprehension and vocabulary means from both groups).

**Figure 7.7 – Mean comprehension scores for the OSE and No OSE groups.**

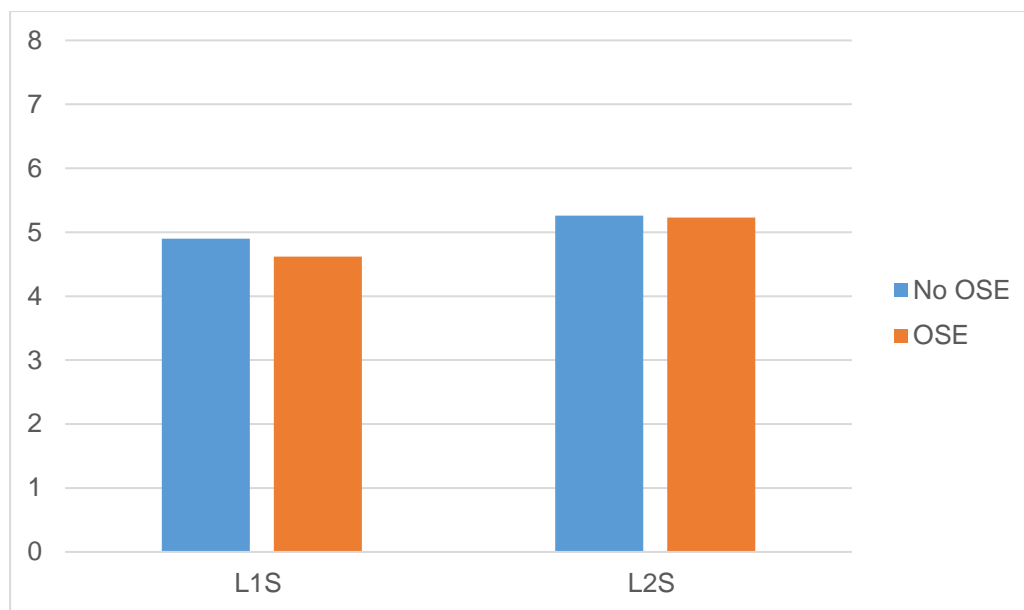


Figure 7.8 – Mean vocabulary scores for the OSE and No OSE groups.

CHAPTER 8 - DISCUSSION

In this chapter, results will be interpreted following the RQs presented in Chapter 6 and related to previous research we have presented in our theoretical framework. This study expands on earlier research in several ways. To our knowledge, it is one of the few studies available that longitudinally (over 5 months) explores L2 comprehension and word-form recognition when watching several episodes of the same animated TV series in young low-proficiency L2 learners (also making a distinction between new and known vocabulary appearing in the TV series). Furthermore, the study also looks at the possible effects that aptitude and proficiency variables (such as VS, RS and English class grade) may have on the L2 comprehension and vocabulary learning from extensive viewing. In addition, it explores whether and how this repeated viewing may have influenced the way young learners watched subtitled episodes after the five months. Finally, in an attempt to capture any possible previous contact time with multimodal input outside school, we enquired about whether exposure to audiovisual materials at home does actually make any difference to the comprehension and vocabulary scores that learners obtained in the treatment.

8.1. Introduction

Based on the existing evidence from several studies (Danan, 2004; Vanderplank, 2010, 2013; Montero-Pérez et al., 2013), we assumed that using multimodal input in English classes could be successful to enhance the L2 proficiency of EFL learners. The use of TV series for EFL learning in schools across Catalonia is not common, but we thought that it would be very engaging for the participants, who can watch TV series

/ programmes in the original language in their daily lives nowadays. We also based our decision on previous evidence, claiming that watching audiovisual materials can be a motivating activity as well (e.g., Kuppens, 2010).

Muñoz (2008) also argues that different options for providing our young learners with large amounts of input that complement classroom instruction need to be considered and this is what we tried to do in this study. The animated TV series chosen offers comprehensible input that is challenging for learners at the same time and offers room for improvement (Krashen, 1985). Authentic input tends to be richer than that found in traditional textbooks (Gilmore, 2007) and has been shown to boost L2 learning (Bahrani & Soltani, 2011; Shabani & Zanussi, 2015). Therefore, its use in the form of subtitled TV series should be considered and carefully evaluated.

8.2. Extensive viewing results in the light of theories on learning from multimodal input

We would like to start by commenting on our results in the light of the three theories we have introduced in the Literature Review section: the DCT, the CTML and CLT. The DCT (Paivio, 1986) and the CTML (Mayer, 2009) argue that the interaction of linguistic and visual input will improve learning and that receiving information through verbal and non-verbal channels will result in better comprehension and recall. According to Sweller (1988; 2005), learning will take place if the brain's cognitive capacity is not overloaded, as proposed in CLT.

The comprehension and vocabulary results in our study reveal that participants were learning from input coming in different forms (aural, written and images). L2 learners were able to merge and process the written, aural and visual input they were exposed to, which, in principle, would support the DCT (Paivio, 1986) and the CTML (Mayer, 2009). The fact that there is learning also suggests that learners were not overwhelmed by the input and that, despite a possible extra CL (Sweller, 1988; 2005), they were making sense of the episodes they were exposed to.

The gains in vocabulary obtained by the L2S group in the L2 word-form recognition test indicate that the combination of visual and written input can work effectively to help participants identify L2 word forms, which supports the CTML (Mayer, 2009). When our L2 learners watched the subtitled videos, they received the verbal and pictorial information of the lexical items through the visual (images and written text) and auditory channels (except E10 and E20 which did not include subtitles). When watching the videos, they processed this information using their WM and categorised the input received into verbal or pictorial models. We believe that by employing these cognitive processes, the low-proficiency L2 learners were led to an effective L2 word identification from the multimodal input. Our results are in line with those obtained from adult learners by Rodgers (2013), who suggested that watching several episodes of the same TV series, which combined visual and aural input, aided L2 vocabulary acquisition.

It should also be noted that in E10 and E20, where textual information was not present, no significant differences were found between groups and the scores were quite low in comprehension (even if the mean is a bit higher in E20 than in E10). As a result,

textual support seems to be playing a role in understanding the plot. When we check the vocabulary results, not having the subtitles does not mean that means are lower than in other episodes. It seems that getting information through text was aiding comprehension to a larger extent than vocabulary learning, even if L2S helped learners more in recognising word-forms.

We should also consider the fact that students are learning and that they are understanding the episodes as in other studies such as Sherman (2003), Danan (2004), Vandergrift (2007), Kuppens (2010) and Webb (2010). These considerations do not provide full support to the DCT and CTML, as we do not have a CG that is learning without being exposed to the TV series. Consequently, we cannot ensure that these children are more successful at learning than if they were not exposed to multimodal input and had not followed another approach (e.g., explicit vocabulary teaching).

8.2.1. Cartoons and L2 learning

The results from this study suggest that the selection of the *Curious George* animated TV series was appropriate for the participants. The L2 learning shown in the results reinforces the idea of cartoons as a potential linguistic tool (Curtis, 2015). Results evidence that there is learning on the part of participants, in line with previous research conducted with young low-proficiency L2 learners watching cartoons to enhance L2 vocabulary learning (e.g., Alexiou, 2015; Kokla, 2016), thus hinting at the potential cartoons have for L2 learning in formal environments. However, the treatment in these previous studies did not include subtitles and the participants were younger (4-6 year-

olds) than our students. Even so, the scarce research available on young low-proficiency L2 learners watching subtitled TV series suggests that the presence of written input was most likely to have aided students in their L2 learning (Gesa, 2019). Due to the few studies in the field with young learners, more longitudinal research is called for.

The findings in this study would support previous research suggesting that cartoons include a plain dialogue with constant recurrences at a low rate of speech (Bahrani, 2014), which might have aided the participants of the study in their comprehension. In addition, Koolstra and Beentjes (1999), whose study was performed with primary students, claimed that cartoons might be the best option for enhancing comprehension and vocabulary acquisition for young L2 learners. Eye-tracking studies conducted with cartoons (d'Ydewalle & Vanrensbergen, 1989; d'Ydewalle & Bruycker, 2007; Tragant & Pellicer-Sánchez, 2019) have talked about the appropriateness of these materials for L2 learning and shown that the children were able to process the multimodal input in the same way as adults, even though more time was devoted to the subtitles.

Moreover, Webb and Rodgers (2009b) and Webb (2011; 2015) claim that it would be less challenging for L2 learners if a TV series where the episodes have the same main characters and related content is selected. Following their research, the *Curious George* TV series has recurring characters in all its episodes and the development of the action can only take place in two settings (city or country). In addition, the episodes are not very long (10-11 minutes each), due to the age group that this TV series was originally aimed at. The length was also optimal for the low L2 proficiency of the learners. Young L2 learners are not able to maintain their attention for too long when

exposed to L2 multimodal input and, due to that, the short length of the *Curious George* episodes might have been crucial for them to remain focused throughout the whole episode.

The teacher-researcher also observed that the TV series immediately attracted the attention of the participants and all paid close attention to the video. The exclusive features that Bahrani (2014) claims cartoons contain, such as bright colors, images and an exaggerated intonation may have helped in capturing their attention. Moreover, the imagery included in the animated TV series also provided non-linguistic input to our participants and aided their understanding of the story. The traits of this specific TV genre allowed the students to watch the video episodes in a low affective filter environment (Bahrani & Soltani, 2011). This may also positively affect L2 learning.

Therefore, our research would support Bahrani (2014) and Akcan and Demirhan (2016) in that cartoons should be included in formal L2 learning environments as teaching tools. Furthermore, we also think that the selection of the animated TV series was a key element in our research: as Khan (2015) argues, if the TV series is adequate for the learners, there is a connection between them and the audiovisual materials that facilitates learning and we believe this was the case for our study.

8.3. L2 comprehension and vocabulary learning: L1 vs. L2

We continue our discussion by addressing the first RQ of the study, which concentrates on the findings related with L2 comprehension and written vocabulary recognition of known and new words. We seek to analyse the results by comparing

the L1S and L2S scores on the tests conducted after watching the *Curious George* episodes using quantitative analyses.

In order to answer this question, we examined the results obtained in the 18 episode-based comprehension and vocabulary tests. In addition, we looked at the longitudinal progress after watching all of the twenty TV episodes selected for the study, and we performed L1S - L2S comparisons through different statistical techniques. In the sections to follow, we review the main L2 comprehension and written vocabulary recognition findings from the analyses and we examine how they compare with the relevant empirical literature in our discussion of RQ1.

8.3.1. Effects of subtitled TV viewing on comprehension: L1S vs. L2S

In response to the first part of RQ1, which focuses on L2 comprehension when watching L1 or L2 subtitled videos, the comparison of the results for the L1S and L2S groups revealed significant differences between them in most of the episodes. On the basis of the analysis presented, we can confirm that the mean comprehension scores tended to be favourable to the L1S learners (with the exception of E17). In addition, the significant differences obtained by the L1S learners in 13 out of the 18 episodes indicate that watching the videos with textual support in the L1 enhanced comprehension. This result can be explained by the low L2 proficiency level of the learners: their reading skills were also better in the L1 (as RS results show), so this allowed them to read L1S faster and comprehension was improved. These results are consistent with those in previous research (Koolstra & Beentjes, 1999; Danan, 2004; Lin & Siyanova, 2015). On the other hand, the lower results obtained by the L2S group

suggest that L2S might have been too challenging for the students to help comprehension, due to their low L2 level and still poor reading skills in comparison to the L1. These findings would be in line with those in Muñoz (2017) in that young learners, due to their low proficiency level, require more fixations on L2S than on L1S.

The findings obtained in previous research comparing the effects of L1S and L2S videos also confirm that the presence of L1S imply better results for L2 comprehension (Markham et al., 2001; Markham & Peter, 2003; Bianchi & Ciabattini, 2008; Latifi et al., 2011). However, it should be noted that these studies were conducted with adult intermediate-level learners and only included a very limited number of sessions. On the other hand, the few studies available comparing L1S vs. L2S for L2 comprehension with young L2 learners (Başaran & Köse, 2012; Galimberti & Miralpeix, 2018) did not obtain significant outcomes, i.e., significant differences were not found for comprehension, independently of whether subtitles were provided in the L1 or L2. It should be noted, though, that these were one-off studies, watching one episode that lasted about 18-20 minutes. Therefore, it is recommended that studies use more than one episode to obtain reliable results since, as also happens in the present study, results are episode-dependent. For example, differences were not observed in E17 and in several episodes differences were not significant. However, they were significant in the majority of episodes, showing that L1S enhanced comprehension. The fact that the behaviour is analysed in several episodes provides more valid evidence that it is taking place, as results in one-off studies may be attributable to a variety of factors that may escape the researchers' control.

It is also worth commenting that, even if L1 subtitles favoured comprehension in the case of the L1S group, the presence of both L1S and L2S in our study was helpful in the enhancement of comprehension for both groups (which also confirms the appropriateness of the materials). The proof is the lower score obtained by both groups in E10 in the middle of the treatment, which included no written input (the results were the lowest of the whole treatment so far in both groups). This suggests that the inclusion of subtitles (either in the L1 or L2) increased the participants' understanding of the episodes. This result is in consonance with Garza (1991) and Becker and Sturm (2017), who claim that the added textual support helps L2 learners in deciphering dubious input. These outcomes are also in line with previous research that has compared the effect of watching videos with or without subtitles to enhance comprehension for older learners too (Garza, 1991; Huang & Eskey, 1999; Winke et al., 2010; Etemadi, 2012; Rodgers, 2013; Gowhary et al., 2015; Ebrahimi & Bazaei, 2016; Gesa, 2019). However, our results do not support Taylor (2005), who concluded that L2S did not improve comprehension in adult beginner learners.

Finally, we should also examine the outcomes obtained by both groups in the last episode of the treatment. The mean for E20 (this episode did not include any subtitles) on comprehension was the second lowest (after E10) for L1S but not for the L2S group, as E8, E13 and E15 obtained lower scores. These findings lead us to interpret that the L1S group experienced more problems when subtitles were not present. First of all, they may have been used to their support all the time and relied more on them for comprehension than their peers. That might explain why they particularly 'suffer' in the two episodes where they do not have the subtitles. Secondly, they add further support to the finding commented above that L2S were harder to make sense of by

the learners as it can be seen that even if the E20 score is very low, the scores for 3 other episodes with subtitles is even lower. This suggests that the presence or absence of subtitles was not as important for comprehension in the L2S as it was for the L1S group. The L2S learners were recurrently experiencing more difficulties than their L1S peers to understand the input (with or without subtitles in this case). The fact that the score was not as low in E20 than in E10 for both groups may be explained by a familiarity effect: it was the last episode of the treatment and the learners had accumulated knowledge from the series as well as experience with the tasks they were performing. This may have helped in obtaining a higher result, even if the difference in the mean of E10 and E20 is actually minimal.

In the light of the results, then, we think that overall subtitles were helpful in understanding the TV series (even if it was more so when the subtitles were in the L1). Results suggest that the written input included in the videos aided in bottom-up processes of comprehension: it helped participants in decoding the different parts of the sound stream (students were able to understand the content of the story also through their background knowledge). As pointed out in Pinter (2017), young L2 learners face a big challenge when trying to effectively apply the linguistic and non-linguistic sources involved in both top-down and bottom-up processes, and our learners have found themselves in this situation. The written input probably aided L2 comprehension when the students had to deal with the challenging procedure of understanding the plot instantly through the interaction between bottom-up and top-down processes (Staer, 2009). Previous research has also suggested that, due to their low proficiency level, young learners make more use of bottom-up processes (Goh, 2000). Our learners were certainly taking subtitles into account and using bottom-up

processes to understand the episode, although it is difficult to state to which extent (or whether these were more consistently used than top-down processes).

Nevertheless, as we do not have a group who watched the episodes without any subtitles throughout the treatment, we cannot conclude whether the experience of having the subtitles available (either in the L1 or the L2) was statistically significant in comparison with not having any textual support at all. Even if the inclusion of a third group watching the series without subtitles was carefully considered before the experiment, previous research had already pointed out how hard it can be for students at low levels to understand real multimodal input without textual support and that is why only the options with subtitles were finally considered for the study.

8.3.2. Effects of subtitled TV viewing on word-form recognition: L1S vs. L2S

The results from the second part of RQ1, which focused on written vocabulary recognition after watching L1 or L2 subtitled TV series indicated that watching *Curious George* helped participants to recognise L2 words. The supportive imagery from the episodes might have lowered the lexical demands of the learners and compensated for their low L2 proficiency (Durbahn et al., 2020; Peters & Muñoz, 2020), helping them in picking up new words or in identifying vocabulary they already knew.

However, contrary to the results obtained for comprehension, the L2S group tended to have higher means and when significant differences were found, these were in favour of the L2S group. This could be explained because having written support in the L2 in the form of subtitles facilitated word-form recognition later in the test, and

was more difficult for the L1S group because they had the subtitles in the first language. L1S participants read the subtitles in their L1, which might have implied an extra difficulty in processing the input, even though they also received the input through different channels at the same time. L1S may have hindered the association of phonological information and L2 written forms, impairing a simultaneous processing with L2 soundtrack. Even though the young L2 learners from both groups received the same amount of input, L1 subtitles bring about lower L2 lexical recognition scores (L1S learners were never exposed to the written form of these items and were not able to identify as many as the L2S group). However, these results might be different for both groups if another variable such as meaning recognition, requiring a connection between form and meaning of the lexical items, had been assessed in the study.

These results are in line with King (2002), who found out that L2S reinforce the aural input through the written form appearing in the subtitles. In our study, the inclusion of L2S provided an extra source of input consisting of the correct written forms of the new and known words that students were also exposed to in the oral form. The addition of L2S allowed learners to better segment and isolate the words as separate units (Danan, 2004; Winke et al., 2010; Peters et al., 2016) and this may have facilitated L2 word-form recognition later on for them (Danan, 2004). That is, written subtitles might have helped participants to segment the speech stream and might have made word identification easier (Vanderplank, 1988).

Similar results have already been found for older learners. For example, several studies on written word-form recognition comparing the effect of watching videos with L2S or without subtitles for adult L2 learners obtained favourable results for those

watching videos with subtitles in the L2 (Markham, 1999; Yuksel & Tanriverdi, 2009; Sydorenko, 2010; Nagira, 2011; Montero-Pérez et al., 2014). In addition, previous research examining this lexical variable with adults but comparing the effects of L1 and L2 subtitles has also obtained favourable results for the latter (Hui, 2007; Frumuselu et al., 2015). Furthermore, Peters et al. (2016) conducted a study on L2 vocabulary recognition with teenagers, which obtained significant differences in favour of the L2S group. The results in our study would also be in line with Hui (2007), who included a listening word-form recognition test in his study with adult L2 beginner learners and obtained better results for the group with subtitles in the L2. However, aural word recognition was not a variable that we tested in the present study. What is relevant, though, is that findings in studies with older learners are confirmed in the present study including younger and less proficient learners.

It is also remarkable that more significant differences were obtained by the L2S group at the end of the treatment (contrary to what was happening with comprehension, where differences were found throughout the treatment). This suggests some accumulation of input is needed for learners to get accustomed to TV viewing with L2S in order for them to learn from this input. These findings would be in line with Gesa (2019), who explored TV viewing in primary and secondary school learners in a very similar context to ours (semi-private school in Catalonia). Taking into account that the learners in primary school in his study were at Grade 6 (and ours at Grades 5 and 6), and that he was assessing both word-form and meaning recognition (and we also assess word-form recognition), it is interesting that more evidence is showing that the effects of exposure to multimodal input cannot be seen immediately but take some time, especially with young children at low proficiency levels.

It is also worth noting that the Mann-Whitney analyses performed which compared both groups in terms of the total number of known and new words they recognised, confirmed the results that the L2S group recognised significantly more words (both known and new) than the L1S group. Close attention should also be paid to the finding that both groups recognised significantly more words in the videos that were already known to them (as the Wilcoxon Signed Ranks Test results gave significant results for each group). On the one hand this finding supports what we were explaining above: learning new vocabulary from multimodal input is difficult (for both groups, independently of the type of subtitles provided) and it requires time and considerable exposure, even if we are just assessing word-form recognition (which does not imply making links with meaning or learning about language use). On the other hand, it also suggests that exposure to subtitled TV series can be a good way of providing new encounters with vocabulary that has already been introduced in class. This has not usually been assessed in L2 studies, when just the learning of new words is the focus of research. However, we should be aware that vocabulary we have previously been exposed to is not permanently activated (as already mentioned by Ebbinghaus 1913/1885) and that it should be revisited so as to become an active part of the lexicon (i.e., that students can fall back on when needed). It is true, though, that little effort is put into vocabulary recycling in teaching materials. Therefore, alternative forms of input (such as audiovisual) can be a good way to promote further encounters with vocabulary once introduced in EFL classes.

The combination of explicit teaching in class and finding words later in TV viewing can be beneficial for learning (even if episodes are subtitled in the L1), as has also been found in Gesa (2019). Previous studies analysing word-form recognition through

multimodal input without any previous knowledge of the items in young low-proficiency L2 learners may not bring about positive results. For example, no significant differences were found in Galimberti and Miralpeix (2018) when comparing primary school children watching TV series with or without subtitles.

To sum up, the young low-proficiency learners who watched the videos with L2S obtained higher L2 vocabulary recognition scores. When a distinction was made between known and new words scores, the results were still favourable for the L2S group in both categories. However, results tend to be significantly better for known word recognition (also in the L1S group). To different extents, the study shows that multimodal input could be beneficial for recognition of new words, but especially for re-activating already known lexical items that have been introduced in previous formal instruction: the exposure through different channels might have aided participants in recycling L2 vocabulary, supporting the proposal of focused instruction prior to TV watching in Gesa (2019) and Pujadas (2019). Considering the existing research on written vocabulary recognition, the distinction between known and new words had not been previously examined in any other longitudinal studies that used animated TV series and further research is encouraged to generalise the results.

The results from research on ER have already shown that this practice facilitated L2 word-form recognition in older learners (e.g., Horst, 2005). ER has also been claimed to be beneficial for all proficiency levels (Woodinsky & Nation, 1988), but ER and listening have been claimed to be especially well-suited for intermediate and advanced learners (Beglar & Hunt, 2002). We suggest that extensive viewing can also be a good practice for beginner learners, similarly to the suggestions made by Gesa (2019).

The vocabulary results in the present study show that after watching each episode, participants were able to identify some of the known and new words encountered several times. Previous research has suggested that young low-proficiency learners might have problems processing formulaic language (Conklin & Schmitt, 2012), especially if they do not have enough exposure to these words. Our results show that the learners from both groups were able to recognise both single words and multiple word units in multimodal input. Further research is needed, however, in determining, for example, if there are any word-specific features that may make certain lexical items more readily learnable from multimodal input for children at this level, or whether there is a necessary number of encounters in this type of input that is most suitable for learning to take place. There is still controversy nowadays regarding the number of times an item needs to be encountered when reading before it is successfully acquired (e.g., Huckin & Coady, 1999; Webb, 2007), even if there is agreement that, in both reading and listening, word-form identification requires less occurrences than providing its meaning or definition, as suggested by Webb (2007), Pellicer-Sánchez and Schmitt (2010), Van Zeeland and Schmitt (2013).

8.4. The role of aptitude and proficiency on L2 comprehension and vocabulary learning through multimodal input

The second RQ of the study analysed the possible relationship that aptitude and proficiency may have with language learning from multimodal input. Aptitude was measured with LLAMA B (Meara, 2005), which basically tests rote memory and associative learning. L2 proficiency was operationalised using three variables: namely L2 VS, RS (in L1 and L2) and English class grade.

8.4.1. Aptitude

It has been claimed that linguistic aptitude plays an important role in L2 learning (Skehan, 1989; Muñoz, 2010; Grañena, 2012) and our results suggest that this learner-related factor plays a role in L2 comprehension (as suggested in Dörnyei & Skehan, 2003). The results from our study show that linguistic aptitude influenced the L2 comprehension scores of our L1S and L2S learners. Therefore, in this case aptitude has a positive influence for children learning second languages, indicating that those with a higher aptitude also obtain better scores, in line with previous research (Dekeyser, 2000; Sparks et al., 2009; Hummel, 2009; Li, 2019; Doughty, 2019; Rogers et al., 2017). These findings are also consistent with previous studies analysing aptitude and language learning in young learners in formal settings, where aptitude was measured by means of the MLAT-Elementary (instead of LLAMA). Therefore, this suggests that this ID has a strong influence on L2 learning and should be taken into account when assessing language development (Suárez, 2010; Rosa, 2011; Muñoz, 2014b). It is remarkable that the regression indicated that aptitude was accounting for 24% of the variance in the comprehension scores of both groups (more specifically, 24.4% in L1S group and 24.9% in the L2S group). This heavy weight in the scores deserves close attention. Furthermore, it is also revealing that there are strong correlations in both groups between L2 comprehension in E10 and E20 and LLAMA_B scores, showing that aptitude is relevant for understanding even when subtitles are not present: it plays a role in making sense of the input as well.

However, the results were not the same when we analysed aptitude in relation to vocabulary learning. Only in the L1S group was there a moderate significant

relationship between vocabulary scores and aptitude (LLAMA scores explained up to 10% of the variance of the vocabulary scores). These results indicate that aptitude influenced and helped L2 vocabulary recognition in the L1S group, whereas it did not help in the case of the L2S group (where no correlations were found and the regression analysis did not reveal any significant results either). It should be noted that, when examining the vocabulary test results in RQ1, the L2S group tended to score significantly higher than their L1S peers. However, the results from the analyses in RQ2 do not show any significant correlation for the L2S learners between their aptitude and L2 word-form recognition scores. A possible explanation for these results could be that L1S students, due to the different languages included in the input (L1S and L2 soundtrack), were faced with more difficulties in identifying the L2 words in the episodes and performing the vocabulary tests. They might have been forced to rely on their aptitude to complete the tests and, due to that, the L1S participants with a higher aptitude might have obtained better results. However, in the case of the L2S group, learners were exposed to the written form of the TWs in the L2, facilitating the task of matching aural and written forms (and maybe eliminating the need to rely on aptitude for that purpose). It is worth recalling here that LLAMA B assesses associative learning and rote learning, and possibly more demanding associations had to be established in the case of L1 written forms and L2 oral forms than between two L2 forms (oral and written). The results obtained when focusing on known and new vocabulary mirror exactly those obtained for L2 vocabulary in both groups, indicating that language aptitude did play a role in enhancing word-form recognition only for L1S learners. Again, this means that L1S learners with a higher aptitude were able to identify more known and new lexical items than other L1S students with lower aptitude scores.

It is interesting at this point to compare our results with the very few we have available on aptitude and video viewing. First of all, Gesa and Suárez (2022), which was also longitudinal and carried out in an EFL context with primary school learners showed that aptitude influenced vocabulary scores in the word-form test, which is the same as we find in this study even if we have used a different vocabulary test format and the vocabulary was not introduced immediately before watching the video. Secondly, a study by Miralpeix et al. (in press) where Catalan / Spanish bilingual learners watched an advert in English with Polish subtitles (an unknown language for participants), also suggested a strong influence of aptitude on meaning recognition test results within the very first minutes of being exposed to a completely new language. Therefore, these two studies show a mediating role of language aptitude when learning from multimodal input. In another study in which participants were older (Suárez & Gesa, 2019), the influence of aptitude was shown to be minor, while proficiency level was gaining ground on test results. In Pattemore and Muñoz (2020), no influence of aptitude was observed in learning grammar constructions from video viewing, but this study was conducted with adult learners. An influence of aptitude on vocabulary learning from a documentary was also found in Teng (2022), although in this case, again, participants were adult Chinese learners of English in a formal setting, not young learners. Our results would thus show that aptitude influences L2 comprehension in subtitled TV viewing to quite a large extent (independently of the type of subtitles used) and to a lesser extent would aid vocabulary learning (word-form recognition) when the task implies the L1 (in the subtitles) and the L2 (in the audio), as this is more challenging for students than having the L2 both in the audio and in the subtitles.

Finally, it should be noted that our study was conducted in a formal setting where our participants watched authentic input. Skehan (1989) had argued that the scores from any L2 aptitude test would be clearer and more revealing if the learners received the input from naturalistic environments, where more significant differences might be found. However, previous research has also found evidence of the role of aptitude in instructed settings (Dekeyser, 2000; Abrahamsson & Hyltenstam, 2008). Even though our research was conducted in an EFL setting, the input the L2 students were exposed to was authentic. This is a characteristic of naturalistic environments, which could be another possible reason for our significant findings.

8.4.2. Proficiency

Proficiency was a key variable in this study, not only because it was one of the variables measured but because the learners participating in it had a very low level. Due to that, care had to be taken in the selection of the series so that the experience was successful. Even though previous research has claimed that authentic materials might not be suitable for low-proficiency learners (Zanon, 2007; Martínez, 2002), several studies have defended the suitability of cartoons for children's L2 learning (Bahrani & Soltani, 2011; Bahrani & Sim, 2012). In addition, our results are also evidence that the inclusion of subtitles was helpful, even if students had a poor L2 reading ability, as the non-subtitled episodes proved to be harder (especially for the L1S group). Special care was taken in selecting the proficiency variables that previous research had suggested could be relevant (e.g., in studies with adult learners or one-off studies) or that we thought could have an influence on the type of input we were investigating (e.g., RS). Based on our results, we discuss below how proficiency

variables were related to understanding and learning from subtitled videos in young learners.

Starting with comprehension, previous research has highlighted that L2 learners need a minimum VS to understand multimodal materials (Webb & Rodgers, 2009b), and this tends to be lower than what we need for reading or listening alone (as the use of different input sources can make up for possible deficiencies in lexical knowledge). The correlations conducted between the VS and the comprehension scores in both groups indicate a strong relationship between these variables, with VS accounting for 24.2% of the variance in comprehension scores (in the L1S group) and for 10.8% in the L2 group. Significant correlations were also found in the case of E10 and E20, without any subtitles. These results corroborate findings in studies with older and more proficient students, which have also found that learners with higher VS scores obtain greater comprehension results (Montero-Pérez et al., 2013; Montero-Pérez et al., 2014). Therefore, the present longitudinal study is one of the first in highlighting the role of L2 receptive VS in very low-proficiency learners watching subtitled television.

It should be considered that most research so far deals with participants having much bigger VSs. In this case, the vocabulary knowledge of our young L2 learners was poor (and, due to that, the participants from our study only took the first three bands of the Vocabulary Size Test -VST-, following Schmitt, 2000, who claims that the use of the 1,000 and 2,000-word level bands for beginner learners is adequate). Our participants had an approximate VS of about 1,200 words, so finding an adequate TV series from which they could learn was challenging.

Nation and Beglar (2007) argue that in order to be able to understand a children's film 6,000 WFs are needed. However, Nation (2006) claimed that only 4,000 WFs were required to comprehend the children's film *Shrek*. Furthermore, he added that learners 'only' needed to be at the 2k word-frequency band to understand 87.91% of the film's content and concludes that the VS for animated films is the lowest one when compared to other TV genres. Nation's findings are also supported by Webb and Rodgers (2009b), who claim that only 2,000 WFs are needed to understand children's programmes. Our analysis of the TV series scripts and the coverage (see Appendix H) suggest that learners with knowledge of the first 2k bands have between 91.7 and 96.5% of coverage, so coverage was above 90% in all episodes. As not all of our students had reached complete mastery of 2k band (as we pointed out, the average VS was of about 1,200 words), coverage was a bit lower than 90% for some children. Given the fact that lower coverage is needed for TV watching than for reading or listening, we think that the coverage was appropriate for learners to understand the input provided (and that was also corroborated by the comprehension scores that both groups obtained). At the same time, it gave room for vocabulary learning (and that was corroborated by the vocabulary test scores in both groups).

Furthermore, the selection of several episodes from the same TV series implied that our L2 learners were doing narrow reading (i.e., the subtitles contained recurrent elements) in each episode. Narrow reading has proved to be an effective method for L2 vocabulary learning and it can also be applied to watching episodes from the same TV series (as Webb & Rodgers, 2009b, have also suggested). Narrow reading implied a decrease in low-frequency words for our learners, due to the repeated number of encounters of recurrent words, and a lower VS demand in order to understand the

input (Webb & Rodgers, 2009a). Our results would also support Bravo (2008) and Bahrani and Sim (2012) in that animated TV series require less cognitive processing than other TV genres and could have played a facilitative role in better understanding of the plot (Peters et al., 2016).

We have also seen that results obtained for VS and our two variables (comprehension and word-form recognition) are very similar to those for English class grade and the same variables. Actually, we have also observed a close correspondence between class grades and VS (that is why we did not enter English class grade into the regression). First of all, this confirms that VS scores could be used as a surrogate for general English proficiency, as already pointed out, for example, in Miralpeix and Muñoz (2018) in older learners. It seems that this correspondence exists since the very first stages of learning a language. It could also be stated that our results are in line with the beginners' paradox (Coady, 1997), which was posed first for reading, but can also be true for watching. That is, learners acquire vocabulary from TV watching, but if they do not have a minimum VS (and proficiency level), they will not understand the message and will have fewer probabilities of learning, so we should find the right materials with an adequate coverage.

Secondly, it should also be noted that, even if results for VS and English class grade are similar in terms of the correlations with comprehension and vocabulary learning, the strongest correlation can be found between English class grade and comprehension scores in the L2S group (the correlation is stronger than the one found for VS). This can be explained because English class grade also takes into account

other abilities such as reading and listening (not just vocabulary knowledge), which are also employed when watching subtitled television.

In relation to RS, results revealed a medium significant correlation for the L1S group but none for the L2S learners. These findings suggest that L1S learners were able to read and follow the subtitles in order to understand the multimodal input from the videos: their L1 RS helped them to understand better and get better results in the comprehension tests. Actually, results from RQ3 also prove that participants were reading the subtitles. These findings may also suggest that reading subtitles is an automatic behavior, especially when they are in a language we fully understand (d'Ydewalle & Gielen, 1992; Pavakanun & d'Ydewalle, 1992) and the images do not distract the learners from paying attention to the subtitles as well (d'Ydewalle & Vanrensbergen, 1989; d'Ydewalle & Bruycker, 2007; Tragant & Pellicer-Sánchez, 2019).

Nevertheless, the lack of significant correlations and significant results in the regression for the L2S group suggests that the L2 subtitles might have been too challenging to read for L2S learners and, due to that, their L2 RS did not have an impact on their understanding of the story. Usually, learners have slower RSs in their L2, as also happens with bilinguals and their weak language (Whitford & Joanisse, 2018). Our L2S students might have had difficulties in reading the input due to a combination of factors, such as the limited amount of time that the subtitles appeared on screen and their poor L2 reading ability. Previous research claims that reading subtitles in the L2 is a continuous challenge for low-proficiency L2 learners (Roberts & Felser, 2011; Vanderplank, 2013), especially if these subtitles are presented at a

quick rate (Zárate, 2008). However, Marzá and Torralba (2015) claim that young low-proficiency learners will be able to follow the six-second rule of subtitles (as was also the case in our study) when they start Grade 5 and our students were in Grades 5 and 6 during the treatment. In our study, the L1S group obtained higher comprehension results in the episodes where the subtitles were included. In addition, even though the results from the L2S group were not as high as their L1S peers, the lower means obtained in the episodes without any written input show that these learners still benefited from the presence of the subtitles to some extent.

Our results might suggest that when audio and subtitles are both in the L1, the subtitles could be edited so as to match the L2 RS of the participants, which is slower than L1 RS. Other researchers conducting research with young low-proficiency viewers also proposed that the subtitles appear on the screen for a longer time, depending on the RS of the learners (e.g., Koolstra et al., 1999; Fresno, 2018). However, there has not been much research on that. A slower subtitle speed and the addition of pauses between subtitles might have provided our learners with an adequate reading pace for cognitive processing (Tamayo, 2016), which could be the way to obtain better results in comprehension and possibly learning.

In relation to those proficiency variables influencing word-form recognition scores, only VS could explain up to 8.41% of the variance of these scores in the L1S group, but no significant results were found for the L2S group. Similar results were obtained when the known and new words were analysed separately: the results show that L1S learners needed to rely on their previous lexical knowledge in order to identify the known words due to the lack of L2 words in the written input. The positive correlation

obtained in E10 (without subtitles) for the L1S group also implies that these learners tried to rely on their previous L2 vocabulary knowledge more than the L2S group when no subtitles were included. We think that the presence of L2 written and aural input in the episodes watched by the L2S group undoubtedly helped L2 word-form recognition for these students and did not generate a need for them to build on their previous L2 vocabulary knowledge.

Furthermore, when examining the new words variable correlations, there is a medium significant relationship between the L2 word-form recognition of new words and the L2 VS of L1S learners. Even though the words were completely new for these learners, the fact that their written form was provided in the L1 probably meant that their already known vocabulary in the L2 was helping them decipher what was being said. The recency effect of having seen the words in the L2 written form for the L2S group meant that, independently of their VS, they could identify the TWs. That is why VS was just relevant for their comprehension, but not for L2 word identification.

The few existing previous studies analysing the possible relationship between learners' VS and lexical knowledge gained from TV viewing has provided positive results in the case of Montero-Pérez et al., (2014); Peters et al., (2016) and Suárez et al. (2021), but no significant results in the study by Rodgers (2013). However, none of these studies were conducted with young low-proficiency L2 learners.

It is worth commenting that, just as happened with comprehension, results for the relationship between English class grade and the correct identification of vocabulary from the video follow the same pattern as those found for VS. However, correlations

between the English class grade of the participants and the vocabulary scores showed a medium significance in the L1S group only with the known words. This again suggests that those that were more proficient in English could correctly identify more words in the video that had already been introduced in English classes, but it was still hard for them to remember new word forms appearing for the first time in the video if the subtitles were offered in the L1. Furthermore, the only significant finding for the L2S group was in E20 (the last episode without subtitles): a significant correlation between the English class grade variable and the vocabulary scores suggests that just at the end of the treatment students with a higher proficiency (taking into account also reading and listening scores) were identifying more vocabulary from the L2S than those with a lower class grade.

Finally, RS did not have any role in L2 word-form recognition, even if, when examining the mean vocabulary scores for RQ1, the L2S group obtained better results than their L1S peers. This does not mean, though, that they were reading more quickly, but probably that they could isolate TWs better. We know from the results obtained for RQ3 that they were reading the subtitles, but according to the results for RQ2, RS did not influence L2 word identification (as word-form recognition does not actually imply reading the whole subtitle but only the recognition of the word itself). Regarding the L1S learners, they could not isolate the TWs because these students were presented with the subtitles in their L1 (even if they could read more quickly, this facilitated comprehension but not L2 word recognition in the immediate post-test).

8.5. Extensive viewing: participants' perceptions and viewing behaviour at the beginning and end of the treatment

The third RQ examined whether repeated viewing over a long period of time may influence the way in which participants watched the *Curious George* episodes, their perceptions of understanding and learning. It was also a way to know learners' opinions that would help us to triangulate some of the results that we obtained for the other RQs. Very few studies have collected data about viewing experience from questionnaires (or interviews). However, we think the point of view of these students could reveal interesting insights into how young learners perceive this experience and possible changes this practice may bring about in the long run that may or not may influence L2 learning. That is why apart from the descriptive data from their answers, McNemar's tests were conducted for each group to check whether any significant changes had taken place during the months that the treatment was taking place.

8.5.1. Self-reported level of understanding

Results of the McNemar's test conducted to compare the responses from each group at the two different times the questionnaire was answered showed significant differences for both groups, indicating that self-perception of learning clearly improved over time, independently of the type of subtitles learners were exposed to. These findings can also be explained as a result of cumulative background knowledge about the series and the familiarity effects of being recurrently exposed to this type of video (and of being assessed with the same test format). The fact that students perceived that their level of understanding was higher does not correspond to the empirical

results found for comprehension in RQ1, which does not show a steady progression over time. Even if an interaction was found between time and comprehension in the ANOVA performed, we could clearly see that scores were episode-dependent and in line with what other studies have found (both for older adults as in Rodgers, 2013, or in younger learners such as in Gesa, 2019). Therefore, even if learners from both groups believed they were understanding more, this could not be confirmed by the results of the comprehension tests. It is clear, though, that learners from both groups understood the episodes from the beginning of the treatment (which confirms that coverage was adequate for their VSs).

8.5.2. Viewing behaviour

Regarding viewing behaviour, no significant differences were found when comparing the learners' behavior when reading the subtitles at the two different administration times in each group. This implies that the experience did not modify the way students were watching the episodes, which is mainly reading and listening at the same time (they were not just 'listening' or 'reading' more at the beginning or end of the treatment).

Only at the descriptive level do we notice that the L1S group was reading a bit more than their peers, who had the subtitles in the L2 (they have also been shown to have a quicker RS) and this tends to happen both at the beginning and at the end of the treatment. However, differences do not reach significance.

It is also worth commenting that only a small number of students selected options 3 ('Listening only. Never read the subtitles') and 4 ('Reading only. Not paying attention to the audio') at both testing times, which suggests that the L2 learners in our study tried to read the subtitles and paid attention to the audio since the very beginning of the treatment. Previous research has proved that reading subtitles is an automatic behavior (d'Ydewalle & Gielen, 1992; Pavakanun & d'Ydewalle, 1992) and it has been argued that if they are displayed on the screen, they are processed and read (Gielen, 1988). However, we think that our results could be explained by the fact that they lacked the knowledge to just listen in order to understand what was going on. At the descriptive level, we also see that some students read the subtitles less at the end of the treatment, but just a small number and therefore this did not reach significant difference in the statistical test.

The fact that they were reading and listening at the same time throughout the treatment, independently of the type of subtitles they were offered, can also explain the complaints the students made when E10 and E20 were shown to them without subtitles. Our participants were used to the presence of the written input and they lacked it in order to understand the episode. In addition, they were reading all along because subtitles were also helpful in deciphering the input and in identifying possible unknown words (Neuman & Koskinen, 1992; Danan, 2004; Winke et al., 2010). We should remember that results were lower when the learners were exposed to an episode without textual support (E10 and E20).

8.5.3. Self-perception of learning from the subtitles

The video-watching questionnaire also enquired about the role subtitles may have for learning. Only those students saying that they were reading the subtitles were required to answer these subsequent questions. However, as we have seen, most of the sample read the subtitles during the treatment.

8.5.3.1. Time spent reading subtitles

Both at the beginning and end of the treatment, most L1S learners selected those options that indicated they 'often', 'very often' or 'always' had enough time to read the subtitles. This result does not evidence any longitudinal changes in the L1S participants' reading proficiency throughout the intervention, which is backed up by the lack of significant differences in the McNemar's test results. The responses from the L1S learners indicate that they did not have major problems following the subtitles and could possibly process them quite automatically, as has been claimed by previous research (d'Ydewalle & Gielen, 1992; Pavakanun & d'Ydewalle, 1992; Tragant & Pellicer-Sánchez, 2019).

On the other hand, the L2S learners may have encountered some more difficulties and lacked enough time when trying to read the subtitles. These students might not have had enough time to read all the input provided on the screen, which might have been too challenging for their low RS (Zárate, 2008; Muñoz, 2017). When looking at the individual changes in the answers of these learners, only 28.57% maintained their responses whereas half of the L2S learners (50%) selected an option that implied they

had more time to read the subtitles at the end of the treatment. Despite this fact, some still expressed that they lacked time, which implies that L2S were still hard to read for some of our young low-proficiency L2 learners. No significant differences were found for the L2S group either, although it should be noted that many L2S students changed their responses towards a better reading ability at the end of the intervention (they felt they needed less time towards the end of the treatment), and that is why we find a nearly significant result in the McNemar's test.

Even though it has been claimed that young low-proficiency L2 learners (10-11 years old) can follow the subtitles when watching cartoons (Marzá & Torralba, 2015), this may not always be the case. However, L2S learners may have been used to reading the subtitles after several weeks: the sustained exposure to multimodal input might have made them increase their reading ability in the L2. Unfortunately, however, L2 RS was not assessed again at the end of the treatment and this finding cannot be corroborated. Some more extensive viewing may have led to significant results, observing the pattern of the answers for students in the L2S group, which is a bit different from that of students in the L1S group.

These results might reinforce the need to extend the presentation time of L2S on the screen for this specific audience (Koolstra et al., 1999; Fresno, 2018) to improve comprehension of the story, especially when subtitles are in the L2. On the contrary, if we think that half of the L2S learners (nearly) significantly improved their reading ability, it should be concluded that the longitudinal effect of the study also had a positive effect on these students.

The results from the questionnaire can be related to the correlations performed between the RS variable and the comprehension scores in RQ2. There was only a significant correlation for the L1S students, which showed that learners from this group with a higher RS were able to understand the input better. The responses provided in the questionnaire by L1S learners confirm that they could read the subtitles in a potentially sufficient way since the beginning of the treatment. On the other hand, the lack of significant correlations between the L2S comprehension scores and their L2 RS is also related to the fact that L2S students had problems following the subtitles, especially at the beginning (although these differences are just appreciated in the raw percentages and only nearly approached significance).

Therefore, we cannot say that the subtitle reading behaviour of both groups significantly improved after watching the TV episodes included in the intervention. It is likely that some more time would be needed for this to occur, especially for the L2S group. Thus, the need for longitudinal studies taking this factor into consideration is clear, so as to find out the suitability of L2S for young low-proficiency L2 learners in longitudinal studies. Students' answers would suggest a beneficial effect of the longitudinal treatment and confirm the positive impact of ER (Paribakht & Wesche, 1997; Huckin & Coady, 1999) and narrow reading (Webb & Rodgers, 2009b), also when watching subtitled TV. However, no definite conclusions can be reached given the borderline significance in the present study.

8.5.3.2. Reasons for reading subtitles

More than half of the learners in the two groups in the study, both at the beginning and end of the treatment, selected options that implied that the written input helped them in understanding the content of the story. Furthermore, learners in both groups acknowledged that by reading them they were learning new things. There was not a remarkable change in these opinions from beginning to end of treatment. However, some more students in the L2S group chose the option related to learning more at the end of the treatment. This might be related with the L2 word-form recognition task in the vocabulary tests: L2S learners watched the subtitles in the L2 and the words included in the test were also in this language. This might have helped them realise that when reading L2S they can identify the words they are asked later on more easily.

It is also noticeable that the only significant difference arises in the option ‘because they appear on the screen and I cannot avoid reading them’ in the case of the L1S group. They seem to exhibit a more automatic reading behaviour at the beginning, as has been found in other studies (Gielen, 1988; d’Ydewalle & Gielen, 1992; Pavakanun & d’Ydewalle, 1992; Tragant & Pellicer-Sánchez, 2019), but then at the end fewer students choose this option and they do not read ‘automatically’: they are more aware that the subtitles help them understand what is happening. The findings suggest that the sustained exposure to multimodal input might have made them realise that there were more important reasons for reading the subtitles than doing it because they appear on the screen (in the case of the L1S group for better comprehension and in the case of the L2S group for better learning). It should also be noted that most learners selected options two and four at both testing times, which shows their

awareness (since the beginning of the treatment) of the positive advantages that written input entails for L2 comprehension and written vocabulary recognition. Therefore, young low-proficiency L2 learners can benefit from the assistance provided by the subtitles watching animated TV series.

8.5.3.3. Subtitles for vocabulary learning

Both groups of students realised they learned L2 vocabulary from the written input. It should be mentioned that even though both groups mostly selected that they had learnt 'some words', the percentage was higher for the L2S group. In addition, the number of students who said subtitles helped them learn 'a lot' of words was higher in the L2S group as well, and the McNemar's test showed significant differences. These results are in agreement with those found for the first RQ and with previous research conducted with young L2 learners that obtained significant benefits for L2 word recognition when the learners were only exposed to one session of the multimodal input (Neuman & Koskinen, 1992) or when they watched several episodes (Gesa, 2019).

On the other hand, the L2S students who indicated learning no words through the reading of the subtitles were already a very reduced number at the beginning of the study, and only one student in this group said they had not learned anything. The percentages of learners in both groups that claimed learning more L2 words at the end of the treatment are higher when compared to the students who expressed learning fewer words (13.3% for L1S and 0% for L2S).

These answers corroborate that L2S learners acknowledged that the presence of L2S helped them to recognise L2 word-forms since early in the beginning of the intervention (also the students in Pattemore et al., 2020, stated that L2S helped in learning at all proficiency levels, even if in that case a comparison was not established with L1S). On the contrary, at the end of the intervention about 20% of the students in the L1S group still thought that subtitles were unhelpful for learning vocabulary. This could be at least partially motivated by the fact that students were tested on word-form recognition, not on meaning. These findings suggest that L1 subtitles combined with L2 soundtrack did not aid the L1S learners as much as L2 subtitles did for their L2S peers. The results also corroborate previous findings that L2S are better than L1S for L2 vocabulary recognition (Hui, 2007; Frumuselu et al., 2015; Peters et al., 2016).

To sum up, results suggest that L2S participants perceived themselves to have learnt more L2 words than the L1S group (especially at the end of the treatment), which is in line with the significant results obtained by the L2S group in RQ1 for L2 word-form recognition. However, the results for the L1S group also followed this pattern and students tended to say they had learnt some vocabulary from reading, although there were more students stating that they were not learning any vocabulary than there were in the L2S group.

8.6. Out-of-school exposure and learning from multimodal input

The fourth RQ of our study focused on the possible relationship between the participants' OSE involving multimodal activities, such as watching subtitled TV series or movies and playing videogames in English, and their performance on the tests after

watching the subtitled cartoons in class. Some of the learners from each group in our research reported viewing TV series or movies and playing videogames in English and, due to that, each group was divided into two subgroups. It should be noted that the amount of learners with OSE was lower than those with no OSE (both L1S and L2S group), evidencing that, in our context, not many participants were receiving OSE in their L2 at home, even if the numbers were quite balanced in each group.

When examining our results, the *t*-tests conducted showed that OSE to multimodal input did not have an effect on the scores obtained in the tests (neither for comprehension nor for vocabulary). These findings might be related to the amount of dubbed audiovisual input the learners had actually received at home. Previous research conducted in other countries has obtained significant results for the effects of L2 OSE on comprehension (Kuppens, 2010; Sylvén & Sundqvist, 2012; Lindgren & Muñoz, 2013; De Wilde & Eyckmans, 2017; Muñoz & Cadierno, 2021) and vocabulary learning (Turgut & Irgin, 2009; Jensen, 2016). In order to explain these results, however, we should be aware that European learners from most of these studies were exposed to subtitled audiovisual input from a very young age at home, which means that the number of hours for which they have received multimodal input in the L2 is high. As results from the questionnaire also show, our participants in Spain had very limited exposure compared to learners from other European countries where all the audiovisual materials are subtitled. In our context, they were only minimally exposed to English through audiovisual input. It should be taken into account that Spain is a dubbing country compared to others like Sweden, Denmark or Belgium, in which subtitling is promoted and, therefore, more opportunities are provided to learn the L2 through sustained exposure to TV series or films (Sundqvist & Wilstrom, 2015; Jensen,

2016; De Wilde & Eyckmans, 2017). Previous research conducted with L2 learners from countries favouring the use of subtitled TV programmes and films (or where children play more videogames in the L2), such as Sweden, has quite often found a significant impact on their L2 learning (Sylvén & Sundqvist, 2012; Sundqvist & Sylvén, 2014; Jensen, 2016). Nevertheless, these studies showed that their participants were exposed to English OSE a total of 9.4 (Sylvén & Sundqvist, 2012) and 7.2 hours per week (Sundqvist & Sylvén, 2014). More specifically, participants received OSE to English more than two hours per week through watching TV, which was not the case for our students. The difference between the broadcasting techniques (dubbing or subtitling) in Sweden and Spain might explain these results: our L2 learners lived in a country that usually dubs audiovisual materials in English and the amount of OSE these learners were exposed to was minimal. Therefore, recurrent exposure to multimodal input was quite new to them, which might have led to the lack of significant results. Even though some L2 learners from our study indicated that they were exposed to OSE at home through playing video games and /or watching subtitled movies with L2 soundtrack (with L1 or L2 subtitles), their engagement might not have been high enough to provide significant outcomes (Muñoz et al., 2018; Peters et al., 2019; Muñoz & Cadierno, 2021).

Even though children in Spain can choose the language of their TV programmes, they need to be willing to include L1 or L2 subtitles and change from their L1 soundtrack to the L2. It is expected that at this young age the children would prefer, if they had the choice, to watch TV shows or movies in their L1 rather than in their L2. Even though their L2 proficiency would be enhanced if they switched to their L2, it is not compulsory for them to watch TV in this language (and it is not the 'default' option). Furthermore,

it would be a challenging task for them: their low L2 proficiency does not encourage language switch (despite, sometimes, the preference and willingness of the parents). However, in countries where the learners have no other option but to include subtitles, it is assumed that these students receive an unlimited amount of L2 input from an early age. This idea can be backed up by the study performed by Muñoz and Cadierno (2021), who compared the OSE of same-age teenagers from Denmark and Spain. The longer hours of OSE were possibly the reason for the significant outcomes favouring the Danish students in most of the language tests.

Moreover, the students' L1 from most of these studies shared a close proximity to English, which might have also facilitated the task to understand the input. This idea has been suggested in previous research conducted by Muñoz et al. (2018) and Muñoz and Cadierno (2021), who compared L2 learning in Spanish and Danish students and mention that one of the causes in favour of the Danish students is the close proximity of the learners' L1 to English. The positive outcomes obtained by Jensen (2016) with Danish students and Sylvén and Sundqvist (2012) and Sundqvist and Sylvén (2014) with Swedish learners could also be explained by the close proximity of their L1s (North Germanic languages) with the L2 (English). Finally, another factor that could have influenced the results in several studies analysing the effect of OSE in Iceland (Lefever, 2010), Flanders (Kuppens, 2010; De Wilde & Eyckmans, 2017), Denmark (Muñoz et al., 2018) and Sweden (Sylvén & Sundqvist, 2012; and Sundqvist & Sylvén, 2014) could be the strong social presence the L2 has in these geographical contexts. On the other hand, English has a limited presence in Spain in comparison with these settings, which means that learners are mostly exposed to English just in class.

Even though there is a growing interest in this area of research, to our knowledge only a few studies have been conducted in Spain analysing the relationship between OSE and L2 learning (Lindgren & Muñoz, 2013; Muñoz, 2020; Muñoz & Cadierno, 2021). Previous studies have claimed that incidental L2 learning through watching multimodal input at home might be the key to contributing towards young L2 learning success and would compensate for the limited amount of school instruction L2 learners receive (Jensen, 2016; De Wilde & Eyckmans, 2017). Even though the results with Spanish students in this study were not significant, probably due to the low number of hours for which our participants had received OSE, recent research in Spain has shown a correlation between the L2 participants' English grades and OSE (Muñoz, 2020). This study took into account other aspects of English proficiency (not just comprehension and vocabulary test scores as in the present study) and suggests that the trend may be changing. The data from the current study was collected in 2016 and at the time, streaming services such as Netflix, Amazon Prime or HBO were not available. However, Muñoz (2020) explains that participants' impatience to watch TV series made them watch them in English with subtitles (this is a reason that would promote exposure to multimodal input in the future as well, since the dubbed versions are always released a bit later). Due to that, we believe that the learners' OSE, especially in the form of multimodal input, should continue to be considered in further research when performing SLA research (Fresno, 2018). Even if no effects of OSE were found in this study, given the results obtained in other settings, L2 learners should be encouraged to engage in extensive L2 TV viewing outside school by their teachers anyway (Lindgren & Muñoz, 2013).

CHAPTER 9 - CONCLUSION

This PhD thesis aimed at investigating L2 comprehension and written word-form recognition in young low-proficiency L2 learners after watching several episodes from a subtitled TV series. It also explored the relationship of aptitude (measured by the LLAMA test) and proficiency (VS, RS and English class grade) on the participants' results. A further objective of the present research was to enquire about possible changes that could be taking place in the ways children dealt with this input (their viewing behaviour –while also reading and listening–, perceptions of learning, etc.), and to check whether previous experience in dealing with multimodal input had any significant effect on the comprehension and vocabulary scores in the intervention.

In this study, participants (11–12 years old) were divided into two groups and watched a *Curious George* episode weekly, subtitled either with L1S or L2S (except E10 and E20, which did not include subtitles). Participants also took a VST, LLAMA_B and L1 and L2 RS tests. They answered two questionnaires as well: one related to the experience of viewing subtitled videos (once after watching the first episode and once again after the intervention) and another questionnaire with biodata enquiring about OSE (before starting the intervention). After viewing each episode, learners completed a test to examine their L2 comprehension and written word-form recognition (including known and new words appearing in the video). It should be recalled here that two of the episodes (E10 and E20) were watched without subtitles in order to compare the behaviour of the two groups in the middle and at the end of the intervention, when textual support in the L1 or L2 was not provided.

This final chapter first provides a summary with the main findings of current research and its contributions to the field (9.1). Next, the limitations of the study will be presented (9.2), followed by several recommendations for further research (9.3) and some pedagogical implications (9.4).

9.1. Review of findings

For the first RQ we saw that there were usually significant differences between watching the episodes either in L1S or L2S, favouring the L1S group for comprehension and the L2S group for word-form recognition. The results of the Mann-Whitney tests also indicated that both groups recognised more known than new words, as there were significant differences between these variables for both L1S and L2S learners. It is interesting to note that significant differences were not always found for each episode, and that cross-sectional comparisons between groups may overlook these differences. That is why conducting longitudinal studies helps us see patterns that would not be obvious in one-off studies.

Our results suggest that learning is taking place and that watching subtitled TV series can be beneficial as well for young low-proficiency L2 learners. However, the low L2 proficiency level of the participants might have implied challenges especially for the L2S group in understanding the written input provided in a language they have poor knowledge of. Our outcomes strongly suggest, as previous research has also pointed out, that in order to have a good understanding of the input, low-proficiency L2 learners should start watching subtitled TV series with L1S (Markham et al., 2001; Markham & Peter, 2003; Danan, 2004; Lin & Siyanova, 2015).

However, L2S favoured L2 word-form recognition more than L1S (and known words were better recognised than new words in both groups). This leads us to conclude that when the goal is identification of written L2 word forms (one of the first steps in vocabulary learning), L2S are more effective. These significant differences in favour of the L2S group can be explained by the fact that, in our study, participants were only required to identify the word form; i.e., they were not asked to provide the meaning or the translation of the lexical item. Due to that, L2S were more beneficial because they included the written L2 word form and the words selected for the tests were also in the L2. The L1S group had more problems in doing so, as they read the forms in the L1. Even though the L2S group obtained higher scores in vocabulary recognition than their L1S group peers, results were still poor for the word-form recognition variable (especially when dealing with completely new vocabulary). As previous research has found, a higher number of encounters might be needed in order for new words to be recognised (Rott, 1999; Waring & Tatak, 2003; Webb, 2007; Suárez & Gesa, 2019). Moreover, explicit teaching is also encouraged in order to enhance new TW acquisition (Gesa, 2019).

Regarding the second RQ, we can conclude that language aptitude significantly influenced the comprehension results in both groups. Findings are consistent both in the correlation and regression analyses performed. On the other hand, a correlation between aptitude and L2 word-form recognition was only found in the L1S group (it thus seems that aptitude was helpful when the task was more challenging, but was not in the case of the L2 group as they might have easily recognised word forms in the L2 appearing in the input). In relation to proficiency, analysed by means of the L1 / L2 RS scores, VS scores and the English class grade, we observed the following: first,

RS was only related to comprehension scores in the L1S group. The lack of a relationship with the comprehension scores of the L2S group is due to the fact that they were reading the subtitles in the L2 and their L2 speed was lower than their L1S, making the reading of the subtitles more difficult, as has also been shown in the questionnaire results. Regarding VS, the VST scores were related to the results obtained by both groups, although the L1S group scores showed a stronger relationship than the L2S scores. The separate analyses between the learners' VS and the known and new words scores also showed significant correlations for the L1S participants. This shows the importance of VS for episode comprehension once more, and also for learning from multimodal input at early ages. Finally, it was found that the participants' English class grade strongly influenced the comprehension outcomes (in both groups), and the correlations between this variable and vocabulary outcomes only showed significant results for the L1S group in the vocabulary total scores and the known words variable (no significant outcomes were found for L2S). The similarity of the results found for VS and English class mark should be highlighted, evidencing that VS is also closely related to general proficiency at low levels.

The third RQ allowed us to glimpse learners' opinions and perceptions, which are not often taken into account in this type of study. We observed that students in both groups acknowledged they had a better understanding of the videos at the end of the treatment (and that the L1S group said that they had a better understanding since the beginning of the intervention when compared to their L2S peers, which is corroborated by the results of the first RQ). This shows that, even for the L2S group, who seemed to struggle more with comprehension, cumulative exposure has an effect on learners, making them better at understanding what was happening in the episodes. This was

probably also due to the fact that TV series contain repetitive elements and situations that enhance comprehension. Secondly, results from the questionnaire confirmed that students experienced no changes in their watching behaviour throughout the duration of the treatment, as most of them devoted attention to reading and listening from the beginning, not favouring one over another at the beginning or end of the treatment. Furthermore, it was seen that learners in the L1S group 'could not avoid reading' (and this can be related to the automatic reading behaviour of subtitles pointed out in d'Ydewalle & Gielen, 1992; Pavakanun & d'Ydewalle, 1992, etc.), while learners in the L2S group made the effort to read as reading in the L2 was more challenging, but they needed to do it to understand the plot. Finally, information from the questionnaire also allowed us to discover that the L2S group significantly felt they were learning vocabulary (a feeling that was not shared with the L1S group). The language in the subtitles was possibly key for this outcome, as well as the vocabulary test administered. There were no participants in the L2S group indicating they were learning fewer words at the end, which was not the case for some L1S learners, who perceived they were not learning words from the episodes.

The fourth RQ provided interesting insights into the effect of OSE on the learners' L2 comprehension and vocabulary outcomes, as we clearly saw that previous exposure to multimodal input was not affecting the results for these children, contrary to what happens in previous research in non-dubbing countries. It is also true that the amount of OSE to audiovisual materials was low for these participants, even for those students claiming to watch TV/videos in original version at home. Our results for this RQ are not in line with similar studies performed in other European countries, observing a significant impact of OSE to multimodal input on participants' outcomes for

comprehension (e.g., Kuppens, 2010; Sylvén & Sundqvist, 2012) and vocabulary learning (e.g., Turgut & Irgin, 2009; Jensen, 2016). Most of the previous research on the topic has been performed in countries that do not dub the programmes they broadcast and, therefore, students are naturally exposed to this kind of material from a very early age and for long periods of time, which was not the case in our study (reflecting also the current situation in a dubbing context).

9.2. Limitations of the study

The present study is not without limitations, which should be acknowledged and taken into account when interpreting the results. We mention below some of the major limitations that we also consider should be overcome when planning future research on the topic. First, the study does not include a CG in which students watched the videos without subtitles: only two episodes (E10 and E20) from the whole intervention were watched without subtitles to control what students were able to do at these two points in time. The inclusion of a comparison group watching all the episodes without subtitles would have provided relevant information on the usefulness of written textual support in cartoon TV series. Nevertheless, the organisation of courses in the institution where the study took place, as well as the logistics of data collection made it impossible to include a third group in the study. It is not clear, either, that students at this low proficiency level could have been able to watch all episodes from the treatment without being discouraged, as they will probably have had problems to follow.

Furthermore, despite the fact that a long treatment allows us to gather interesting information on learning, we should be aware of some side-effects it may imply. First, learners got accustomed to being tested after watching each episode: they already knew the format and instructions and this routine familiarity could have altered the outcomes of the weekly sessions. Next, it should be noted that the number of participants that were always present in class was lower than we expected at the beginning of the intervention, as some of them missed some sessions and were not included in the final sample. This particularly affected the analyses in RQ4, when participants were subdivided into groups according to OSE. The limited number of learners in the analyses for this RQ makes us take the results with due caution. A study with a larger sample would definitely be needed to corroborate or reject the findings of the present dissertation.

Another limitation is that long term retention of vocabulary learned was not examined. A delayed post-test could have been devised to assess learning, although we focused here on immediate learning. It should also be acknowledged that a test of this kind could have favoured those lexical items appearing in the latest episodes, more than those learned in the first chapters of the series, as the intervention was two trimesters long.

An additional shortcoming of the study is related to the aspects of vocabulary knowledge assessed. Here we focus on word-form identification and meaning has not been tested. It would have been interesting to see whether the learners who had L1 translations in the subtitles were actually better able to make form-meaning connections than those to whom subtitles were offered in the L2. This would need to

be assessed in further research. Possibly, checking first meaning recall and immediately after meaning recognition would be the best way to explore whether and how form-meaning connections are established in the two conditions studied.

Regarding the selection of the known and new words, it should be noted that it was made taking into account the criteria of the participants' teacher, who was also the researcher of the study. As she had been their teacher since these students started to learn English, she was aware of the vocabulary that had been introduced in class so far and, thus, inferences could be made regarding which words were known and which were new. However, the fact that words had been taught at school does not actually mean that all students knew those words and, similarly, the fact that some had not been introduced does not mean that someone may have known them (even if care was taken that those learners in the study had not taken extracurricular English classes). Therefore, pre-testing would be recommended in order not to depend on the teacher's criteria in the future. However, as a large amount of target vocabulary would need to be assessed, our recommendation is that words in the pre-test appear in different subsets and students are assessed in different sessions so as to avoid fatigue effects (in addition, distractors should be added not to discourage them, increasing the amount of vocabulary that would need to be pre-tested). Finally, even though there was a wide range of lexical items available, on a few occasions we had to select words or MWUs that only appeared twice in the episode. This is problematic, as we know that we need several encounters with words in order for us to notice and learn them. However, working with real input also implies having to deal with constraints of this sort (which makes word selection difficult at some points, resulting in some episodes being discarded). All in all, however, the final sample of TWs was representative

enough of different types of vocabulary appearing in real language: nouns, verbs, adjectives, adverbs, collocation compounds, phrasal verbs, linking adverbials and free word combinations (as presented in the literature review section) to make inferences about the lexical acquisition processes.

Finally, it could also be considered that the fact that the researcher conducting the study was the teacher herself is a limitation. As we have seen in the methodology section, this may bring about advantages and disadvantages. We actually consider it an advantage that students already knew and trusted her since the beginning of the intervention (this was a key point in our case, due to the young age of the participants). The students took the tests performing at their best, as it was considered a regular class activity (no new researchers had to be introduced to students). As the teacher had known the learners for a long time, she could also predict their behaviour and any issues that could arise from the beginning of the intervention (Kosnik & Beck, 2000), which would have been difficult for an outsider.

9.3. Directions for further research

After considering the limitations pointed out in the previous section (and how they should be addressed in future studies), there are also several aspects that need to be considered for further research on language learning from multimodal input in young learners. First of all, the sample of participants was relatively small (L1S = 47; L2S = 45) and it would be of great interest to replicate the study with a larger pool of learners. Even though statistical requirements were met, a larger amount of students is recommended in order to strengthen the results. Another aspect that could not be

addressed was, as we mentioned above, the inclusion of a CG: a longitudinal research design with a CG where young low-proficiency L2 learners watched all *Curious George* episodes without subtitles would provide interesting data on the role of subtitles when young learners watch television.

In the present study, extreme care was taken in the selection of TWs, as we wanted them to be representative of the many different lexical items that naturally occur in language. Therefore, different types of lexical items were selected. It has been observed that different types of vocabulary were learned, including MWUs formed by more than one lexical item. The challenging task of learning MWUs might have been compensated with the type of input selected, which was key to the obtention of good results as has been claimed in previous research (Schmitt & Redwood, 2011; González Fernández & Schmitt, 2015; Boers, 2020). Further research should definitely investigate if different kinds of words are more readily learned than others from subtitled series, or whether frequency of exposure required to learn MWUs is different from that of single words. Probably, though, some input manipulation would be needed to study frequency effects in video watching. Other dimensions that could be assessed in relation to word knowledge could be pronunciation accuracy or productive written recall, as only passive recognition has been examined in this thesis: development in other language areas has been missed. Other specific questions that can be addressed with the data we already have are related to the way tests were corrected and to the distractors in the vocabulary tests. For example, following Waring and Tatakai (2003), wrong answers could be computed (i.e., those words that were not circled correctly) so as to see whether guessing was taking place and to which extent.

Additionally, it is interesting to see that in Suárez and Gesa (2019), proficiency was influencing the results more than aptitude (in high school and university learners). In Gesa and Suárez (2022), aptitude was influencing word-form recognition scores in primary school children. The present study has also confirmed that aptitude influences learning from multimodal input at low proficiency levels (independently of the type of subtitles that are used) but just one aptitude test, LLAMA B, was used (assessing associative memory and rote learning). A study assessing different aptitude subtypes in young learners would provide useful insights on how this variable affects learning. For example, LLAMA D (assessing listening for new words) or LLAMA E (assessing the linkage of sounds and symbols in a new language) could be used to deepen our knowledge of language aptitude and young learners. There is also the question about whether students' memory affected test results in the present study (i.e., some may have noticed the TWs but did not remember them in the test, or answered they were present in the video when they remembered them from classes, for example). Adding a WM test in the initial test battery could help elucidate these issues.

Finally, with the data obtained in the present dissertation, our questionnaires could be revised and adapted so that they are more meaningful for further studies that can be used with learners of a similar age and in similar settings. Following McNemar's analyses in RQ3, some categories could also be adjusted to facilitate researchers' work in the future. In addition, learners' preferences for specific TV genres should also be taken into account. The present work chose cartoons and its use was piloted with different samples of learners of similar ages, but sitcoms addressed to children could also constitute good input as long as coverage is adequate (e.g., as in Galimberti & Miralpeix, 2018). We think that questionnaires are also necessary in order to give

account of the learners' perceptions of the treatment. These are crucial to inform further research and take advantage of their learning potential (Pinter, 2018).

9.4. Pedagogical implications

Results in the present dissertation can be useful for researchers, but also for educators. This study has provided answers to several issues pointed out in previous research and filled in several gaps that SLA experts considered to be in need of attention. The study shows that sustained exposure to animated TV series is also beneficial for young low proficiency learners: as long as coverage is appropriate, they make sense of the input and are able to understand the plots. They are also capable of learning new word forms (as in Avello, 2023) and identify others that have previously been learned in class. However, it should also be noticed that gains are not remarkable and that incidental learning is difficult for these populations even if extra exposure is provided (the 20 episodes amount to just about 3 hours of exposure to subtitled video in total).

Therefore, some explicit instruction could be desirable so that learners make the most of the viewing experience. For example, pre-teaching could be considered for low-proficiency L2 learners in order to improve their comprehension and vocabulary outcomes. The results from the L1S learners suggest that they did not have enough input to identify all the selected words by only watching the videos once and they might have benefitted from a pre-teaching of the lexical items before viewing each episode, as Gesa (2019) did in his study. Moreover, this could be a good practice to improve the low results obtained by both groups in the new words variable. Another possibility

to maximise learning can be viewing the episodes more than once (as in Avello, 2023). Especially when subtitles are in the L2, learners at this age usually like repeated viewing of the same episode and they do not find such practice redundant.

Such a classroom activity can also prompt learners to do the same at home, empowering them to continue watching the same series (or similar ones) and this would undoubtedly facilitate exposure to the L2. The context, of course, would be different anyway from that in non-dubbing countries, but recent developments available online (e.g., Language Reactor) may help learners make the most of the input they get in the FL. Teacher training along these lines could be encouraged, as well as close collaboration between families and schools to improve the language learning experience of young students.

REFERENCES

- Abrahamsson, N., & Hyltenstam, K. (2008). The robustness of aptitude effects in near-native second language acquisition. *Studies in Second Language Acquisition* 30(4), 481-509.
- Aghlara, L., & Hadidi Tamjid, N. (2011). The effect of digital games on Iranian children's vocabulary retention in foreign language acquisition. *Procedia - Social and Behavioral Sciences*, 29, 552-560.
- Agustín Llach, M. P., & Terrazas, M. (2009). Examining the relationship between receptive vocabulary size and written skills of primary school learners. *Atlantis*, 31(1), 129-47.
- Ahmad, J. (2011). Intentional vs. incidental vocabulary learning. *Interdisciplinary Journal of Contemporary Research in Business*, 3(5), 67–75.
- Akcan, P. İ., & Demirhan, U. U. (2016). The database of cartoons and commercials as visually-oriented language input: Motivations, the procedure and methodological issues (2015), *Dilbilim Araştırmaları Dergisi*, 1, 1–15.
- Alali, F., & Schmitt, N. (2012). Teaching formulaic sequences: The same as or different from teaching single words? *TESOL Journal*, 3(2), 153-180.
- Alderson, J. C., & Banerjee, J. (2002). Language testing and assessment (Part 2). *Language Teaching*, 35(2), 79-113.
- Alexiou, T. (2015). Vocabulary uptake from Peppa Pig: A case study of preschool EFL learners in Greece. In C. Gitsaki, & T. Alexiou (Eds.), *Current issues in second/foreign language teaching and teacher development: Research and practice* (pp.285-301). Cambridge Scholars Publishing.
- Alshehri, M. (2014, Jul. 14). *The effects of frequency of exposure, elaboration, and individual differences on incidental vocabulary learning*. 9th International PG Conference in Linguistics and Language Teaching, Lancaster University.
- Anderson, D. R., & Hanson, K. G. (2009). Children, media, and methodology. *American Behavioral Scientist*, 52(8), 1204–1219.
- Arnon, I., & Snider, N. (2010). More than words: Frequency effects for multi-word phrases. *Journal of Memory and Language*, 62(1), 67–82.
- Artieda, G., & Muñoz, C. (2016). The LLAMA tests and the underlying structure of language aptitude at two levels of foreign language proficiency. *Learning and Individual Differences*, 50, 42-48.
- Avello, D. (2023). *L2 learning from captioned-video viewing in primary school students*. [Unpublished PhD dissertation. University of Barcelona, English Studies Department].

References

- Bacon, S. M., & Finnemann, M. D. (1990). A study of the attitudes, motives, and strategies of university foreign language students and their disposition to authentic oral and written input. *The Modern Language Journal*, 74(4), 459–473.
- Baddeley, A. (2003). Working memory: Looking back and looking forward. *Nature Reviews Neuroscience*, 4(10), 829-839.
- Badecker, W., & Allen, M. (2002). Morphological parsing and the perception of lexical identity: A masked priming study of stem homographs. *Journal of Memory and Language*, 47(1), 125–144.
- Bahrani, T. (2014). An overview on the effectiveness of watching cartoons as authentic language input for language learning development. *International Journal of Language Learning and Applied Linguistics World*, 5(2), 550–556.
- Bahrani, T., & Sim, T. S. (2012). Audiovisual news, cartoons and films as sources of authentic input. *The Turkish Online Journal of Educational Technology (TOJET)*, 11(4), 56-64.
- Bahrani, T., & Soltani, R. (2011). The pedagogical values of cartoons. *Research on Humanities and Social Sciences*, 1(4), 19–23.
- Baltova, I. (1999). *The effect of subtitled and staged video input on the learning and retention of content and vocabulary in a second language*. [Unpublished doctoral dissertation, University of Toronto.]
- Barclay, S. (2017, Mar. 18-21). *The effect of word class and word length on the decay of lexical knowledge*. Annual Conference of the American Association of Applied Linguistics, Portland, USA.
- Bardovi-Harlig, K. (2009). Conventional expressions as a pragmalinguistic resource: recognition and production of conventional expressions in L2 pragmatics. *Language Learning*, 59(4), 755-795.
- Bardovi-Harlig, K. (2010). Recognition of conventional expressions in L2 pragmatics. In G. Kasper, H.T. Nguyen, D.R. Yoshimi, & J.K. Yoshioka (Eds.), *Pragmatics and language learning* (Vol. 12, pp. 141-162). University of Hawai'i, National Foreign Language Resource Center.
- Bardovi-Harlig, K. (2012). Formulas, routines, and conventional expressions in pragmatics research. *ARAL*, 32, 206-227.
- Bardovi-Harlig, K., Nickels, E., & Rose, M. (2008). The influence of first language and level of development in the use of conventional expressions of thanking, apologizing, and refusing. In M. Bowles, R. Foote, S. Perpiñán, & R. Bhatt (Eds.) *Selected proceedings of the 2007 Second Language Research Forum* (pp. 113-130). Cascadilla Proceedings Project.
- Başaran, H., & Köse, G. (2012). The effects of captioning on EFL listening comprehension. *Procedia - Social and Behavioural Sciences*, 70, 702-708.

References

- Becker, S. R., & Sturm, J. L. (2017). Effects of Audiovisual Media on L2 Listening Comprehension: A Preliminary Study in French. *CALICO Journal*, 34(2), 147–177.
- Beglar, D., & Hunt, A. (2002). Implementing task-based language teaching. In J. Richards, & W. A. Renandya (Eds.), *Methodology in language teaching: An anthology of current practice* (pp. 96-106). Cambridge University Press.
- Benavent, G. T., & Peñamaría, S. S. (2011). Use of authentic materials in the ESP classroom. *Encuentro*, 89–94.
- Berwald, J.P. (1979). Teaching foreign languages by means of subtitled visuals. *Foreign Language Annals*, 12(5), 375-378.
- Berwald, J.P. (1986). *Au Courant: Teaching French vocabulary and culture using the mass media*. Center for Applied Linguistics.
- Bianchi, F., & Ciabattini, T. (2008). Captions and subtitles in EFL learning: An investigative study in a comprehensive computer environment. In A. Baldry, C. M. Pavesi, T. Torsello, & C. Taylor (Eds.), *From didactas to ecolingua* (pp. 69–80). Edizioni Università di Trieste.
- Biber D., Johansson, S., Leech, G., Conrad, S., & Finegan, E. (1999). *The Longman grammar of spoken and written English*. Longman.
- Bialystok, E. (1997). The structure of age: In search of barriers to second language acquisition. *Second Language Research* 13(2), 116-137.
- Bird, S. A., & Williams, J. N. (2002). The effect of bimodal input on implicit and explicit memory: An investigation into the benefits of within-language subtitling. *Applied Psycholinguistics*, 23(4), 509–533.
- Birdsong, D. (1999). Introduction: Whys and why nots of the critical period hypothesis for second language acquisition. In D. Birdsong (Ed.), *Second language acquisition and the critical period hypothesis* (pp. 1-22). Lawrence Erlbaum Associates.
- Birulés-Muntané, J., & Soto-Faraco, S. (2016). Watching subtitled films can help learning foreign languages. *PLoS ONE* 11(6), 1-10.
- Bisson, M. J., van Heuven, W. J. B., Conklin, K., & Tunney, R. J. (2014). The role of repeated exposure to multimodal input in incidental acquisition of foreign language vocabulary. *Language Learning*, 64(4), 855–877.
- Bland, J. (2018). Learning through literature. In S. Garton, & F. Copland (Eds.) *The Routledge handbook of teaching English to Young Learners*. (pp.269-287). Routledge.
- Bod, R. (2000). *The storage vs. computation of three-word sentences*. AMLaP2000, University of Leiden, Leiden, the Netherlands.

References

- Bod, R. (2001). *Sentence memory: Storage vs. computation of frequent sentences*. 14th Annual CUNY Conference on Human Sentence Processing. University of Pennsylvania, Philadelphia, PA.
- Boers, F. (2020). Factors affecting the learning of multiword items. In S. Webb (Ed.), *The Routledge handbook of vocabulary studies* (pp. 158-173). Routledge.
- Bokander, L., & Bylund, E. (2020). Probing the internal validity of the LLAMA language aptitude tests. *Language Learning*, 70(1), 11-47.
- Bongaerts T., Van Summeren, C., Planken, B., & Schils, E. (1997). Age and ultimate attainment in the pronunciation of a foreign language. *Studies in Second Language Learning* 19(4), 447-465.
- Borgia, E.T., & Schuler, D. (1996). *Action research in early childhood education*, University of Illinois.
- Borrás, I., & Lafayette, R. C. (1994). Effects of multimedia courseware subtitling on the speaking performance of college students of French. *Modern Language Journal*, 78(1), 61–75.
- Bradbard, D. A., Parker, D. F., & Stone, G. L. (2004). An alternate multiple-choice scoring procedure in a macroeconomics course. *Decision Sciences Journal of Innovative Education*, 2(1), 11-26.
- Bravo, M. (2008). *Putting the reader in the Picture: screen translation and foreign-language learning*. [Unpublished Doctoral Dissertation, Universitat Rovira i Virgili]
- Brett, P. (1995). Multimedia for listening comprehension: The design of a multimedia-based resource for developing listening skills. *System*, 23(1), 77–85.
- Brett, P. (1997). A comparative study of the effects of the use of multimedia on listening comprehension. *System*, 25(1), 39–53.
- Brindley, G. (1998). Assessing listening abilities. *Annual Review of Applied Linguistics*, 18(1), 171–191.
- Brown, D. (2001). *Teaching by principle—an interactive approach to language pedagogy*. Addison Wesley Longman.
- Brown, R., Waring, R., & Donkaewbua, S. (2008). Incidental vocabulary acquisition from reading, reading-while-listening, and listening to stories. *Reading a Foreign Language* 20(2), 136-163.
- Buck, G. (2001). *Assessing listening*. Cambridge University Press.
- Butler, Y. G. (2019). Gaming and young learners. In S. Garton, & F. Copland *The Routledge handbook of teaching English to Young Learners*. (pp.305-319). Routledge.

References

- Caimi, A. (2006). Audiovisual translation and language learning: The promotion of intralingual subtitles. *The Journal of Specialised Translation*, 6, 85–98.
- Cameron, L. (2002). Measuring vocabulary size in English as an additional language. *Language Teaching Research*, 6(2), 145–173.
- Canning-Wilson, C. (2000). Practical aspects of using video in the foreign language classroom. *The Internet TESL Journal*, 6(11). Retrieved from <http://iteslj.org/Articles/Canning-Video.html>
- Carroll, J. B. (1962). *The prediction of success in intensive foreign language training*. Training Research and Education, University of Pittsburgh Press.
- Carroll, J. B. (1973). Implications of aptitude test research and psycholinguistic theory for foreign-language teaching. *International Journal of Psycholinguistics*, 2, 5–14.
- Carroll, J. B. (1981). Twenty-five years of research on foreign language aptitude. In K. C. Diller (Ed.), *Individual differences and universals in language learning aptitude* (pp. 83-118). Newbury House.
- Carroll, J. B. (1989). The Carroll Model: A 25-year retrospective and prospective view. *Educational Researcher*, 18(1), 26-31.
- Carroll, J. B., & Sapon, S. (1959). *Modern Language Aptitude Test: Form A*. The Psychological Corporation.
- Carroll, J. B., & Sapon, S. (1967). *Modern Language Aptitude Test-Elementary*. The Psychological Corporation.
- Casulleras, M. (2014). *Electronic materials and vocabulary learning in the primary EFL classroom* [Unpublished Master's Thesis, English Studies Department, Applied Linguistics Program, Universitat de Barcelona].
- Cenoz, J. (2002). Age differences in foreign language learning. *ITL-International Journal of Applied Linguistics*, 135(1), 125-142.
- Charles, T. J., & Trenkic, D. (2015). Speech segmentation in a second language: the role of bi-modal input. In Y. Gambier, A. Caimi, & C. Mariotti (Eds.), *Subtitles and language learning: Principles, strategies and practical experiences* (pp. 173-198). Peter Lang.
- Chen, M. (2013). Phrasal verbs in a longitudinal learner corpus: Quantitative findings. In S. Granger, G. Gilquin, & F. Meunier (Eds.) *Twenty years of learner corpus research: Looking back, moving ahead. Corpora and Language in Use – Proceedings 1* (pp.89-101). Presses universitaires de Louvain.
- Chen, Y. (2016). Dictionary use for collocation production and retention: A CALL-based study. *International Journal of Lexicography*, 30(2), 225–251.

References

- Chun, D. M., & Plass, J. (1996). Effects of multimedia annotations on vocabulary acquisition. *The Modern Language Journal*, 80(2), 183–198.
- Chun, D. M., & Plass, L. P. (1997). Research on text comprehension in multimedia environments. *Language, Learning and Technology*, 1(1), 60-81.
- Chung, J. (1999). The effects of using video texts supported with advance organizers and captions on Chinese college students' listening comprehension: An empirical study. *Foreign Language Annals*, 32(3), 295-308.
- Clarke, D. (1989). Communicative theory and its influence on materials production. *Language Teaching* 22(2), 73–86.
- Coady, J. (1997). L2 vocabulary acquisition through extensive reading. In J. Coady, & T. Huckin (Eds.), *Second Language Vocabulary Acquisition* (pp. 225-237). Cambridge University Press.
- Cobb, T. (ongoing). VocabProfile (version 2) [Computer software]. Retrieved from: <https://www.lexutor.ca/vp/comp/>
- Cobb, T., & Horst, M. (1999). Vocabulary sizes of some city university students. *Journal of the Division of Language Studies of the City University of Hong Kong*, 1(1), 59-68.
- Collins, M. F. (2010). ELL preschoolers' English vocabulary acquisition from storybook reading. *Early Childhood Research Quarterly* 25(1), 84–97.
- Concord Evaluation Group (2012). *Evaluation of Curious George*. Retrieved from https://www-tc.pbs.org/parents/curiousgeorge/program/pdf/Curious_George_Evaluation.pdf
- Conklin, K., & Schmitt, N. (2012). The processing of formulaic language. *Annual Review of Applied Linguistics*, 32, 45-61.
- Crystal, D., & Davy, D. (1975). *Advanced Conversational English*. Longman.
- Curtis, A. (2015). Learning English through the language of the Pokémon: “Just watching ‘TV, Ma’”. In D. Nunan, & J. C. Richards (Eds.), *Language learning beyond the classroom* (pp. 129-138). Routledge.
- Dahlen, K., & Caldwell-Harris, C. (2013). Rehearsal and aptitude in foreign vocabulary learning. *The Modern Language Journal*, 97(4), 902–916.
- Danan, M. (1992). Reversed subtitling and dual coding theory: New directions for foreign language instruction. *Language Learning*, 42(4), 497-527.
- Danan, M. (2004). Captioning and subtitling: Undervalued language learning strategies. *Meta: Translators' Journal*, 49(1), 67–77.

References

- Danan, M. (2015). Subtitling as a language learning tool: Past findings, current applications, and future paths. In Y. Gambier, A. Caimi, & C. Mariotti (Eds.), *Subtitles and language learning* (pp.63-82). Peter Lang.
- De Bot, K., Jagt, J., Janssen, H., Kessels, E., & Schils, E. (1986). Foreign television and language maintenance. *Second Language Research*, 2(1), 321-328.
- DeKeyser, R. M. (1998). Beyond focus on form: Cognitive perspective on learning and practical second language grammar. In C. Doughty, & J. Williams (Eds.), *Focus on form in classroom second language acquisition* (pp.42–63). Cambridge University Press.
- DeKeyser R. M. (2000). The robustness of critical period effects in second language acquisition. *Studies in Second Language Acquisition* 22(4), 499-533.
- Dekeyser, R. M. (2008). Implicit and explicit learning. In C. J. Doughty & M. H. Long (Eds.), *The handbook of second language acquisition*. Blackwell.
- DeKeyser, R. M., Alfi-Shabtay, I., & Ravid, D. (2010). Cross-linguistic evidence for the nature of age effects in second language acquisition. *Applied Psycholinguistics*, 31(3), 413–438.
- DeKeyser, R. M., & Larson-Hall, J. (2005). What does the Critical Period really mean? In J. F. Kroll, & A. M. B. DeGroot (Eds.), *Handbook of bilingualism: Psycholinguistic perspectives* (pp. 88–108). Oxford University Press.
- De Wilde, V., & Eyckmans, J. (2017). Game on! Young learners' incidental language learning of English prior to instruction. *Studies in Second Language Learning and Teaching*, 7(4), 673-694.
- Díaz-Cintas, J., & Remael, R. (2007). *Audiovisual translation: Subtitling*. St. Jerome.
- Dockrell, J. E., Braisby, N., & Best, R. M. (2007). Children's acquisition of science terms: Simple exposure is insufficient. *Learning and Instruction*, 17(6), 577–594.
- Dörnyei, Z. (2005). *The psychology of the language learner*. Lawrence Erlbaum.
- Dörnyei, Z., & Skehan, P. (2003). Individual differences in second language learning. In C. J. Doughty, & M. H. Long (Eds.), *The handbook of second language acquisition*. Blackwell.
- Doughty, C. (2019). Cognitive language aptitude. *Language Learning*, 69(1), 101–126.
- Durbahn, M., Rodgers, M.P.H. & Peters, E. (2020). The relationship between vocabulary and viewing comprehension. *System*, 88, 1-13.
- d'Ydewalle, G., & De Bruycker, W. (2007). Eye movements of children and adults while reading television subtitles. *European Psychologist*, 12(3), 196–205.

References

- d'Ydewalle, G., & Gielen, I. (1992). Attention allocation with overlapping sound, image, and text. In K. Rayner (Ed.), *Eye movements and visual cognition* (pp. 415–427). Springer.
- d'Ydewalle, G., & Pavakanun, U. (1995). Acquisition of a second/foreign language by viewing a television program. In P. WinterhoffSpurk (Ed.), *Psychology of media in Europe: The state of the art, perspectives for the future* (pp. 51-64). Westdeutscher Verlag.
- d'Ydewalle, G., & Pavakanun, U. (1997). Could enjoying a movie lead to language acquisition? In P. WinterhoffSpurk, & T. H. A. van der Voort (Eds.), *New horizon in media psychology* (pp. 145-155). Westdeutscher Verlag.
- d'Ydewalle, G., & Van De Poel, M. (1999). Incidental foreign-language acquisition by children watching subtitled television programs. *Journal of Psycholinguistic Research*, 28(3), 227–245.
- d'Ydewalle, G., & Van Rensbergen, J. (1989). Developmental studies of text–picture interactions in the perception of animated cartoons with text. In M. Heinz, & J. R. Levin (Eds.), *Knowledge acquisition from text and pictures* (pp. 313-321). Elsevier.
- d'Ydewalle, G., Van Rensbergen, J., & Pollet, J. (1987). Reading a message when the same message is available auditorily in another language: The case of subtitling. In J. K. O'Reagan, & A. Lévy Schoen (Eds.), *Eye movements: From physiology to cognition* (pp. 313–321). Elsevier.
- Ebel, R. L. (1975). Can classroom teachers write good true-false test items? *Journal of Educational Measurement*, 12(1), 31-35.
- Ebel, R. L., & Frisbie, D. A. (1991). *Essentials of education measurement*. Prentice Hall.
- Ebeling, S. O., & Hasselgård, H. (2015). Learner corpora and phraseology. In S. Granger, G. Gilquin, & F. Meunier (Eds.), *The Cambridge handbook of learner corpus research* (pp. 185–206). Cambridge University Press.
- Ebbinghaus, H. (1913/1885). *Memory: A contribution to experimental psychology*. Teachers College, Columbia University.
- Ebrahimi, Y., & Bazaee, P. (2016). The effect of watching English movies with standard subtitles on EFL learners' content and vocabulary comprehension. *Journal of Applied Linguistics and Language Research*, 3(5), 284-295.
- Edmondson, W., & House, J. (1991). Do learners talk too much? The waffle phenomenon in interlanguage pragmatics. In R. Phillipson, E. Kellerman, L. Selinker, M. Sharwood Smith, & M. Swain (Eds.), *Foreign/second language pedagogy research: A commemorative volume for Claus Faerch* (pp. 273-286). Multilingual Matters.

References

- Ehrman, M. E., & Oxford, R. L. (1995). Cognition plus: Correlates of language learning success. *The Modern Language Journal*, 79, 67-89.
- Elley, W. B. (1980). A comparison of content-interest and structuralist reading programs in Niue primary schools. *New Zealand Journal of Educational Studies*, 15, 39-53.
- Elley, W. B. (1989). Vocabulary acquisition from listening to stories. *Reading Research Quarterly*, 24(2), 174–187.
- Elley, W. B., & Mangubhai, F. (1983). The impact of reading on second language learning. *Reading Research Quarterly*, 19(1), 53–67.
- Ellis, N. C. (1996). Sequencing in SLA: Phonological memory, chunking, and points of order. *Studies in Second Language Acquisition*, 18(1), 91–126.
- Ellis, N. C. (2002). Frequency effects in language acquisition: A review with implications for theories of implicit and explicit language acquisition. *Studies in Second Language Acquisition*, 24(2), 143–188.
- Ellis, N. C. (2016). Salience, cognition, language complexity and complex adaptive systems. *Studies in Second Language Acquisition*, 38(2), 341-351.
- Ellis, R. (1994). Introduction: Implicit and explicit language learning: An overview. In N. Ellis (Ed.), *Implicit and explicit learning of languages* (pp. 1-13). Academic Press.
- Ellis, R. (2005). Measuring implicit and explicit knowledge of a second language: A psychometric study. *Studies in Second Language Acquisition*, 27(2), 141–172.
- Ellis, R. (2008). Explicit knowledge and second language learning and pedagogy. In J. Cenoz, & N. H. Hornberger (Eds.), *Encyclopedia of language and education*, 6, (pp.143-153). Springer.
- Etemadi, A. (2012). Effects of bimodal subtitling of English movies on content comprehension and vocabulary recognition. *International Journal of English Linguistics*, 2(1), 239–248.
- European Commission. (2008). *Multilingualism: An asset for Europe and a shared commitment*. (COM(2008) 566 final) Retrieved from <https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:52008DC0566&from=EN>
- European Commission. (2011). *Study on the use of subtitling. The potential of subtitling to encourage foreign language learning and improve the mastery of foreign languages* (Report No. EACEA/2009/01). Retrieved from: <https://publications.europa.eu/en/publication-detail/-/publication/e4d5cbf4-a839-4a8a-81d0-7b19a22cc5ce>
- Frenck-Mestre, C. (2005). Eye-movement recording as a tool for studying syntactic processing in a second language: A review of methodologies and experimental findings. *Second Language Research*, 21(2), 175–198.

References

- Fresno, N. (2018). Watching accessible cartoons: The speed of closed captions for young audiences in the United States, *Perspectives*, 26(3), 405-421.
- Frumuselu, A. D., De Maeyer, S., Donche, V., & Colon Plana, M. del M. G. (2015). Television series inside the EFL classroom: Bridging the gap between teaching and learning informal language through subtitles. *Linguistics and Education*, 32(B), 107–117.
- Fusté, J. (2013). *Characterising the lexical knowledge of EFL learners at the end of primary education* [Unpublished Master's Thesis, English Studies Department, Applied Linguistics Program, Universitat de Barcelona].
- Galimberti, V., & Miralpeix, I. (2018). Multimodal input for Italian beginner learners of English. A study on comprehension and vocabulary learning from undubbed TV series. In C. M. Coonan, & A. Bier, & E. Ballarin (Eds.), *La didattica delle lingue nel nuovo millennio. Le sfide dell'internazionalizzazione* (pp. 615-626). Edizioni Ca'Foscari.
- García Mayo, M. P., & García Lecumberri, M. L. (2003). *Age and the acquisition of English as a foreign language*. Multilingual Matters.
- Garnier, M., & Schmitt, N. (2016). Picking up polysemous phrasal verbs: How many do learners know and what facilitates this knowledge? *System*, 59, 29–44.
- Garza, T. J. (1991). Evaluating the use of captioned video materials in advanced foreign language learning. *Foreign Language Annals*, 24(3), 239–258.
- Gass, S. M. (1999). Incidental Vocabulary Learning. *Studies in Second Language Acquisition*, 21(2), 319–333.
- Gass, S. M., & Mackey, A. (2007). Input, interaction, and output in second language acquisition. In B. Vanpatten, & J. Williams (Eds.), *Theories in second language acquisition* (pp. 175-200). LEA.
- Gesa, F. (2019). *L1 / L2 subtitled TV series and EFL learning: A study on vocabulary acquisition and content comprehension at different proficiency levels*. [Unpublished doctoral dissertation. University of Barcelona].
- Gesa, F., & Suárez, M. M. (2022). Individual differences in young learners' L2 vocabulary development: The case of language aptitude and exposure to subtitled TV series. In V. De Wilde, & C. Goriot. (Eds.), *Second language learning before adulthood* (pp. 65-86). DeGruyter.
- Ghia, E. (2012). *Subtitling matters: New perspectives on subtitling and foreign language learning*. Peter Lang.
- Gielen, M. (1988). *Perception and subtitles: The parafoveal and peripheral information processing of subtitles*. University of Leuven.
- Gilmore, A. (2007). Authentic materials and authenticity in foreign language learning. *Language Teaching*, 40(2), 97-118.

References

- Goh, C. C. M. (2000). A cognitive perspective on language learners listening comprehension problems, *System*, 28(1), 55–75.
- Goldschneider, J. M., & DeKeyser, R. (2001). Explaining the “natural order of L2 morpheme acquisition” in English: A meta-analysis of multiple determinants. *Language Learning*, 51(1), 1–50.
- González Fernández, B., & Schmitt, N. (2015). How much collocation knowledge do L2 learners have?: The effects of frequency and amount of exposure. *ITL*, 166(1), 94-126.
- Gowhary, H., Pourhalashi, Z., Jamalinesari, A., & Azizifar, A. (2015). Investigating the effect of video captioning on Iranian EFL learners' listening comprehension. *Procedia - Social and Behavioral Sciences*, 192, 205–212.
- Granger, S., & Paquot, M. (2008). Disentangling the phraseological web. In S. Granger, & F. Meunier (Eds.) *Phraseology: An Interdisciplinary Perspectives* (pp. 27-49). John Benjamins.
- Grañena, G. (2012). *Age differences and cognitive aptitudes for implicit and explicit learning in ultimate second language attainment*. [Unpublished doctoral dissertation. University of Maryland].
- Grañena, G. (2013a). Cognitive aptitudes for second language learning and the LLAMA Language Aptitude Test. In G. Grañena, & M. Long (Eds.) *Sensitive periods, language aptitude, and ultimate L2 attainment* (pp.105-129). John Benjamins.
- Grañena, G. (2013b). Individual differences in sequence learning ability and second language acquisition in early childhood and adulthood. *Language Learning*, 63(4), 665-703.
- Grañena, G., & Long, M. H. (2012) Age of onset, length of residence, language aptitude, and ultimate L2 attainment in three linguistic domains. *Second Language Research*, 29(3), 311–343.
- Grigorenko, E., Sternberg, R. J., & Ehrman, M. (2000). A theory based approach to the measurement of foreign language learning ability: The Canal-F theory and test. *Modern Language Journal*, 84(3), 390-405.
- Guichon, N., & McLornan, S. (2008). The effects of multimodality on L2 learners: Implications for CALL resource design. *System*, 36(1), 85–93.
- Guillory, H. G. (1998). The effects of keyword captions to authentic French video on learner comprehension. *Calico Journal*, 15(1/3), 89-108.
- Haladyna, T. M., Downing, S. M., & Rodriguez, M. C. (2002). A Review of Multiple-Choice Item-Writing Guidelines for Classroom Assessment, *Applied Measurement in Education*, 15(3), 309-334.
- Halliday, M. A. K. (1989). *Spoken and written language*. Oxford University Press.

References

- Hammersley, M. (1993). On the teacher as researcher. *Educational Action Research*, 1(3), 425-445.
- Hanf, A. (2015). Resourcing authentic language in television series. In D. Nunan, & J.C. Richards (Eds.) *Language Learning Beyond the Classroom* (pp. 138-148). Routledge.
- Harley, B., & Hart, D. (1997). Language aptitude and second language proficiency in classroom learners of different starting ages. *Studies in Second Language Acquisition*, 19(3), 379-400.
- Harper Collins Publishers (n/d). Reading speed. In *Collins English Dictionary*. Retrieved July, 10, 2022, from <https://www.collinsdictionary.com/dictionary/english/reading-speed>
- Hatim, B., & Mason, I. (2000). Politeness in screen translating. In I. Venuti (Ed.). *The Translation Studies reader* (pp. 433 - 445). Routledge.
- Hayati, A., & Mohmedi, F. (2011). The effect of films with and without subtitles on listening comprehension of EFL learners. *British Journal of Educational Technology*, 42(1), 181–192.
- Hemmati, F., & Ghaderi, E. (2014). The effect of four formats of multiple-choice questions on the listening comprehension of EFL learners. *Procedia – Social and Behavioral Sciences*, 98, 637-644.
- Herron, C., Hanley, J., & Cole, S. (1995). A comparison study of two advance organizers for introducing beginning foreign language students to video. *Modern Language Journal*, 79(3), 387-394.
- Higgins, J., & Wallace, R. (1989). Hopalong: A computer reader pacer. *System*, 17(3), 389–399.
- Hirsch, D., & Nation, I. S. P. (1992). What vocabulary size is needed to read unsimplified texts for pleasure? *Reading in a Foreign Language*, 8(2), 689–696.
- Holmes, G. (2008, November 24). Americans can't get enough of their screen time [Press Release]. The Nielsen Company.
- Horst, M. (2005). Learning L2 vocabulary through extensive reading: A measurement study. *The Canadian Modern Language Review*, 61(3), 355-382.
- Horst, M., Cobb, T., & Meara, P. (1998). Beyond a clockwork orange: Acquiring second language vocabulary through reading. *Reading in a Foreign Language*, 11(2), 207–223.
- Hsu, C.-K., Hwang, G.-J., Chang, Y.-T., & Chang, C.-K. (2013). Effects of video caption modes on English listening comprehension and vocabulary acquisition using handheld devices. *Educational Technology and Society*, 16(1), 403–414.

References

- Huang, H. C., & Eskey, D. E. (1999). The effects of closed-captioned television on the listening comprehension of intermediate English as a Second Language (ESL) students. *Journal of Educational Technology Systems, 28*(1), 75–96.
- Huckin, T., & Coady, J. (1999). Incidental vocabulary acquisition in a second language: A review. *Studies in Second Language Acquisition, 21*(2), 181–193.
- Hui, W. (2007). The effects of captions on Chinese EFL students' incidental vocabulary acquisition. *CELEA Journal, 30*(4), 9-16.
- Hughes, A. (2003). *Testing for language teachers*. Cambridge University Press.
- Hulstijn, J. H. (1992). Retention of inferred and given word meanings: Experiments in incidental word learning. In P. J. L. Arnaud, & H. Béjoint (Eds.), *Vocabulary and Applied Linguistics* (pp. 113-125). Macmillan.
- Hulstijn, J. H. (2001). Intentional and incidental second language vocabulary learning: A reappraisal of elaboration, rehearsal and automaticity. In P. Robinson (Ed.), *Cognition and second language instruction* (pp. 258–286). Cambridge University Press.
- Hulstijn, J. H., Hollander, M., & Greidanus, T. (1996). Incidental vocabulary learning by advanced foreign language students: The influence of marginal glosses, dictionary use, and reoccurrence of unknown words. *The Modern Language Journal, 80*(3), 327–339.
- Hummel, K. M. (2009). Aptitude, phonological memory, and second language proficiency in non-novice adult learners. *Applied Psycholinguistics, 30*(2), 225–249.
- Hunt, A., & Beglar, D. (2005). A framework for developing EFL reading vocabulary. *Reading in a Foreign Language, 17*(1), 23-59.
- Hwang, K., & Nation, I. S. P. (1989). Reducing the vocabulary load and encouraging vocabulary learning through reading newspapers. *Reading in a Foreign Language, 6*(1), 323-335.
- In'nami, Y., & Koizumi, R. (2009). A meta-analysis of test format effects on reading and listening test performance: Focus on multiple-choice and open-ended formats. *Language Testing, 26*(2), 219–244.
- Ioup, G., Boustagui, E., Tigi, M., & Moselle, M. (1994). Reexamining the Critical Period Hypothesis: A case study of successful adult SLA in a naturalistic environment. *Studies in Second Language Acquisition 16*(1), 73–98.
- Jensen, L. (1986). Advanced reading skills in a comprehensive course. In F. Dubin, D. Eskey, & W. Grabe (Eds.). *Teaching second language reading for academic purpose*. Wesley Publishing.
- Jensen, S. H. (2016). Gaming as an English language learning resource among young children in Denmark. *CALICO Journal, 34*(1), 1-19.

References

- Jiménez Catalán, R. M., & Ruíz de Zarobe, Y. (2009). The receptive vocabulary of EFL learners in two instructional contexts: CLIL versus non-CLIL instruction. In Y. Ruiz de Zarobe, & R.M. Jiménez (Eds.), *Content and language integrated learning. Evidence from research in Europe*. (pp. 81-92). Multilingual Matters.
- Jiménez Catalán, R. M., & Terrazas Gallego, M. (2008). The Receptive Vocabulary of English Foreign Language Young Learners. *Journal of English Studies*, 5, 173–191.
- Johnson, J., & Newport, E. L. (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(1), 60–99.
- Jones, L. C., & Plass, J. L. (2002). Supporting listening comprehension and vocabulary acquisition in French with multimedia annotations. *The Modern Language Journal*, 86(4), 546–561.
- Juhász, B. (2007). The influence of semantic transparency on eye movements during English compound word recognition. In R. van Gompel, M. Fischer, W. Murray, & R. Hill (Eds.), *Eye movements: A window on mind and brain* (pp. 374–389). Elsevier.
- Kaan, E., Ballantyne, J., & Wijnen, F. (2015). Effects of reading speed on second-language sentence processing. *Applied Psycholinguistics*, 36(4), 799-830.
- Karakas, A., & Sariçoban, A. (2012). The impact of watching subtitled animated cartoons on incidental vocabulary learning of ELT students. *Teaching English with Technology*, 12(4), 3–15.
- Karamitroglou, F. (1998). A proposed set of subtitling standards for Europe. *Translation Journal*, 2(2). Retrieved from <http://translationjournal.net/journal/04stndrd.htm>
- Khan, M. A. (2015). Using films in the ESL classroom to improve communication skills of non-native learners. *International Journal of Teachers of English*, 52(5), 46–52.
- King, J. (2002). Using DVD feature film in the EFL classroom. *ELT Newsletter. Article 88*. Retrieved from https://www.moviesgrowenglish.com/PDF/CATESOL/Films_for_EFL,%20J.King.pdf
- Kokla, N. (2016). Dora the Explorer: A TV character or a preschoolers' foreign language teacher? *Selected Papers of the 21st International Symposium on Theoretical and Applied Linguistics (ISTAL 21)*, 666-683.
- Koolstra, C. M., & Beentjes, J. W. J. (1999). Children's vocabulary acquisition in a foreign language through watching subtitled television programs at home. *Educational Technology Research and Development*, 47(1), 51–60.
- Koolstra, C. M., Peeters, A. L., & Spinhof, H. (2002). The pros and cons of dubbing and subtitling. *European Journal of Communication*, 17(3), 325–354.

References

- Koolstra, C. M., Van der Voort, T., & d'Ydewalle, G. (1999). Lengthening the presentation time of subtitles on television: Effects on children's reading time and recognition. *Communication & Medicine, 24*(4), 407-422.
- Koolstra, C. M.; Van der Voort, T. H. A., & Van der Kamp, L. J. T. (1997). Television's impact on children's reading comprehension and decoding skills: a 3-year panel study. *Reading Research Quarterly 32*(2), 128-152.
- Kosnik, C., & Beck, C. (2000). The action research process as a means of helping student teachers understand and fulfil the complex role of the teacher. *Educational Action Research, 8*(1), 115-136.
- Kost, C. R., Foss, P., & Lenzini, J. J. (1999). Textual and pictorial glosses: Effectiveness on incidental vocabulary growth when reading in a foreign language. *Foreign Language Annals, 32*(1), 89–97.
- Krashen, S. D. (1982). *Principles and practice in second language acquisition*. Pergamon Press.
- Krashen, S. D. (1985). *The input hypothesis: Issues and implications*. Longman.
- Krashen, S. D., Long, M. H., & Scarcella, R. C. (1979). Age, rate, and eventual attainment in second language acquisition. *TESOL Quarterly, 13*(4), 573–582.
- Kramsch, C. (1983). Culture and constructs: Communicating attitudes and values in the foreign language classroom. *Foreign Language Annals, 16*(6), 437-448.
- Kremmel, B., Brunfaut, T., & Alderson, J. C. (2017). Exploring the role of phraseological knowledge in foreign language reading. *Applied Linguistics, 38*(6), 848–870.
- Kruger, J.-L., Hefer-Jordaan, E., & Matthew, G. (2013). Measuring the impact of subtitles on cognitive load: Eye tracking and dynamic audiovisual texts. In *Proceedings of the 2013 Conference on Eye Tracking South Africa* (pp. 62-66). Association for Computing Machinery.
- Kuppens, A. H. (2010). Incidental foreign language acquisition from media exposure. *Learning, Media and Technology, 35*(1), 65–85.
- Latifi, M., Mobalegh, A., & Mohammadi, E. (2011). Movie subtitles and the improvement of listening comprehension ability: Does it help? *The Journal of Language Teaching and Learning, 1*(2), 18–29.
- Lambert, W. E., & Holobow, N. E. (1984). Combinations of printed script and spoken dialogue that show promise for students of a foreign language. *Canadian Journal of Behavioral Science/Revue Canadienne des Sciences du Comportement, 16*(1), 1–11.
- Laufer, B. (1989). What percentage of text lexis is essential for comprehension? In C. Lauren, & M. Nordman (Eds.), *Special language: From humans thinking to thinking machines* (pp.316–323). Multilingual Matters.

References

- Laufer, B. (1991). Some properties of the foreign language learner's lexicon as evidenced by lexical confusions. *International Review of Applied Linguistics in Language Teaching*, 29(4), 317-348.
- Laufer, B. (1997). The lexical plight in second language reading: Words you don't know, words you think you know, and words you can't guess. In J. Coady, & T. Huckin (Eds.), *Second language vocabulary acquisition: A rationale for pedagogy* (pp. 20–34). Cambridge University Press.
- Laufer, B. (2001). Reading, word-focused activities and incidental vocabulary acquisition in a second language. *Prospect*, 16, 44-54.
- Laufer, B. (2003). Vocabulary acquisition in a second language: Do learners really acquire most vocabulary by reading? Some empirical evidence. *Canadian Modern Language Review*, 59(4), 567–587.
- Laufer, B. (2005). Focus on form in second language vocabulary learning. *EUROSLA Yearbook*, 5(1), 223 –250.
- Laufer, B., & Nation, P. (1995). Vocabulary size and use: Lexical richness in L2 written production. *Applied Linguistics*, 16(3), 307-322.
- Laufer, B., & Paribakht, T. S. (1998). The relationship between passive and active vocabularies: Effects of language learning context. *Language Learning*, 48(3), 365–391.
- Lavaur, J., & Bairstow, D. (2011). Languages on the screen: Is film comprehension related to the viewers' fluency level and to the language in the subtitles. *International Journal of Psychology*, 46(6), 455–462.
- Lefever, S. (2010, Dec.). *English skills of young learners in Iceland: "I started talking English when I was 4 years old. It just bang... just fall into me."*. Menntakvika Conference, Reijkjavik, Iceland.
- Lenneberg, E. H. (1967). *Biological foundations of language*. New York: Wiley.
- Lekkai, I. (2014). Incidental foreign language acquisition by children watching subtitled television programs. *The Turkish Online Journal of Educational Technology* (TOJET), 13(4), 81-87.
- Lertola, J. (2015). Subtitling in language teaching: Suggestions for language teachers. In Y. Gambier, A. Caimi, & C. Mariotti (2015). *Subtitles and language learning* (pp.245-268). Peter Lang.
- Lévesque, M. (2013). *Incidental vocabulary acquisition through aural means: What do English television programs have to offer?* [Unpublished MA Thesis. Concordia University].

References

- Li, S. (2015). The associations between language aptitude and second language grammar acquisition: A meta-analytic review of five decades of research. *Applied Linguistics*, 36(3), 385-408.
- Li, S. (2019). Six decades of language aptitude research: A comprehensive and critical review. In Z. Wen, P. Skehan, A. Biedron, S. Li, & R. Sparks (Eds.), *Language aptitude: Advancing theory, testing, research, and practice* (pp.78-96). Routledge.
- Liao, Y., & Fuyuka, Y. J. (2004). Avoidance of phrasal verbs: the case of Chinese learners of English. *Language Learning*, 54(2), 193-226.
- Lightbown, P. M., & Spada, N. (2006) *How languages are learned*. Oxford University Press.
- Lin, P. M. S., & Siyanova-Chanturria, A. (2015). Internet television for L2 vocabulary learning. In D. Nunan, & J.C. Richards (Eds.), *Language learning beyond the classroom* (pp. 149-158). Routledge.
- Lindgren, E., & Muñoz, C. (2013). The influence of exposure, parents, and linguistic distance on young European learners' foreign language comprehension. *International Journal of Multilingualism*, 10(1), 105-129.
- Liu, D. (2008). Linking adverbials: An across-register corpus study and its implications. *International Journal of Corpus Linguistics*, 13(4), 491-518.
- Liu, D. (2011). The most frequently used English phrasal verbs in American and British English: A multicorpus examination. *TESOL Quarterly*, 45(4), 661-688.
- Mackin, R. (1978). On collocations: Words shall be known by the company they keep. In P. Strevens (Ed.), *In honour of A. S. Hornby* (pp.149–165). Oxford University Press.
- Maleki, A., & Safaee Rad, M. (2011). The effect of visual and textual accompaniments to verbal stimuli on the listening comprehension test performance of Iranian high and low proficient EFL learners. *Theory and Practice in Language Studies*, 1(1), 28–36.
- Markham, P. L. (1999). Captioned videotapes and second-language listening word recognition. *Foreign Language Annals*, 32(3), 321–328.
- Markham, P., & Peter, L. A. (2003). The influence of English language and Spanish language on foreign language listening/reading comprehension. *Journal of Educational Technology Systems*, 31(3), 331–341.
- Markham, P., Peter, L. A., & McCarthy, T. J. (2001). The effects of native language vs target language captions on foreign language students' dvd video comprehension. *Foreign Language Annals*, 34(5), 439–445.

References

- Martinez, A. (2002). Authentic materials: An overview. Free Resources for Teachers and Students of English, *Karen's Linguistics Issues*, 1-7.
- Martinez, R., & Schmitt, N. (2012). A phrasal expressions list. *Applied Linguistics*, 33(3), 299-320.
- Marzá, A., & Torralba, G. (2015) Incidental language learning through subtitled cartoons: Is it possible in a dubbing country? In Y. Gambier, A. Caimi, & C. Mariotti, (2015). *Subtitles and language learning* (pp.199-220). Peter Lang.
- Matielo, R., D'Ely, R. C. S. F., & Baretta, L. (2015). The effects of interlingual and intralingual subtitles on second language learning/acquisition: A state-of-the-art review. *Trabalhos Em Linguística Aplicada*, 54(1), 161–182.
- Matsuoka, W., & Hirsch, D. (2010). Vocabulary learning through reading: Does an ELT course book provide good opportunities? *Reading in a Foreign Language*, 22(1), 56-70.
- Mayer, R. E. (2005). *The Cambridge handbook of multimedia learning*. Cambridge University Press.
- Mayer, R. E. (2009). *Multimedia learning*. Cambridge University Press.
- Mayer, R. E. (2010). Applying the science of learning to medical education. *Medical Education*, 44(6), 543–549.
- McCowan, R. J., & McCowan S. C., (1999) *Item analysis for criterion-referenced tests*. State University of New York.
- McKeown, K. R., & Radev, D. R. (2000). Collocations. In R. Dale, H. Moisl, & H. Somers (Eds.), *A handbook of natural language processing*. (pp. 507–523). Marcel Dekker.
- McNiff, J. (2013). *Action research: Principles and practice*. Routledge
- Meara, P. (1989). Matrix models of vocabulary acquisition. *AILA Review*, 6, 66–85.
- Meara, P. (1996). The dimensions of lexical competence. In G. Brown, K. Malmkjær, & J. Williams (Eds.). *Performance and competence in second language acquisition* (pp. 35 –53). Cambridge University Press.
- Meara, P. (2005). *Llama language aptitude tests*. Lognostics. Retrieved from https://www.lognostics.co.uk/tools/LLAMA_3/index.htm
- Meara, P. M. (2010). *V_Words* [Computer software]. Lognostics. Retrieved from: https://www.lognostics.co.uk/tools/V_Words/V_Words.htm
- Meara, P., & Jones, G. (1990) *Eurocentres Vocabulary Size Tests 10*. Eurocentres Learning Service.

References

- Milton, J., & Alexiou, T. (2006). Language aptitude development in young learners. In C. Abello-Contesse, R. Chacón-Beltrán, M. D. López-Chiménez, & M. M. Torreblanca-López (Eds.), *Age in L2 acquisition and teaching* (pp. 177-192). Peter Lang.
- Miralpeix, I., Gesa, F., & Suárez, M. M. (in press). Vocabulary learning after minimal exposure to subtitled input. In B. L. Reynolds (Ed.), *Vocabulary Learning in the Wild*. Springer.
- Miralpeix, I., & Muñoz, C. (2018). Receptive vocabulary size and its relationship to EFL language skills. *International Review of Applied Linguistics in Language Teaching*, 56(1), 1–24.
- Mondini, S., Jarema, G., Luzzatti, C., Burani, C., & Semenza, C. (2002). Why is “red cross” different from “yellow cross”? A neurophysiological study of non-adjective agreement within Italian compounds. *Brain and Language*, 81(1-3), 621–634.
- Montero Perez, M., & Desmet, P. (2012). The effect of input enhancement in L2 listening on incidental vocabulary learning: a review. *Procedia - Social and Behavioral Sciences*, 34, 153-157.
- Montero Perez, M., Desmet, P., & Peters, E. (2015). Enhancing Vocabulary Learning Through Captioned Video: An Eye-Tracking Study. *The Modern Language Journal*, 99(2), 308–328.
- Montero Perez, M., Peters, E., & Desmet, P. (2014). Is less more? Effectiveness and perceived usefulness of keyword and full captioned video for L2 listening comprehension. *ReCALL*, 26(1), 21-43.
- Montero Perez, M., Peters, E., & Desmet, P. (2018). Vocabulary learning through viewing video: The effect of two enhancement techniques. *Computer Assisted Language Learning*, 31(1–2), 1–26.
- Montero Perez, M., Van Den Noortgate, W., & Desmet, P. (2013). Captioned video for L2 listening and vocabulary learning: A meta-analysis. *System*, 41(3), 720–739.
- Mueller, G. (1980). Visual contextual cues and listening comprehension: An experiment. *The Modern Language Journal*, 64 (3), 335–340.
- Muñoz, C. (1997). Age, exposure and foreign language learning. *APAC Monographs*, 2, 16-24.
- Muñoz, C. (2006). *Age and the rate of foreign language learning*. Multilingual Matters.
- Muñoz, C. (2008a). Symmetries and asymmetries of age effects in naturalistic and instructed L2 learning. *Applied Linguistics*, 29(4), 578-596.

References

- Muñoz, C. (2008b). Age-related differences in foreign language learning. Revisiting the empirical evidence. *International Review of Applied Linguistics (IRAL)*, 46(3), 197-220.
- Muñoz, C. (2010). Staying abroad with the family: a case study of two siblings' second language development during a year's immersion. *ITL. International Journal of Applied Linguistics*, 160, 24-48.
- Muñoz, C. (2014a). Exploring young learners' foreign language learning awareness. *Language Awareness*, 23(1-2), 24-40.
- Muñoz, C. (2014b). The association between aptitude components and language skills in young learners. In M. Pawlak, & L. Aronin (Eds.). *Essential topics in Applied Linguistics and Multilingualism. Studies in honor of David Singleton*, (pp.51-68). Springer International Publishing.
- Muñoz, C. (2017). The role of age and proficiency in subtitle reading. An eye-tracking study, *System*, 67, 77-86.
- Muñoz, C. (2020). Boys like games and girls like movies. Age and gender differences in out-of-school contact with English. *RESLA*, 33(1), 172-202.
- Muñoz, C., & Cadierno, T. (2021). How do differences in exposure affect English language learning? A comparison of teenagers in two learning environments. *Studies in Second Language Learning and Teaching*, 11(2), 185-212.
- Muñoz, C., Cadierno, T., & Casas, I. (2018). Different starting points for early foreign language learning: A comparative study with Danish and Spanish young learners of English. *Language Learning*, 68(4), 1076-1119.
- Muñoz, C., & Singleton, D. (2019). Age and multilingualism. In D. Singleton, & L. Aronin (Eds.) *Twelve lectures on multilingualism*. (pp.213-230). Multilingual Matters.
- Muñoz, C., Pattemore, A., & Avello, D. (2022). Exploring repeated captioning viewing as a way to promote vocabulary learning: time lag between repetitions and learner factors. *Computer Assisted Language Learning*, 1-27.
- Nagira, Y. (2011). Vocabulary learning through captions. *International Proceedings of Economics Development and Research*, 26, 95-99.
- Nagy, W. E., Herman, P. A., & Anderson, R. C. (1985). Learning words from context. *Reading Research Quarterly*, 20(2), 233-253.
- Nagy, W., Anderson, R. C., & Herman, P. (1987). Learning word meanings from context during normal reading. *American Educational Research Journal*, 24(2), 237-270.
- Nation, I. S. P (1983). Learning vocabulary. *Language Teacher* 9(1), 10-11.
- Nation, I. S. P. (1990). *Teaching and learning vocabulary*. Newbury House.

References

- Nation, I. S. P. (1995). Best practice in vocabulary teaching and learning. *EA Journal*, 3(2), 7-15.
- Nation, I. S. P. (2001). The role of the first language in foreign language learning. *Asian EFL Journal*, 5(2), 1–8.
- Nation, I. S. P. (2006). How large a vocabulary is needed for reading and listening? *The Canadian Modern Language Review*, 63(1), 59–82.
- Nation, I. S. P. (2007). The four strands. *Innovation in Language Learning and Teaching*, 1(1), 2-13.
- Nation, I. S. P. (2012). *The vocabulary size test*. Retrieved from: <https://www.wgtn.ac.nz/lals/resources/paul-nations-resources/vocabulary-tests/the-vocabulary-size-test/Vocabulary-Size-Test-information-and-specifications.pdf>
- Nation, I. S. P., & Beglar, D. (2007). A vocabulary size test. *The Language Teacher*, 31(7), 9-13.
- Nation, I. S. P., & Wang, K. (1999). Graded readers and vocabulary. *Reading in a Foreign Language*. 12(2), 355–380.
- Neuman, S. B., & Koskinen, P. (1992). Captioned television as comprehensible input: Effects of incidental word learning from context for language minority students. *Reading Research Quarterly*, 27(1), 94-106.
- Newton, P., & Burgess, D. (2008). Exploring types of educational action research: Implications for research validity. *International Journal of Qualitative Methods*, 7(4), 18–30.
- Noblitt, J. S. (1995): "Cognitive approaches to listening comprehension" in R. M. Terry (Ed.) *Dimension '95: The Future is Now. Selected Proceedings of the 1995 Joint Conference of the Southern Conference on Language Teaching and the South Carolina Foreign Language Teachers' Association*, Valdosta State University.
- Norris, J. M., & Ortega, L. (2000). Effectiveness of L2 instruction: A research synthesis and quantitative meta-analysis. *Language Learning*, 50(3), 417–528.
- Nunan, D. (1999). *Second language teaching and learning*. Heinle and Heinle.
- Nuttall, C. (1996). *Teaching reading skills in a foreign language*. Heinemann.
- Ockey, G. J. (2007). Construct implications of including still image or video in computer-based listening tests, *Language Testing*, 24(4), 517–537.
- Paivio, A. (1986). *Mental representations: A dual coding approach*. Oxford University Press.
- Paradis, M. (2009). *Declarative and procedural determinants of second languages*. John Benjamins.

References

- Paribakht, T.S., & Wesche, M. (1997). Vocabulary enhancement activities and reading for meaning in second language vocabulary acquisition. In J. Coady, & T. Huckin (Eds.), *Second language vocabulary acquisition: A rationale for pedagogy* (pp. 174–200). Cambridge University Press.
- Paribakht, T.S., & Wesche, M. (1999). Reading and 'incidental' L2 vocabulary acquisition. *Studies in Second Language Acquisition*, 21, 195-229.
- Parry, K. (1993). Too many words: Learning the vocabulary of an academic subject. In T. Huckin, M. Haynes, & J. Coady (Eds.), *Second language reading and vocabulary learning* (pp. 109–129). Ablex.
- Pattemore, A., & Muñoz, C. P. (2020). Learning L2 constructions from captioned audio-visual exposure: The effect of learner-related factors. *System*, 93, 1-13.
- Pattemore, A., Suárez, M. M., & Muñoz, C. (2020). Exploring L2 TV mode preferences and perceptions of learning. In K.-M. Frederiksen, S. Larsen, L. Bradley & S. Thouësny (Eds), *CALL for widening participation: short papers from EUROCALL 2020* (pp. 272-278).
- Patterson, J. L. (2002). Relationships of expressive vocabulary to frequency of reading and television experience among bilingual toddlers, *Applied Psycholinguistics*, 23(4), 493-508.
- Pavakanun, U., & d'Ydewalle, G. (1992). Watching foreign television programs and language learning. In F. L. Engel, D. G. Bouwhuis, T. Bosser, & G. d'Ydewalle (Eds.), *Cognitive modelling and interactive environments in language learning* (pp. 193-198). Springer.
- Pawlak, M. (2011). Anxiety as a factor influencing the use of language strategies. In M. Pawlak (Ed.), *Extending the boundaries of research on second language learning and teaching* (pp. 149-165). Springer.
- Pellicer-Sánchez, A. (2020). Learning single words vs. multiword items. In S. Webb (Ed.), *The Routledge handbook of vocabulary studies*. (pp. 158-173). Routledge.
- Pellicer-Sánchez, A., & Schmitt, N. (2010). Incidental vocabulary acquisition from an authentic novel: Do Things Fall Apart? *Reading in a Foreign Language*, 22(1), 31–55.
- Perego, E., Del Missier, F., Porta, M., & Mosconi, M. (2010). The cognitive effectiveness of subtitle processing. *Media Psychology*, 13(3), 243–272.
- Peters, E. (2014). The effects of repetition and time of post-test administration on EFL learners' form recall of single words and collocations. *Language Teaching Research*, 18(1), 75–94.
- Peters, E. (2018). The effect of out-of-class exposure to English language media on learners' vocabulary knowledge. *ITL - International Journal of Applied Linguistics*, 169(1), 142-168.

References

- Peters, E., & Muñoz, C. (2020). Introduction to Special Issue Language Learning from Multimodal Input. *Studies in Second Language Acquisition*, 42, 489-497.
- Peters, E., & Webb, S. (2018). Incidental vocabulary acquisition through viewing L2 television and factors that affect learning. *Studies in Second Language Acquisition*, 40(3), 551-577.
- Peters, E., Heynen, E., & Puimège, E. (2016). Learning vocabulary through audiovisual input: The differential effect of L1 subtitles and captions. *System*, 63, 134–148.
- Peters, E., Noreillie, A.-S., Heylen, K., Bulté, B., & Desmet, P. (2019). The impact of instruction and out-of-school exposure to foreign language input on learners' vocabulary knowledge in two languages. *Language Learning*, 69(3), 747-782.
- Pigada, M., & Schmitt, N. (2006). Vocabulary acquisition from extensive reading: a case study. *Reading in a Foreign Language*, 18(1), 1-28.
- Pimsleur, P. (1966). *Language aptitude battery*. Harcourt Brace.
- Pinter, A. (2017). *Teaching young language learners*. Oxford University Press.
- Pinter, A. (2018). Research issues with young learners. In S. Garton, & F. Copland (Eds.), *The Routledge handbook of teaching English to young learners* (pp.411-424). Routledge.
- Plass, J. L., & Jones, L. C. (2005). Multimedia learning in second language acquisition. In R. E. Mayer (Ed.), *The Cambridge handbook of multimedia learning* (pp. 467–488). Cambridge University Press.
- Prensky, B. M. (2001). Digital natives, digital immigrants. *On the Horizon*, 9(5), 1-6.
- Pujadas, G. (2019). *Language learning through extensive viewing. A study with adolescent EFL learners*. [Unpublished doctoral dissertation. University of Barcelona].
- Pujolà, J. T. (2002). CALLing for help: Researching language learning strategies using help facilities in a web-based multimedia program. *ReCALL*, 14(2), 235–262.
- Pulido, D. (2003). Modeling the role of second language proficiency and topic familiarity in second language incidental vocabulary acquisition through reading. *Language Learning*, 53(2), 233-284.
- Pulido, D. (2004). The relationship between text comprehension and second language incidental vocabulary acquisition: A matter of topic familiarity? *Language Learning*, 54(3), 469–523.
- Reber, A. S. (1976). Implicit learning of synthetic languages: The role of instructional set. *Journal of Experimental Psychology: Human Learning and Memory*, 2(1), 88–94.

References

- Rice, M. L., & Woodsmall, L. (1988). Lessons from television: Children's word learning when viewing. *Child Development*, 59(2), 420–429.
- Richards, J. C. (1983). Listening comprehension: Approach, design, procedure. *TESOL Quarterly*, 17(2), 219–240.
- Richards, J. C. (1990). *The language teaching matrix*. Cambridge University Press.
- Richards, J. C., & Schmidt, R. (2002). *Longman dictionary of language teaching and applied linguistics*. Pearson Education.
- Rieder, A. (2003). Implicit and explicit learning in incidental vocabulary acquisition. *Views* 12(2), 24-39.
- Roberts, L., & Felser, C. (2011). Plausibility and recovery from garden paths in second language sentence processing. *Applied Psycholinguistics*, 32(2), 299–331.
- Robin, R. (2007). Commentary: learner-based listening and technological authenticity. *Language Learning & Technology*, 11(1), 109–115.
- Robinson, P. J. (1988). A Hallidayan framework for vocabulary teaching—An approach to organizing the lexical content of an EFL syllabus. *International Review of Applied Linguistics in language Teaching*, 16(3), 229-238.
- Robinson, P. J. (2001). Individual differences, cognitive abilities, aptitude complexes and learning conditions in second language acquisition. *Second Language Research*, 17(4), 368–392.
- Robinson, P. J. (2005). Aptitude and second language acquisition. *Annual Review of Applied Linguistics*, 25, 46-73.
- Rodgers, M. P. H. (2013). *English language learning through viewing television, an investigation of comprehension, incidental vocabulary acquisition, lexical coverage, attitudes, and captions*. [Unpublished doctoral dissertation. Victoria University of Wellington].
- Rodgers, M. P. H., & Webb, S. (2011). Narrow viewing: The vocabulary in related television programs. *TESOL Quarterly*, 45(4), 689-717.
- Rodgers, M. P. H., & Webb, S. (2017). The effects of captions on EFL learners' comprehension of English-language television programs, *Journal of the Computer Assisted Language Instruction Consortium (CALICO)*, 34(1), 20–38.
- Rodriguez, M. C. (2005). Three options are optimal for multiple-choice items: A meta-analysis of 80 years of research. *Educational Measurement: Issues and Practice*, 24(2), 3–13.
- Roever, C. (2005). *Testing ESL pragmatics*. Peter Lang.

References

- Roever, C. (2011). Testing of second language pragmatics: Past and future. *Language Testing*, 28(4), 463–481.
- Rogers, V., Meara, P., Barrett-Leigh, T., Curry, C., & Davie, E. (2017). Examining the LLAMA aptitude tests. *Journal of the European Second Language Association*, 1(1), 49–60.
- Rosa, D. (2011). *The effects of aptitude and motivation on the acquisition of EFL in young learners*. [Unpublished Master's Thesis, English Studies Department, Applied Linguistics Program, Universitat de Barcelona].
- Rott, S. (1999). The effect of exposure frequency on intermediate language learners' incidental vocabulary acquisition and retention through reading. *Studies in Second Language Acquisition*, 21(4), 589–619.
- Rubin, J. (1994). A review of second language listening comprehension research. *Modern Language Journal*, 78(2), 199–221.
- Sawyer, M., & Ranta, L. (2001). Aptitude, individual differences, and instructional design. In P. Robinson (Ed.), *Cognition and second language acquisition* (pp. 319-353). Cambridge University Press.
- Selinker, L. (1972). Interlanguage. *International Review of Applied Linguistics*, 10(3), 209–231.
- Senechal, M., & Cornell, E. H. (1993). Vocabulary acquisition through shared reading experiences. *Reading Research Quarterly*, 28(4), 361–373.
- Schmidt, R., & Frota, S. (1986). Developing basic conversational ability in a second language. A case study of an adult learner of Portuguese. In R. Day (Ed.), *Talking to learn: Conversation in second language acquisition*, (pp. 237-326). Newbury House.
- Schmitt, N. (2000). *Vocabulary in language teaching*. Cambridge University Press.
- Schmitt, N. (2004). *Formulaic sequences: Acquisition, processing and use*. John Benjamins.
- Schmitt, N., & Redwood, S. (2011). Learner knowledge of phrasal verbs. A corpus-informed study. In F. Meunier, S. De Cock, G. Gilquin, & M. Paquot (Eds.), *A taste for corpora: In honour of Sylviane Granger* (pp.173–209). John Benjamins.
- Schmitt, N., Schmitt, D., & Clapham, C. (2001). Developing and exploring the behaviour of two new versions of the Vocabulary Levels Test. *Language Testing*, 18(1), 55–88.
- Scovel, T. (2000). A critical review of the critical period research. *Annual Review of Applied Linguistics*, 20, 213-223.

References

- Shabani, K., & Zanussi, M. P. (2015). The impact of watching captioned TV series on vocabulary development of EFL students. *Journal for the Study of English Linguistics*, 3(1), 118.
- Sherman, J. (2003). *Using authentic video in the language classroom*. Cambridge University Press.
- Shroyer, E. H., & Birch, J. (1980). Captions and reading rates of hearing-impaired students. *American annals of the deaf*, 125(7), 916–922.
- Singleton, D. (1989). *Language acquisition. The age factor*. Multilingual Matters.
- Singleton, D. (1995). Introduction: A critical look at the Critical Period Hypothesis in second language acquisition research. In D. Singleton, & Z. Lengyel (Eds.), *The age factor in second language acquisition* (pp. 1–29). Multilingual Matters.
- Skehan, P. (1986). The role of foreign language aptitude in a model of school learning. *Language Testing*, 3, 188–221.
- Skehan, P. (1989). *Individual differences in second language learning*. Edward Arnold.
- Skehan, P. (1998). *A cognitive approach to language learning*. Oxford University Press.
- Skehan, P. (2002). Theorising and updating aptitude. In P. Robinson (Ed.), *Individual differences and instructed language learning* (pp. 69–93). John Benjamins.
- Skehan, P. (2012). Language aptitude. In S. M. Gass, & A. Mackey (Eds.). *The Routledge handbook of second language acquisition* (pp. 381-395). Routledge.
- Skehan, P., & Ducroquet, L. (1988). *A comparison of first and foreign language ability*. London University.
- Sparks, R., & Ganschow, L. (1991). Foreign language learning differences: Affective or native language aptitude differences? *Modern Language Journal*, 75(1), 3-16.
- Sparks, R., & Ganschow, L. (2001). Aptitude for learning a foreign language. *Annual Review of Applied Linguistics* 21(1), 90–111.
- Sparks, R., Patton, J., Ganschow, L., Humbach, N., & Javorsky, J. (2006). Native language predictors of foreign language proficiency and foreign language aptitude. *Annals of Dyslexia*, 56, 129–160.
- Sparks, R. L., Patton, J. O. N., Ganschow, L., & Humbach, N. (2009). Long-term relationships among early first language skills, second language aptitude, second language affect, and later second language proficiency, *Applied Psycholinguistics* 30(4), 725–755.
- Spaan, M. (2007). Evolution of a test item. *Language Assessment Quarterly*, 4(3), 279–293.

References

- Staehr, L. S. (2009). Vocabulary knowledge and advanced listening comprehension in English as a foreign language. *Studies in Second Language Acquisition*, 31(4), 577–607.
- Stewart, M. A., & Pertusa, I. (2004). Gains to foreign language while viewing target language closed-caption films. *Foreign Language Annals*, 37(3), 438–443.
- Suárez, M. M. (2010). *Language aptitude in young learners: The Elementary Modern Language Aptitude Test in Spanish and Catalan*. [Unpublished doctoral dissertation. University of Barcelona].
- Suárez, M. M., Gilabert, R., & Moskvina, N. (2021). The mediating role of vocabulary size, working memory, attention and inhibition in early vocabulary learning under different TV genres: An exploratory study. *TESOL Journal*, 12(4), 1-29.
- Suárez, M. M., & Gesa, F. (2019). Learning vocabulary with the support of sustained exposure to captioned video: Do proficiency and aptitude make a difference? *Language Learning Journal*, 47(4), 497–517.
- Suárez, M. M., & Muñoz, C. (2011). Aptitude, age and cognitive development: The MLAT-E in Spanish and Catalan. In L. Roberts, G. Pallotti, & C. Bettoni (Eds.), *EuroSLA Yearbook 11*, (pp.5-29). John Benjamins.
- Sundqvist, P. (2009). Extramural English matters: Out-of-school English and its impact on Swedish ninth graders' oral proficiency and vocabulary. [Doctoral Dissertation, Karlstad University, Karlstad].
- Sundqvist, P., & Sylvén, L. K. (2014). Language-related computer use: Focus on young L2 English learners in Sweden. *ReCALL*, 26(1), 3–20.
- Sundqvist, P., & Wikström, P. (2015). Out-of-school digital gameplay and in-school L2 English vocabulary outcomes. *System*, 51, 65–76.
- Swain, M., & Lapkin, S. (1995). Problems in output and the cognitive processes they generate: A step towards second language learning. *Applied Linguistics*, 16(3), 371-391.
- Sweller, J. (1988). Cognitive load during problem solving: Effects on learning. *Cognitive Science*, 12(2), 257-285.
- Sweller, J. (2005). Implications of cognitive load theory for multimedia learning. In R. E. Mayer (Ed.), *The Cambridge handbook of multimedia learning* (pp. 19–30). Cambridge University Press.
- Sydorenko, T. (2010). Modality of input and vocabulary acquisition. *Language Learning & Technology*, 14(2), 50–73.
- Sylvén, L. K., & Sundqvist, P. (2012). Gaming as extramural English: L2 learning and L2 proficiency among young learners. *ReCALL*, 24(3), 302–321

References

- Tamayo, A. (2016). Reading speed in subtitling for hearing impaired children: An analysis in Spanish television. *Journal of Specialised Translation*, 26, 275-294.
- Taylor, G. (2005). Perceived processing strategies of students watching captioned video, 1992–1993. *Foreign Language Annals*, 38(3), 422-427.
- Teng, M. F. (2022). Incidental L2 vocabulary learning from viewing captioned videos: Effects of learner-related factors. *System*, 105, 1-12.
- Tragant, E., & Pellicer-Sánchez, A. (2019). Young learners' engagement with multimodal exposure: An eyetracking study. *System*, 80, 212–223.
- Turgut, Y., & Irgin, P. (2009). Young learners' language learning via computer games. *Procedia—Social and Behavioral Sciences*, 88(1), 760-764.
- Vandergrift, L. (2007). Recent developments in second and foreign language listening comprehension research. *Language Teaching*, 40(3), 191–210.
- Vandergrift, L., & Goh, C. C. M. (2012). *Teaching and learning second language listening*. Routledge.
- Vanderplank, R. (1988). The value of teletext sub-titles in language learning. *ELT Journal*, 42(4), 272–281.
- Vanderplank, R. (1993). A very verbal medium: Language learning through closed captions. *TESOL Journal* 3(1), 10–14.
- Vanderplank, R. (1999). Global medium – global resource? Perspectives and paradoxes in using authentic broadcast material for teaching and learning English. In C. Gnutzmann (Ed.), *Teaching and learning English as a global language: Native and non-native perspectives*. (pp.253-266). Stauffenberg.
- Vanderplank, R. (2010). Déjà vu? A decade of research on language laboratories, television and video in language learning. *Language Teaching*, 43(1), 1–37.
- Vanderplank, R. (2013). 'Effects of' and 'effects with' captions: How exactly does watching a TV programme with same-language subtitles make a difference to language learners? *Language Teaching*, 47(1), 1-16.
- Vanderplank, R. (2016). *Captioned media in foreign language learning and teaching: subtitles for the deaf and hard-of-hearing as tools for language learning*. London, Palgrave Macmillan.
- VanPatten, B. (Ed.). (2004). *Processing instruction: Theory, research, and commentary*. Routledge.
- Van Der Zee, T., Admiraal, W., Paas, F., Saab, N., & Giesbers, B. (2017). Effects of subtitles, complexity, and language proficiency on learning from online education videos. *Journal of Media Psychology*, 29(1), 18–30.

References

- Van Zeeland, H., & Schmitt, N. (2013). Incidental vocabulary acquisition through L2 listening: A dimensions approach. *System*, 41(3), 609–624.
- Vidal, K. (2011). A comparison of the effects of reading and listening on incidental vocabulary acquisition. *Language Learning*, 61(1), 219–258.
- Vulchanova, M., Aurstad, L. M. G., Kvitnes, I.E.N., & Eshuis, H. (2015). As naturalistic as it gets: Subtitles in the English classroom in Norway. *Frontiers in Psychology*, 5(1), 1-10.
- Waring, R., & Tatakai, M. (2003). At what rate do learners learn and retain new vocabulary from reading a graded reader?. *Reading in a Foreign Language* 15(2), 130–160.
- Webb, S. (2007). The effects of repetition on vocabulary knowledge. *Applied Linguistics*, 28(1), 46-65.
- Webb, S. (2008). The effects of context on incidental vocabulary learning. *Reading in a Foreign Language*, 20(2), 232–245.
- Webb, S. (2010). Using glossaries to increase the lexical coverage of television programs. *Reading in a Foreign Language*, 22(1), 201–221.
- Webb, S. (2011). Selecting television programs for language learning: Investigating television programs from the same genre. *International Journal of English Studies*. 11(1), 117–135.
- Webb, S. (2015). Extensive viewing: Language learning through watching television. In D. Nunan, & J. C. Richards (Eds.), *Language learning beyond the classroom* (pp. 159-168). Routledge.
- Webb, S., & Nation, P. (2017). *How vocabulary is learned*. Oxford University Press.
- Webb, S., & Rodgers, M. P. H. (2009a). Vocabulary demands of television programs. *Language Learning*, 59(2), 335–366.
- Webb, S., & Rodgers, M. P. H. (2009b). The lexical coverage of movies. *Applied Linguistics*, 30(4), 407–427.
- Webb, S., Newton, J., & Chang, A. (2013). Incidental learning of collocation. *Language Learning*, 63(1), 91-120.
- Wen, Z., & Skehan, P. (2011). A new perspective on foreign language aptitude research: building and supporting a case for “working memory as language aptitude”. *Ilha Do Desterro*, 60, 15-43.
- Wen, Z., Biedron, A., Skehan, P. (2017). Foreign language aptitude theory: Yesterday, today and tomorrow. *Language Teaching*, 50(1), 1-31.

References

- Whitford, V., & Joanisse, M. (2018). Do eye movements reveal differences between monolingual and bilingual children's first-language and second-language reading? A focus on word frequency effects. *Journal of Experimental Child Psychology*, 173, 318-337.
- Winke, P. (2005). *Individual differences in adult Chinese second language acquisition: The relationships among aptitude, memory and strategies for learning*. [Unpublished doctoral dissertation, Georgetown University, Washington, DC].
- Winke, P., Gass, S., & Sydorenko, T. (2010). The effects of captioning videos used for foreign language listening activities. *Language Learning & Technology*, 14(1), 65–86.
- Winke, P., Hall, B. W., Road, R.C., & Gass, S. (2013). Factors influencing the use of captions by foreign language learners: An eye - tracking study. *The Modern Language Journal*, 97(1), 254-275.
- Wood, J. (2008). The teacher as researcher. In E. Provenzo (Ed.), *Encyclopedia of the social and cultural foundations of education*. Sage.
- Woodinsky, M., & Nation, I. S. P. (1988). Learning from graded readers. *Reading in a Foreign Language*, 5(1), 155-161.
- Wray, A. (1999). Formulaic language in learners and native speakers. *Language Teaching*, 32(4), 213–231.
- Wray, A. (2002). *Formulaic language and the lexicon*. Cambridge University Press.
- Young, D. J. (1999). Linguistic simplification of SL reading material: Effective instructional practice? *The Modern Language Journal*, 83(3), 350–366.
- Young, E. (1989). *Lon Po Po: A Red Riding Hood Story from China*. Penguin Putnam Books for Young Readers.
- Yuksel, D., & Tanriverdi, B. (2009). Effects of watching captioned movie clip on vocabulary development of EFL learners. *Turkish Online Journal of Educational Technology*, 8(2), 48–54.
- Zafar, S., & Meenakshi, K. (2012). Individual learner differences and second language acquisition: A review. *Journal of Language Teaching and Research*, 3(4), 639-646.
- Zahar, R., Cobb, T., & Spada, N. (2001). Acquiring vocabulary through reading: Effect of frequency and contextual richness. *The Canadian Modern Language Review*, 57(3), 541-572.
- Zanon (2007). Learning vocabulary through authentic video and subtitles. *TESOL-SPAIN*, 31, 5-8.

References

- Zárate, S. (2008). Subtitling for deaf children on British television. *The Sign Language Translator and Interpreter*, 2(1), 1-20.
- Zárate, S. (2010). Bridging the gap between Deaf Studies and AVT for deaf children. In J. Díaz Cintas, A. Matamala, & J. Neves (Eds.). *New insights into audiovisual translation and media accessibility: Media for all 2*. (pp.159-173). Rodopi.
- Zimmerman, C. B. (1997). Do reading and interactive vocabulary instruction make a difference? An empirical study. *TESOL Quarterly*, 31(1), 121-140.

APPENDIX A – VOCABULARY SIZE TEST (VST)

Vocabulary Size Test: Version A

1. see: They <saw it>.
 - a closed it tightly
 - b waited for it
 - c looked at it
 - d started it up
2. time: They have a lot of <time>.
 - a money
 - b food
 - c hours
 - d friends
3. period: It was a difficult <period>.
 - a question
 - b time
 - c thing to do
 - d book
4. figure: Is this the right <figure>?
 - a answer
 - b place
 - c time
 - d number
5. poor: We <are poor>.
 - a have no money
 - b feel happy
 - c are very interested
 - d do not like to work hard
6. microphone: Please use the <microphone>.
 - a machine for making food hot
 - b machine that makes sounds louder
 - c machine that makes things look bigger
 - d small telephone that can be carried around
7. nil: His mark for that question was <nil>.
 - a very bad
 - b nothing
 - c very good
 - d in the middle
8. pub: They went to the <pub>.
 - a place where people drink and talk
 - b place that looks after money
 - c large building with many shops
 - d building for swimming
9. circle: Make a <circle>.
 - a rough picture
 - b space with nothing in it
 - c round shape
 - d large hole
10. dig: Our dog often <digs>.
 - a solves problems with things
 - b creates a hole in the ground
 - c wants to sleep
 - d enters the water
11. soldier: He is a <soldier>.
 - a person in a business
 - b person who studies
 - c person who uses metal
 - d person in the army
12. restore: It has been <restored>.

Appendices

- a said again
b given to a different person
c made like new again
d given a lower price
13. pro: He's <a pro>.
a someone who is employed to find out important secrets
b a stupid person
c someone who writes for a newspaper
d someone who is paid for playing sport
14. compound: They made a new <compound>.
a agreement
b thing made of two or more parts
c group of people forming a business
d guess based on past experience
15. deficit: The company <had a large deficit>.
a spent a lot more money than it earned
b went down a lot in value
c had a plan for its spending that used a lot of money
d had a lot of money stored in the bank
16. strap: He broke the <strap>.
a promise
b top cover
c shallow dish for food
d strip of strong material
17. weep: He <wept>.
a finished his course
b cried
c died
d worried
18. haunt: The house is <haunted>.
a full of decorations
b rented
c empty
d full of ghosts
19. cube: I need one more <cube>.
a sharp thing used for joining things
b solid square block
c tall cup with no saucer
d piece of stiff paper folded in half
20. butler: They have a <butler>.
a man servant
b machine for cutting up trees
c private teacher
d cool dark room under the house
21. nun: We saw a <nun>.
a long thin creature that lives in the earth
b terrible accident
c woman following a strict religious life
d unexplained bright light in the sky
22. olive: We bought <olives>.
a oily fruit
b scented flowers
c men's swimming clothes
d tools for digging
23. shudder: The boy <shuddered>.
a spoke with a low voice
b almost fell
c shook

- d called out loudly
24. threshold: They raised the <threshold>.
- a flag
 - b point or line where something changes
 - c roof inside a building
 - d cost of borrowing money
25. demography: This book is about <demography>.
- a the study of patterns of land use
 - b the study of the use of pictures to show facts about numbers
 - c the study of the movement of water
 - d the study of population
26. malign: His <malign> influence is still felt.
- a good
 - b evil
 - c very important
 - d secret
27. strangle: He <strangled her>.
- a killed her by pressing her throat
 - b gave her all the things she wanted
 - c took her away by force
 - d admired her greatly
28. dinosaur: The children were pretending to be <dinosaurs>.
- a robbers who work at sea
 - b very small creatures with human form but with wings
 - c large creatures with wings that breathe fire
 - d animals that lived an extremely long time ago
29. jug: He was holding <a jug>.
- a a container for pouring liquids
 - b an informal discussion
 - c a soft cap
 - d a weapon that blows up
30. crab: Do you like <crabs>?
- a very thin small cakes
 - b tight, hard collars
 - c sea creatures that always walk to one side
 - d large black insects that sing at night

APPENDIX B - SPANISH AND ENGLISH READING SPEED TESTS

B.1 Spanish reading speed test: text, questions and answers.

EL PRINCIPITO



Cuando yo tenía seis años vi en un libro sobre la selva virgen que se titulaba "Historias vividas", una magnífica lámina. Representaba una serpiente boa que se tragaba a una fiera.

En el libro se afirmaba: "La serpiente boa se traga su presa entera, sin masticarla. Luego ya no puede moverse y duerme durante los seis meses que dura su digestión".

Reflexioné mucho en ese momento sobre las aventuras de la jungla y a mi vez logré trazar con un lápiz de colores mi primer dibujo.

Enseñé mi obra de arte a las personas mayores y les pregunté si mi dibujo les daba miedo.

-¿Por qué habría de asustar un sombrero? - me respondieron.

Mi dibujo no representaba un sombrero. Representaba una serpiente boa que digiere un elefante. Dibujé entonces el interior de la serpiente boa a fin de que las personas mayores pudieran comprender. Siempre estas personas tienen necesidad de explicaciones.

Las personas mayores me aconsejaron abandonar el dibujo de serpientes boas, ya fueran abiertas o cerradas, y poner más interés en la geografía, la historia, el cálculo y la gramática. De esta manera a la edad de seis años abandoné una magnífica carrera de pintor. Había quedado desilusionado por el fracaso de mis dibujos número 1 y número 2. Las personas mayores nunca pueden comprender algo por sí solas y es muy aburrido para los niños tener que darles una y otra vez explicaciones.

Tuve, pues, que elegir otro oficio y aprendí a pilotar aviones. He volado un poco por todo el mundo y la geografía, en efecto, me ha servido de mucho; al primer vistazo podía distinguir perfectamente la China de Arizona. Esto es muy útil, sobre todo si se pierde uno durante la noche.

A lo largo de mi vida he tenido multitud de contactos con multitud de gente seria. Viví mucho con personas mayores y las he conocido muy de cerca; pero esto no ha mejorado demasiado mi opinión sobre ellas.

Cuando me he encontrado con alguien que me parecía un poco lúcido, lo he sometido a la experiencia de mi dibujo número 1 que he conservado siempre. Quería saber si verdaderamente era un ser comprensivo. E invariablemente me contestaban siempre: "Es un sombrero". Me abstenia de hablarles de la serpiente boa, de la selva virgen y de las estrellas. Poniéndome a su altura, les hablaba del bridge, del golf, de política y de corbatas. Y mi interlocutor se quedaba muy contento de conocer a un hombre tan razonable.



PREGUNTAS DE COMPRENSIÓN



¿De qué temática era el libro que leyó el narrador?

- De la selva virgen De la sabana De una llanura De un bosque

SIGUIENTE >



PREGUNTAS DE COMPRENSIÓN



¿Cómo come la boa?

- Se traga entera la presa La mastica lentamente La come a trozos La boa no tiene presas

SIGUIENTE >



PREGUNTAS DE COMPRENSIÓN



¿Qué representaba el dibujo del narrador?

Una serpiente boa que digiere un elefante Una serpiente boa que se tragaba a una fiera
 Un sombrero Una serpiente boa que se tragaba un sombrero

SIGUIENTE >



PREGUNTAS DE COMPRENSIÓN



¿Qué otro oficio aprendió el narrador tras dejar la pintura?

Piloto de naves espaciales Piloto de carreras
 Piloto de barcos Piloto de Avión

TERMINAR

PREGUNTA	TUS RESPUESTAS	RESPUESTAS CORRECTAS
1	✓	<i>EL libro trata de la selva virgen</i>
2	✓	<i>La boa se traga la presa entera</i>
3	✓	<i>El dibujo representa a una boa que se traga un elefante.</i>
4	✓	<i>El narrador se hizo piloto de avión.</i>

B.2 English reading speed test: text, questions and answers

Fun Facts Level 01 Story 01

People have hair. Cats and dogs have fur. Sheep have wool. We use wool to make yarn. The yarn is used to make clothes. We also use it for blankets. It keeps us warm.

Farmers cut the wool off the sheep. It doesn't hurt them. It's like getting your hair cut.

Wool clothes can shrink when they get wet. The wool does not shrink on the sheep. Their skin makes lanolin. It is a grease. It keeps the wool from changing size.

180 Words/Minute Cancel Done Reading

Fun Facts Level 01 Story 01

1. We use wool to make

- A. clothes.
- B. shoes.
- C. yarn.

2. Wool keeps us

- A. comfortable.
- B. warm.
- C. cool.

3. Lanolin is a(n)

- A. oil.
- B. grease.
- C. moisturizer

4. Wool clothes can

- A. shrink.
- B. expand.
- C. be greasy.

Cancel Done

Fun Facts Level 01 Story 01

1. We use wool to make

- A. clothes.
- B. shoes.
- C. **yarn.**

2. Wool keeps us

- A. comfortable.
- B. **warm.**
- C. cool.

3. Lanolin is a(n)

- A. oil.
- B. **grease.**
- C. moisturizer

4. Wool clothes can

- A. **shrink.**
- B. expand.
- C. be greasy.

0% Comprehension [Cancel](#) [View Results](#)

APPENDIX C – OUT-OF-SCHOOL EXPOSURE QUESTIONNAIRE

Nom _____ Curs _____ Edat _____

QÜESTIONARI (5è i 6è Primària)

Encercla i escriu (en cas necessari) la resposta que creguis més adient de les següents preguntes:

1. Quina llengua/gües parles habitualment amb la família?

Pare _____ Mare: _____ Germans: _____ Altres (avis etc.) _____

2. Quina llengua/gües parles amb els amics? _____

3. Fas classes extraescolars d'anglès?

Sí No

a. Si has respost que Sí:

a1. Des de quan fas extraescolars d'anglès? _____

a2. Quantes hores a la setmana hi vas? _____

4. Has estat mai a un país on es parli anglès? Sí No

a. On? _____ b. Quant de temps hi has estat? _____

5. Amb quina freqüència realitzes aquestes activitats?

Activitats	Mai	Menys d'1 cop al mes	Entre 1-3 cops al mes	Entre 1-3 cops per setmana	Entre 4-6 cops per setmana	Cada dia
Escoltar música en anglès						
Llegir llibres, còmics, revistes, etc. en anglès						

Veure pel·lícules, dibuixos i/o sèries en anglès:	Mai	Menys d'1 cop al mes	Entre 1-3 cops al mes	Entre 1-3 cops per setmana	Entre 4-6 cops per setmana	Cada dia
Sense subtítols						
Amb subtítols en català / castellà						
Amb subtítols en anglès						

Jugar a jocs d'ordinador, mòbil o videojocs en anglès	Mai	Menys d'1 cop al mes	Entre 1-3 cops al mes	Entre 1-3 cops per setmana	Entre 4-6 cops per setmana	Cada dia
Sense subtítols						
Amb subtítols en català / castellà						
Amb subtítols en anglès						

Nom _____ Curs _____ Edat _____

QÜESTIONARI (5è i 6è Primària)

A emplenar pel mestre segons les respostes que li proporciona l'alumne/a.

6. Si veus pel·lícules, dibuixos i/o sèries en anglès:

a. Quines acostumes a veure?

Títol

Amb Subtítols?

Català/Castellà

Anglès

Català/Castellà

Anglès

Català/Castellà

Anglès

b. Amb qui els/les veus?

- Sol/a
- Amb els germans/es
- Amb els pares
- Amb els avis
- Altres _____

c. T'agrada fer-ho?

- Molt Bastant No gaire Gens

7. Si veus pel·lícules, dibuixos i/o sèries en anglès amb subtítols, pensa com mires aquests vídeos:

a. Ho fas...

- Escoltant i llegint els subtítols a la vegada.
- Escoltant i algun cop llegint els subtítols.
- Només escoltant i no he llegit mai els subtítols.
- Només llegint i no he parat mai atenció al que escoltava.

b. Creus que els subtítols passen massa ràpid?

- No, sempre que els llegeixo tinc temps d'acabar-los.
- A vegades no tinc temps d'arribar al final.
- Sí, van molt ràpid i mai tinc temps de llegir-los tots.

c. Si llegeixes els subtítols, per què creus que llegeixes els subtítols? Pots triar més d'una resposta.

- Perquè si no, no entenc el que diuen.
- Perquè m'ajuden a entendre el que passa.
- Perquè estan allà i no puc evitar llegir-los.
- Perquè trobo que aprenc més coses (p.ex. vocabulari).

Altres.....

APPENDIX D – VIDEO-WATCHING QUESTIONNAIRE

Nom i cognoms _____ Curs _____ Data _____

Respon a les següents preguntes marcant la resposta amb una creu.

1. Has entès el que ha passat en aquest vídeo? (Marca amb una creu)

- Gens No gaire Una mica Bastant Molt Tot

2. Aquest vídeo tenia subtítols.... pensa sobre com has mirat aquest vídeo:

- Escoltant i llegint els subtítols a la vegada.
- Escoltant i algun cop llegint els subtítols.
- Només escoltant i no he llegit mai els subtítols.
- Només llegint i no he parat mai atenció al que escoltava

3. Si has intentat llegir els subtítols,

a. Has tingut temps de llegir-los sempre?

- No, mai No gaire Una mica Bastant sovint Força sovint Sí, sempre

b. Per què creus que llegies els subtítols? Posa una creu en els motius que pensis (n'hi pot haver més d'un)

- Perquè si no, no entenc el que diuen.
- Perquè m'ajuden a entendre el que passa.
- Perquè estan allà i no puc evitar llegir-los.
- Perquè trobo que aprenc més coses (p.ex. vocabulari).

Altres.....

c. Creus que amb el subtítols has après paraules noves?

- Sí, moltes
- Sí, alguna
- No, cap

Appendices

1. Responde a las siguientes preguntas sobre el vídeo que acabas de ver:

1. ¿En qué estación del año nos encontramos?

2. ¿Qué consejo le da Ted a George para ser buen futbolista?

3. ¿Qué harán George y Charkie?

4. ¿Por qué no se lleva el perro con ella la propietaria?

5. ¿Qué le dijeron Betsy y Steve a la dueña del perro sobre Charkie?

2. Marca con un círculo las palabras y expresiones que han aparecido en este vídeo.

KICK

FLOWERS

CHAIN

RUN

KEEP AN EYE ON HER

SLIPPED AWAY

IMPROVING

YOU ARE A GOOD BOY

GOALKEEPER

YOU'D BETTER GET GOING

DOGSIT

OVER AND OVER

Vocabulary words in the episode-based reading questionnaire.

Words in the video	Words not in the video
Kick	Run
Improving	Flowers
Dogsit	Goalkeeper
Over and over	You are a good boy
You'd better get going	Chain
Keep an eye on her	
Slipped away	

APPENDIX E – EPISODE-BASED COMPREHENSION AND VOCABULARY TESTS

Nombre y apellidos _____ Fecha _____ Curso _____

“A Zoo Night”

1. Marca si estos enunciados son verdaderos (V) o falsos (F).

- El zoo está en el mismo barrio donde vive George. V / F
- Cuando se va al zoo, Ted le dice que no tarde mucho en volver. V / F
- Cuando los animales están sueltos, George piensa que todos los zoolos deberían ser así. V / F
- Los animales piden a George que los devuelva a sus lugares en el zoo. V / F
- Los animales están tristes porque no pueden ir a casa con George. V / F

2. Escoge la respuesta correcta (A, B o C).

Mientras George mira el panda en el ordenador, Ted le dice...

- A) Que lleva muchas horas en el ordenador.
- B) Que él necesita el ordenador.
- C) Que en el zoo hay muchos animales para visitar.

¿Por qué George da hojas de una planta al oso panda?

- A) Porque el panda está flaco y tiene que crecer.
- B) Porque el panda solo no podrá coger hojas para comer.
- C) Porque el panda debe estar aburrido de comer siempre las mismas hojas de su jaula.

¿Por qué George decide coger las llaves que ve en el zoo?

- A) Para salir del zoo.
- B) Para entrar en la jaula del panda.
- C) Para liberar a los animales encerrados.

¿Qué piensa Ted cuando ve que George no vuelve?

- A) Que lo debe estar pasando muy bien.
- B) Que algún animal lo ha atacado.
- C) Que se ha hecho amigo de los animales.

¿Qué piensa el orangután cuando ve a George y Ted irse juntos del zoo?

- A) Que Ted es su padre.
- B) Que George es un mono afortunado.
- C) Que George se ha olvidado los plátanos.

3. Ordena cronológicamente (del 1 al 5) estos hechos que pasan en el capítulo que acabas de ver. El número 3 te puede servir de guía para ordenar los demás.

- ___ Ted piensa que el zoo debe estar cerrado.
- ___ George coge un mapa de un panel del zoo.
- 3 George libera las jirafas y los pingüinos.
- ___ El ruido de los animales puede despertar al pequeño panda.
- ___ George da de comer al panda.

4. Marca con un círculo las palabras y expresiones que han aparecido en este vídeo.

keys	baby	animals	lock up
snake	shown live	map	gently
cage	go the wrong way	door	good morning

Nombre y apellidos _____ Fecha _____ Curso _____

“Jorge Monta en Cohete”

1. Marca si estos enunciados son verdaderos (V) o falsos (F).

- Ted será la primera persona que vaya al espacio sin ser astronauta V / F
- Los astronautas se han quedado sin alimentos. V / F
- La profesora Wiseman ha diseñado el cohete. V / F
- George, desde el cohete, puede ver la tierra a su izquierda. V / F
- Cuando lance la carga a la estación espacial, Jorge volverá automáticamente a la Tierra. V / F

2. Escoge la respuesta correcta (A, B o C).

A Jorge le gusta visitar a la Profesora Wiseman porque...

- A) Ella le da golosinas para comer.
- B) Se lo pasa muy bien estando con ella.
- C) Ella hace muchas cosas diferentes.

Ted no puede pilotar el cohete porque...

- A) No se entiende con los dos profesores.
- B) Sólo tiene dos manos.
- C) Es demasiado peligroso.

Mientras Jorge está dentro del cohete y se pierde la comunicación con ellos, Ted ...

- A) Está celoso porque Jorge va en el cohete y él no.
- B) Espera que Jorge no esté asustado.
- C) Dice que no dejará que Jorge se vaya otra vez solo nunca más.

Los científicos, al ver que Jorge ha abierto las cajas de los experimentos, piensan que...

- A) Al menos no ha liberado a las hormigas.
- B) Jorge está probando antes de tiempo los experimentos.
- C) Jorge puede hacer lo que quiera ya que es muy inteligente.

Si Jorge no tira de la palanca a tiempo, el cohete...

- A) Aterrizará en Marte.
- B) Puede aterrizar en cualquier parte de la Tierra.
- C) Se volverá invisible y tendrán que buscarlo.

Appendices

3. Ordena cronológicamente (del 1 al 5) estos hechos que pasan en el capítulo que acabas de ver. El número 3 te puede servir de guía para ordenar los demás.

___ Jorge puede ver desde el cohete la estación espacial a la derecha.

___ La profesora Wiseman le dice a Jorge que lance la carga.

3 Jorge piensa que el espacio es un sitio fantástico para jugar.

___ Ted le dice a Jorge que sea bueno en el espacio.

___ La Profesora Wiseman le dice a Jorge que guarde los experimentos.

4. Marca con un círculo las palabras y expresiones que han aparecido en este vídeo.

toy

gleeful

keypads

pull the lever

bicycle

supply

great job

experiments

launch

food

pilot

endorsement

Nombre y apellidos _____ Fecha _____ Curso _____

“Jorge el Arquitecto”

1. Marca si estos enunciados son verdaderos (V) o falsos (F).

- Aunque en la obra hay muchas cosas para ver, Jorge prefiere comer el bocadillo. V / F
- El constructor sabe por qué se mueven los cimientos del edificio. V / F
- Ted le pregunta a Jorge si quiere leche para desayunar. V / F
- Ted le pregunta al portero si ha visto a Jorge. V / F
- El constructor acusa enfadado a Jorge por haber derribado el edificio. V / F

2. Escoge la respuesta correcta (A, B o C).

Cuando Jorge está subido en la obra, Ted le dice...

- A) Que le haga una foto.
- B) Que la obra no es un patio para jugar.
- C) Que se puede caer el edificio.

Ted le da un billete de 10 dólares a Jorge y le dice que ...

- A) Compre todo lo que pueda con este dinero.
- B) Que tiene 10 minutos para volver a casa.
- C) Que vaya con mucho cuidado.

Al no haber nadie en la obra, Jorge piensa que...

- A) Es un momento perfecto para echar un vistazo sin distraer a nadie.
- B) Se quedará vigilando hasta que llegue alguien.
- C) Aprovechará para jugar tranquilamente en el terreno.

Lo que debilitaba los cimientos de la obra era...

- A) Agua subterránea.
- B) El peso de las máquinas de la obra.
- C) Una cañería rota.

Cuando el propietario le da el dinero a Jorge, Ted dice que...

- A) Ahora son ricos.
- B) Traerán comida para todo el mundo.
- C) Les invita a comer a un restaurante.

Appendices

3. Ordena cronológicamente (del 1 al 5) estos hechos que pasan en el capítulo que acabas de ver. El número 3 te puede servir de guía para ordenar los demás.

- ___ La conductora de la grúa pregunta a Ted si Jorge es su mono.
- ___ El constructor dice que ha revisado todas las vigas la noche anterior.
- _3_ La gatita Brújula coge el billete de 10 dólares.
- ___ El dueño dice que empezará con un diseño nuevo y moderno.
- ___ El dueño decide cerrar la obra.

4. Marca con un círculo las palabras y expresiones que han aparecido en este vídeo.

- | | | | |
|-------------|--------------------|-------------------|--------------------|
| machine | piping | foundation | waffle |
| foreseeable | construction | wood | shut the site down |
| building | It's three o'clock | pretty impressive | blueberry |

Nombre y apellidos _____ Fecha _____ Curso _____

“Magnetismo Animal”

1. Marca si estos enunciados son verdaderos (V) o falsos (F).

- Ted le dice a Jorge que tiene varios imanes en casa. V / F
- La nevera de Ted es casi nueva. V / F
- El dependiente de la tienda sugiere a Ted que compre una nevera nueva. V / F
- George le pregunta al señor de la furgoneta dónde lleva la nevera de la señora V / F
- En el desguace, Ted le dice a Jorge que vaya a dar una vuelta. V / F

2. Escoge la respuesta correcta (A, B o C).

Ted piensa que el dibujo de Jorge es...

- A) Jorge conduciendo un coche.
- B) Un plátano conduciendo un coche.
- C) El vecino conduciendo un coche.

Cuando ve el dibujo de Jorge, la profesora Wiseman...

- A) Le pregunta si lo ha hecho él.
- B) Le dice que no le gusta nada.
- C) Le dice que el dibujo pertenece a un museo de arte.

Al ver el reino de los imanes, Ted...

- A) Cree que el magnetismo es su fuerza invisible favorita.
- B) Decide poner imanes en todas las paredes de su casa.
- C) Decide investigar más sobre el magnetismo.

Después de comprar los imanes, Ted le dice a Jorge...

- A) Que a partir de ahora harán colección y cada mes comprarán uno.
- B) Que los quedarán muy bien en la nevera.
- C) Que devolverán dos de ellos porque no le gustan.

Para dejar caer la nevera en el triturador, el señor de la grúa le dice a Jorge...

- A) Que le dé al botón rojo.
- B) Que puede hacer una pirueta en el aire primero.
- C) Que la deje caer poco a poco en su sitio.

Appendices

3. Ordena cronológicamente (del 1 al 5) estos hechos que pasan en el capítulo que acabas de ver. El número 3 te puede servir de guía para ordenar los demás.

- ___ Jorge cuelga su dibujo justo antes de dormirse.
- ___ Jorge no puede permitir que su dibujo vaya al triturador.
- 3 Ted le dice a Jorge que volverán al reino de los imanes otro día.
- ___ Jorge juega con el imán en el reino de los imanes.
- ___ Ted dice que las grandes obras de arte se cuelgan en la nevera.

4. Marca con un círculo las palabras y expresiones que han aparecido en este vídeo.

- | | | | |
|----------------|----------------|---------------|------------------|
| refrigerator | telephone | be in a hurry | no need to do it |
| magnet | have a mission | handle | gallery |
| hit the button | fair | be right back | hold up |

Nombre y apellidos _____ Fecha _____ Curso _____

“Jorge Jefe de Estación”

1. Marca si estos enunciados son verdaderos (V) o falsos (F).

- Bill llegará a casa en el tren número ocho. V / F
- El hermano del jefe de la estación recomienda a Ted y Jorge viajar en avión. V / F
- El hermano del jefe de la estación trae la cena a su hermano. V / F
- Ted compara el tamaño de los sándwiches de jamón con los hermanos. V / F
- Al llegar a la estación, Bill dice que el viaje ha sido el peor de su vida. V / F

2. Escoge la respuesta correcta (A, B o C).

El tren de Bill llegará a la estación a...

- A) Las cinco en punto.
- B) Las tres en punto.
- C) No llegará hasta el día siguiente.

Aunque es pronto para ir a buscar a Bill, Ted le dice a Jorge que:

- A) Pueden ir antes y ver cómo funciona la estación.
- B) Hablarán con el jefe de estación mientras tanto.
- C) Irán pero sin pasar por la autopista, por no llegar demasiado temprano.

El jefe de la estación reconoce a Ted...

- A) Porque es su vecino.
- B) Por su sombrero amarillo.
- C) Porque lo oye cantar cada día por la mañana.

El jefe de estación le dice a Ted que si pasa algo en la estación...

- A) Vendrá el encargado general de estaciones.
- B) Los conductores de los trenes lo solucionaran solos.
- C) Sonará una alarma.

Cada mediodía, el jefe de estación:

- A) Sale a dar un paseo.
- B) Se hecha una siesta.
- C) Recibe visitas de sus superiores.

Appendices

3. Ordena cronológicamente (del 1 al 5) estos hechos que pasan en el capítulo que acabas de ver. El número 3 te puede servir de guía para ordenar los demás.

- ___ Jorge se da cuenta que los trenes están desordenados.
- ___ Jorge necesita ayuda.
- _3_ El jefe de estación está comiendo.
- ___ El jefe de estación dice que Jorge sabe lo suficiente para dirigir la estación solito.
- ___ Bill avisa de que la convención de cometas ha terminado.

4. Marca con un círculo las palabras y expresiones que han aparecido en este vídeo.

- | | | | |
|---------------|-----------|-------------|--------------|
| out of order | flag | want to see | follow |
| train station | medicine | conductor | cut in half |
| relieved | stick out | steadily | warning bell |

Nombre y apellidos _____ Fecha _____ Curso _____

“Un Monito Embarrado”

***L1S**

1. Marca si estos enunciados son verdaderos (V) o falsos (F).

- Ted descubre pronto que Jorge ha perdido a Crocky. V / F
- La Profesora Wiseman le dice a Jorge que está sucio. V / F
- Ted le sugiere a Jorge que lave su avión. V / F
- Ted tiene que salir a pasear para no pensar en problemas. V / F
- Ted al final entiende por qué Jorge no quería bañarse. V / F

2. Escoge la respuesta correcta (A, B o C).

Ted le dice a Jorge que si tuviera más cuidado con sus juguetes...

- A) Le compraría muchos más.
- B) Sería un mono responsable.
- C) No perdería tantos.

La profesora Wiseman sugiere que Jorge no se quiere bañar porque...

- A) No le gusta el agua.
- B) Ya es mayor para bañarse.
- C) Lo hacía con demasiada frecuencia y lo ha aburrido.

Steve le dice a Jorge que para ganar dinero él y Betsy...

- A) Lavan a perros.
- B) Hacen exhibiciones de burbujas en la calle.
- C) Sacan a pasear a los perros de los vecinos.

Con Steve y Betsy, Jorge...

- A) Decide montar un negocio en casa de Ted.
- B) Necesita más jabón y agua y pide a Ted que traiga más.
- C) Aún echa más de menos a Crocky.

Cuando Charkie llega, Steve le pide a Jorge que...

- A) Vaya a buscar más agua.
- B) Limpie la pelota de Charkie.
- C) Recoja a Crocky.

Appendices

3. Ordena cronológicamente (del 1 al 5) estos hechos que pasan en el capítulo que acabas de ver. El número 3 te puede servir de guía para ordenar los demás.

___ Ted le dice a Jorge que debe limpiarse antes de comer.

___ Ted le dice a la profesora Wiseman que es un genio.

3 Jorge hace cosquillas a Ted.

___ Ted encuentra el barco en el congelador.

___ Ted le sugiere a Jorge que ayude a Betsy y Steve.

4. Marca con un círculo las palabras y expresiones que han aparecido en este vídeo.

you are grounded

help me

frog

truck

clean up

buy sweets

muddy

rubber duck

lunch

sponge

bubble-maker

take a bath

Nombre y apellidos _____ Fecha _____ Curso _____

“Un Monito Embarrado”

*L2S

1. Marca si estos enunciados son verdaderos (V) o falsos (F).

- Ted descubre pronto que Jorge ha perdido a Sproingy. V / F
- La Profesora Wiseman le dice a Jorge que está sucio. V / F
- Ted le sugiere a Jorge que lave su avión. V / F
- Ted tiene que salir a pasear para no pensar en problemas. V / F
- Ted al final entiende por qué Jorge no quería bañarse. V / F

2. Escoge la respuesta correcta (A, B o C).

Ted le dice a Jorge que si tuviera más cuidado con sus juguetes...

- A) Le compraría muchos más.
- B) Sería un mono responsable.
- C) No perdería tantos.

La profesora Wiseman sugiere que Jorge no se quiere bañar porque...

- A) No le gusta el agua.
- B) Ya es mayor para bañarse.
- C) Lo hacía con demasiada frecuencia y lo ha aburrido.

Steve le dice a Jorge que para ganar dinero él y Betsy...

- A) Lavan a perros.
- B) Hacen exhibiciones de burbujas en la calle.
- C) Sacan a pasear a los perros de los vecinos.

Con Steve y Betsy, Jorge...

- A) Decide montar un negocio en casa de Ted.
- B) Necesita más jabón y agua y pide a Ted que traiga más.
- C) Aún echa más de menos a Sproingy.

Cuando Charkie llega, Steve le pide a Jorge que...

- A) Vaya a buscar más agua.
- B) Limpie la pelota de Charkie.
- C) Recoja a Sproingy.

Appendices

3. Ordena cronológicamente (del 1 al 5) estos hechos que pasan en el capítulo que acabas de ver. El número 3 te puede servir de guía para ordenar los demás.

___ Ted le dice a Jorge que debe limpiarse antes de comer.

___ Ted le dice a la profesora Wiseman que es un genio.

3 Jorge hace cosquillas a Ted.

___ Ted encuentra el barco en el congelador.

___ Ted le sugiere a Jorge que ayude a Betsy y Steve.

4. Marca con un círculo las palabras y expresiones que han aparecido en este vídeo.

you are grounded

help me

frog

truck

clean up

buy sweets

muddy

rubber duck

lunch

sponge

bubble-maker

take a bath

Nombre y apellidos _____ Fecha _____ Curso _____

“Jorge y el Puesto de Limonada”

1. Marca si estos enunciados son verdaderos (V) o falsos (F).

- Ted dice que hay limonada acabada de comprar en la nevera. V / F
- Para el portero, medio vaso de limonada no es suficiente. V / F
- Jorge da toda la limonada que quieren a Charkie y Gnocchi. V / F
- Betsy lleva toda la mañana buscando al perro Charkie. V / F
- Betsy dice que han ganado suficiente para comprar dos balones. V / F

2. Escoge la respuesta correcta (A, B o C).

Los pantalones cortos que Ted lleva en la foto...

- A) Son de un safari que hizo su abuelo.
- B) Eran de su hermano.
- C) Se los dejó un amigo.

El consejo que la da el portero a Jorge es que...

- A) Ahorre dinero utilizando siempre los mismos vasos.
- B) La gente prefiere limonada sin babas de animales.
- C) No les dé limonada gratis a Charkie y Gnocchi.

El tendero le dice a Jorge que...

- A) Puede ponerse delante de su tienda a vender limonada.
- B) Él vende la misma limonada que Jorge.
- C) Vendiendo en la calle la limonada se calienta y debe enfriarla.

Jorge intenta vender limonada por la ciudad pero...

- A) No hay demasiada gente en las calles.
- B) A la gente no le gusta la limonada.
- C) La gente está demasiado ocupada.

¿Por qué compran limonada los trabajadores del zoo?

- A) Porque les gusta más que el agua.
- B) Porque en el zoo hoy no hay agua para beber.
- C) Porque los elefantes se han comido todo lo que llevaban los trabajadores en sus bolsas.

Appendices

3. Ordena cronológicamente (del 1 al 5) estos hechos que pasan en el capítulo que acabas de ver. El número 3 te puede servir de guía para ordenar los demás.

___ El tendero le dice a Jorge que puede llevarse una caja de limones.

___ Jorge recuerda cómo convertir dos vasos de limonada en cuatro.

3 Betsy le pregunta si está ganando mucho dinero.

___ Ted le pregunta a Jorge si ha llevado la limonada a pasear.

___ Ted le dice a Jorge que se porte bien y sea un buen monito.

4. Marca con un círculo las palabras y expresiones que han aparecido en este vídeo.

make money

spare a moment

people

licking

ice-cold lemonade

orange juice

stand

soccer ball

looking forward to coming home

business

play basketball

jug

Nombre y apellidos _____ Fecha _____ Curso _____

“Jorge trabaja de Portero”

1. Marca si estos enunciados son verdaderos (V) o falsos (F).

- Hundley piensa que Jorge es organizado y limpio. V / F
- El dueño del apartamento 217 es jugador de béisbol. V / F
- La señora mayor le dice a Hundley que él y Jorge están guapísimos. V / F
- El dueño del tinte dice que tardará una hora exacta en limpiar la camisa. V / F
- Jorge piensa que los paquetes que reciben en el portal son para él. V / F

2. Escoge la respuesta correcta (A, B o C).

Para Hundley, su trabajo más importante es...

- A) Recibir y guardar todos los paquetes hasta que llegue su dueño.
- B) Proteger el portal de Jorge.
- C) Siempre sonreír a todo el mundo.

Los balones que el dueño del 217 ha encargado son...

- A) De China.
- B) De un importante jugador de varios deportes.
- C) Muy caros y valiosos.

En el tinte, el portero dice que puede esperarse porque...

- A) No le importa lo que pase en el portal.
- B) Sabe que Hundley mantendrá el portal perfectamente en orden.
- C) Sabe que Jorge está ayudando a Hundley.

Cuando Jorge entrega al dueño de la 217 una caja abierta,...

- A) Piensa que Jorge la ha abierto.
- B) Piensa que Hundley la ha abierto.
- C) Piensa que los mensajeros la han repartido en mal estado.

El portero, al volver del tinte, ...

- A) Invita a Jorge y Hundley a comer.
- B) Le ofrece a Jorge ser su ayudante el día siguiente.
- C) Le ofrece a Jorge su trabajo y él se va de vacaciones.

Appendices

3. Ordena cronológicamente (del 1 al 5) estos hechos que pasan en el capítulo que acabas de ver. El número 3 te puede servir de guía para ordenar los demás.

___ El portero dice que es la última camisa que tiene limpia al manchársela de aceite.

___ Jorge entiende que las pelotas no son para él.

3 Jorge se da cuenta que el uniforme de portero le gusta mucho.

___ El dueño del apartamento 217 dice que espera muchos paquetes para hoy.

___ Hundley sabe hacer de defensa.

4. Marca con un círculo las palabras y expresiones que han aparecido en este vídeo.

spotless

computer

wrap up

boxes

rushed right up

uniform

monkey

contrive

delivery

clean

clothes

stain

Nombre y apellidos _____ Fecha _____ Curso _____

“Un Mono Esquiador”

1. Marca si estos enunciados son verdaderos (V) o falsos (F).

- La nieve es muy profunda y Jorge no puede jugar. V / F
- Bill tiene unos esquís nuevos para Jorge. V / F
- Bill se va a casa porque no encuentra lo que hace un ruido extraño. V / F
- Una de las niñas quiere darle el trineo a Jorge. V / F
- Ted le dice a Jorge que tiene plátanos para él dentro de la casa. V / F

2. Escoge la respuesta correcta (A, B o C).

Para Ted, la primera regla para jugar en la nieve es...

- A) Saber esquiar.
- B) No tener miedo al frío.
- C) Poder salir de casa.

Bill escucha un ruido y le dice a Jorge que...

- A) Podría ser el monstruo de las nieves.
- B) Irá a echar un vistazo.
- C) Le acompañe a ver qué puede ser.

Las niñas oyen un ruido raro y piensan que...

- A) Serán niños montando en trineo.
- B) Será un oso polar que se ha escapado del zoo.
- C) Alguien se ha perdido y hace ruido para que lo encuentren.

Mientras anda, la única cosa que mantiene a Jorge en pie es pensar que...

- A) Podrá hacer un muñeco de nieve.
- B) Podrá descansar junto al fuego.
- C) Podrá tomar un chocolate caliente.

El dueño del cerdo Mike dice que el cerdo...

- A) Nunca antes había visto nieve.
- B) Quiso salir a buscar a su madre.
- C) Se escapó porque quería dar la vuelta al mundo.

Appendices

3. Ordena cronológicamente (del 1 al 5) estos hechos que pasan en el capítulo que acabas de ver. El número 3 te puede servir de guía para ordenar los demás.

___ Bill le dice a Jorge que para subir una pendiente, haga zigzag.

___ Jorge se pregunta cómo el cerdo Mike ha llegado hasta allí.

3 Jorge puede ver casas y granjas desde la cima de la montaña.

___ Con esquís, Jorge cree poder ir allá donde haya nieve.

___ Para las niñas, la bajada de Jorge es impresionante.

4. Marca con un círculo las palabras y expresiones que han aparecido en este vídeo.

hot cocoa

ice

sled

snowshoes

shovel

leave him here

hill

all the way

skiing

vessels

boots

water cycle

Nombre y apellidos _____ Fecha _____ Curso _____

“Jorge hace una casa para palomas”

1. Marca si estos enunciados son verdaderos (V) o falsos (F).

- Jorge nunca había visto pájaros que llevaran una anilla. V / F
- Ted le dice a Jorge que su dibujo de un pájaro no es muy bueno. V / F
- Ted no quiere que la paloma entre en su casa. V / F
- El portero piensa que George le ha ensuciado la entrada. V / F
- El portero dice que la paloma tiene un buen sentido de la orientación. V / F

2. Escoge la respuesta correcta (A, B o C).

Cuando Jorge se cae de la silla, Ted le dice que...

- A) Ya se lo había advertido.
- B) La vida de un artista está llena de golpes.
- C) Tendrá que llamar a una ambulancia.

Para Jorge, la mejor forma de tener un árbol en casa sería...

- A) Fabricarlo.
- B) Pedir al portero que le ayude a subir uno de la calle.
- C) Comprar y subirlo con una grúa.

Para el portero, la paloma que encuentra George es...

- A) Un animal de compañía.
- B) La paloma perfecta.
- C) Casi una paloma mensajera.

Ted le dice a Jorge que ha comprado el árbol para que...

- A) Las palomas puedan ir a vivir allí.
- B) Se posen los pájaros y él pueda dibujarlos.
- C) Puedan tener un pequeño bosque en la terraza.

¿Qué piensa la paloma sobre Jorge?

- A) Ella no sabe qué tipo de animal es Jorge.
- B) Ella cree que Jorge es el más bueno de todos los animales
- C) Ella piensa que Jorge es amigo de todos los animales.

3. Ordena cronológicamente (del 1 al 5) estos hechos que pasan en el capítulo que acabas de ver. El número 3 te puede servir de guía para ordenar los demás.

- ___ Ted piensa que Jorge quiere intimidad.
- ___ El portero parece conocer la paloma con la anilla amarilla.
- _3_ Jorge cree que ha puesto demasiada agua en el árbol de arcilla.
- ___ Ted le dice a Jorge que los pájaros prefieren árboles auténticos.
- ___ La paloma entra en casa de Ted y Jorge.

4. Marca con un círculo las palabras y expresiones que han aparecido en este vídeo.

- | | | | |
|-----------|--------------------------|---------------|---------|
| draw | this called for research | food | dirt |
| bird | squab | balcony | teacher |
| came back | furthermore | homing pigeon | tree |

Nombre y apellidos _____ Fecha _____ Curso _____

“Jorge el Tendero”

1. Marca si estos enunciados son verdaderos (V) o falsos (F).

- Ted le dice a Jorge que ya tiene un dinosaurio igual al del escaparate. V / F
- Al tendero le vendría bien tener ayuda hasta que su hijo vuelva. V / F
- La niña en el carro de compra le pregunta a su madre si puede comprar patatas. V / F
- El cliente quiere el paquete más pequeño de la pirámide de papel. V / F
- El portero le desea Feliz Verano al tendero. V / F

2. Escoge la respuesta correcta (A, B o C).

Ted le dice a Jorge que el horno de la tienda de juguetes es muy caro y que...

- A) Puede utilizar el horno de la cocina.
- B) La semana pasada ya le compró un juguete caro.
- C) No necesita ningún horno porque puede comprar pasteles en la tienda.

El cliente quiere plátanos muy dulces para...

- A) Hacerle una tarta de plátano a su madre.
- B) Un mono que tiene en casa.
- C) Hacer yogur de plátano y venderlo.

Al caerle la pirámide de papel encima, Jorge da gracias que ...

- A) Él va al gimnasio y puede aguantar que le caigan cosas encima.
- B) En esa tienda no vendan ladrillos.
- C) El tendero no lo ha visto.

El cliente le dice al tendero que no sabe lo que le paga al mono pero que...

- A) Lo contrate para que esté todos los días.
- B) Contrate a más monos ya que alegran la compra.
- C) Se merece el doble.

Dentro del escaparate, el tendero le dice a Jorge que...

- A) Salgan porque tiene calor.
- B) Salgan porque la gente les mira.
- C) Está contratado durante todo el verano.

Appendices

3. Ordena cronológicamente (del 1 al 5) estos hechos que pasan en el capítulo que acabas de ver. El número 3 te puede servir de guía para ordenar los demás.

- ___ Jorge decide buscar clientes que necesiten ayuda.
- ___ Ted le dice a Jorge que llega tarde al trabajo.
- _3_ El tendero dice que para comprar fruta hay que arriesgarse.
- ___ El tendero le dice a Jorge que es el empleado del mes.
- ___ El tendero dice que ojalá tuviera un empleado con cuatro manos.

4. Marca con un círculo las palabras y expresiones que han aparecido en este vídeo.

supermarket	ball	window display	Merry Christmas
bananas	customer	parsley	lemon-cicles
mobile phone	hands	employee	regardless

Nombre y apellidos _____ Fecha _____ Curso _____

“Las Abejas y el Oso”

1. Marca si estos enunciados son verdaderos (V) o falsos (F).

- A Jorge nunca le han gustado los sonidos que se oyen en el campo. V / F
- Según Ted, no se ven osos por allí desde hace más de veinte años. V / F
- Ted le pide a Jorge que le ayude a regar las plantas. V / F
- Jorge piensa que el plan del niño Bill para sacar la colmena les gustará a las abejas. V / F
- La chica no puede quitar la colmena porque hay demasiadas ramas. V / F

2. Escoge la respuesta correcta (A, B o C).

Lo que más le gusta a Jorge es...

- A) Jugar al escondite con las abejas.
- B) Comerse la miel con pan recién hecho.
- C) Comerse la miel directamente del bote.

Al ver que la abeja vuelve, Jorge se piensa que ella...

- A) Quiere ser su amiga.
- B) Quiere que le devuelva la miel que se ha comido.
- C) Quiere dar un paseo con él.

Bill quiere quitar la colmena del árbol...

- A) Porque quiere comerse la miel.
- B) Porque quiere hacer un casa en el árbol.
- C) Porque no quiere que las abejas piquen a sus conejos.

Al ver la colmena, Ted propone...

- A) Llamar a unos profesionales para quitarla.
- B) Que Jorge suba al árbol y la quite.
- C) Llamar al señor y la señora Redkins.

Jorge espera poder volver a ver al oso algún día...

- A) Para poder jugar con él.
- B) Para ir al zoo a ver a los pandas con él.
- C) Para poder compartir sándwiches de miel.

Appendices

3. Ordena cronológicamente (del 1 al 5) estos hechos que pasan en el capítulo que acabas de ver. El número 3 te puede servir de guía para ordenar los demás.

- ___ La señora Redkins le prepara pan casero con miel a Jorge.
- ___ A Jorge le pica una abeja.
- _3_ Ted se queda con los Sres. Redkins a ayudar con las colmenas.
- ___ El Jardinero pregunta si tienen algo de comer.
- ___ Bill quiere mover la colmena con su traje anti-abejas.

4. Marca con un círculo las palabras y expresiones que han aparecido en este vídeo.

- | | | | |
|-------|-------------------|------------|--------|
| bee | wax | homemade | bear |
| hive | polish your nails | proper way | flower |
| swing | honey | buy a farm | sting |

Nombre y apellidos _____ Fecha _____ Curso _____

“Proyecto Reciclaje”

1. Marca si estos enunciados son verdaderos (V) o falsos (F).

- Los bomberos organizan un concierto una vez al año. V / F
- Ted le dice al portero que no se pierda el concierto. V / F
- Jorge sabe exactamente qué instrumentos hará con lo que ha reciclado. V / F
- El portero le da permiso a George para quedarse algunos objetos de la portería. V / F
- El bombero que conduce se pregunta porque nunca reciben llamadas normales. V / F

2. Escoge la respuesta correcta (A, B o C).

Mientras ensayan, la bombero Andie le dice a Jorge que sonaran mejor ya que...

- A) Hoy es sólo su primer día de ensayo.
- B) Ahora sólo están calentando.
- C) Vendrá un famoso director de orquesta a dirigirlos.

Ted le dice a Jorge que no puede tocar su tuba porque...

- A) Primero tiene que ir a clases para aprender cómo se toca.
- B) Los monos no tocan la tuba.
- C) Es un instrumento muy delicado.

Ted le dice a Jorge que con todo lo que ha cogido de la basura...

- A) Tiene para toda una orquesta.
- B) Puede hacer una batería, un violín y una pandereta.
- C) Puede construir un castillo.

Ted le dice a Jorge que mejor no coja las tapas del fregadero ya que...

- A) Hará mucho ruido y los vecinos se quejarán.
- B) Las necesita él para ir a ensayar con los bomberos.
- C) Son las tapas de sus mejores cacerolas.

La primera lección que Jorge aprende como líder del grupo es que...

- A) Antes de tocar en grupo los músicos deben tocar solos.
- B) Tienen que ensayar mucho para que todo salga bien.
- C) A los músicos les encanta la comida gratis.

3. Ordena cronológicamente (del 1 al 5) estos hechos que pasan en el capítulo que acabas de ver. El número 3 te puede servir de guía para ordenar los demás.

- ___ El portero le pregunta a Ted si ha ido bien el ensayo.
- ___ Hundley siente curiosidad por saber por qué Jorge necesita galletas de perro.
- _3_ Ted le dice a Jorge que va a comprar arroz ya que no les queda.
- ___ La bombero Andie le dice a Jorge que lo que ha reciclado son buenos instrumentos.
- ___ Jorge decide montar su propio grupo de música.

4. Marca con un círculo las palabras y expresiones que han aparecido en este vídeo.

- | | | | |
|--------------|-------|-----------|----------|
| instruments | drums | briefcase | play |
| eat a carrot | band | musicians | practice |
| warming up | sure | lute | realized |

Nombre y apellidos _____ Fecha _____ Curso _____

“Jorge se va a buscar conejos”

*L1S

1. Marca si estos enunciados son verdaderos (V) o falsos (F).

- Ted nunca sabe cuando tiene que sacar a Jorge de la ciudad. V / F
- Jorge está acostumbrado a ver conejos de cerca. V / F
- Bill dice que es hora de repartir los periódicos. V / F
- A la ardilla le gusta la sorpresa de que la coja un mono. V / F
- Según Bill, el primer paso para acariciar a un conejo es abrir el pestillo de la jaula. V / F

2. Escoge la respuesta correcta (A, B o C).

Al llegar al campo, Ted le dice a Jorge que...

- A) Vaya a dar una vuelta.
- B) Vaya a comprar la comida para hacer la cena.
- C) Respire hondo y llene los pulmones de aire campestre.

Al ver a los conejos, Ted le pregunta a Bill...

- A) Si los conejos tienen nombre.
- B) Qué edad tienen los conejos.
- C) Si los va a dar en adopción.

Jorge sabrá cuantos conejitos se han escapado...

- A) Preguntándole a Bill cuantos tenía.
- B) Preguntándole a la mama conejo.
- C) Contando los cuencos que hay en la jaula.

Para buscar al conejo que le faltaba, Jorge necesita...

- A) Una brújula.
- B) Subirse a un árbol y buscarlo.
- C) Un experto en conejitos.

Bill le dice a Jorge que la clave para acariciar a un conejo...

- A) Está en mantener el control.
- B) Tiene que hablar con el veterinario.
- C) Tiene que lavarse las manos.

Appendices

3. Ordena cronológicamente (del 1 al 5) estos hechos que pasan en el capítulo que acabas de ver. El número 3 te puede servir de guía para ordenar los demás.

- ___ Bill le dice a Jorge que es un chico muy cuidadoso.
- ___ Jorge piensa que no todas las huellas llevan a los conejitos.
- _3_ La mamá conejo sabe que debe quedarse en su casita.
- ___ Ted le dice a Jorge que suelte el aire limpio.
- ___ Bill dice que el conejo de nombre “peludo” es guai.

4. Marca con un círculo las palabras y expresiones que han aparecido en este vídeo.

- | | | | |
|-----------------|-------------|---------|-----------|
| surprises | tractor | fencing | too fast |
| easy | country air | hutch | landscape |
| petting a bunny | city | plastic | time |

“Jorge se va a buscar conejos”

***L2S**

1. Marca si estos enunciados son verdaderos (V) o falsos (F).

- Ted nunca sabe cuando tiene que sacar a Jorge de la ciudad. V / F
- Jorge está acostumbrado a ver conejos de cerca. V / F
- Bill dice que es hora de repartir los periódicos. V / F
- A la ardilla le gusta la sorpresa de que la coja un mono. V / F
- Según Bill, el primer paso para acariciar a un conejo es abrir el pestillo de la jaula. V / F

2. Escoge la respuesta correcta (A, B o C).

Al llegar al campo, Ted le dice a Jorge que...

- A) Vaya a dar una vuelta.
- B) Vaya a comprar la comida para hacer la cena.
- C) Respire hondo y llene los pulmones de aire campestre.

Al ver a los conejos, Ted le pregunta a Bill...

- A) Si los conejos tienen nombre.
- B) Qué edad tienen los conejos.
- C) Si los va a dar en adopción.

Jorge sabrá cuantos conejitos se han escapado...

- A) Preguntándole a Bill cuantos tenía.
- B) Preguntándole a la mama conejo.
- C) Contando los cuencos que hay en la jaula.

Para buscar al conejo que le faltaba, Jorge necesita...

- A) Una brújula.
- B) Subirse a un árbol y buscarlo.
- C) Un experto en conejitos.

Bill le dice a Jorge que la clave para acariciar a un conejo...

- A) Está en mantener el control.
- B) Tiene que hablar con el veterinario.
- C) Tiene que lavarse las manos.

Appendices

3. Ordena cronológicamente (del 1 al 5) estos hechos que pasan en el capítulo que acabas de ver. El número 3 te puede servir de guía para ordenar los demás.

- ___ Bill le dice a Jorge que es un chico muy cuidadoso.
- ___ Jorge piensa que no todas las huellas llevan a los conejitos.
- _3_ La mamá conejo sabe que debe quedarse en su casita.
- ___ Ted le dice a Jorge que suelte el aire limpio.
- ___ Bill dice que el conejo de nombre “fuzzy” es guai.

4. Marca con un círculo las palabras y expresiones que han aparecido en este vídeo.

- | | | | |
|-----------------|-------------|---------|-----------|
| surprises | tractor | fencing | too fast |
| easy | country air | hutch | landscape |
| petting a bunny | city | plastic | time |

Nombre y apellidos _____ Fecha _____ Curso _____

“El barco de las maravillas”

1. Marca si estos enunciados son verdaderos (V) o falsos (F).

- A Bill le ha ayudado su padre a hacer el barco. V / F
- Este año Bill cree que tiene muchas posibilidades de ganar. V / F
- Hacía muchos años que Jorge quería construir un ferry. V / F
- Al hundirse el barco de Bill, Jorge piensa que lo ha puesto del revés. V / F
- Cuando Bill ve que su barco se hunde, le dice a Jorge que le ha hecho un favor. V / F

2. Escoge la respuesta correcta (A, B o C).

Jorge se queda con Bill, y Ted dice que él...

- A) Estará en las gradas guardándoles el sitio.
- B) Se va a comprar al supermercado y los recogerá más tarde.
- C) Va a construir un barco para ganar el concurso.

Bill le dice a Jorge que para poder hacer el barco...

- A) Ha trabajado en la tienda de su madre todo el verano.
- B) Ha repartido periódicos durante nueve semanas.
- C) Le ha tenido que pedir dinero a su padre.

Jorge piensa que para hacer un barco con prisa...

- A) Lo mejor es preguntar a los demás concursantes cómo lo han hecho.
- B) Un cubo de juguetes resultará útil.
- C) La ardilla le dará pistas desde su árbol

Al ver el barco que ha hecho Jorge, Bill dice que creía que...

- A) Jorge no sabía hacer barcos.
- B) Era casi tan bueno como el de la señora Renkins.
- C) Los chicos de ciudad lo compraban todo.

En la cinta del premio de Jorge pone...

- A) Premio de consolación.
- B) Mejor barco de un mono.
- C) Premio al mejor aprendiz.

Appendices

3. Ordena cronológicamente (del 1 al 5) estos hechos que pasan en el capítulo que acabas de ver. El número 3 te puede servir de guía para ordenar los demás.

- ___ Bill pregunta dónde está su barco.
- ___ Jorge ve un gran barco lleno de coches por el río.
- _3_ Bill dice que va a devolver su bici a casa.
- ___ Bill aconseja a Jorge que apunte su barco a la competición.
- ___ Bill le pide a Jorge que cuide un rato de su barco.

4. Marca con un círculo las palabras y expresiones que han aparecido en este vídeo.

- | | | | |
|-------|------------|----------|-------------|
| cars | fetch | floating | competition |
| swim | contest | model | badger |
| built | over there | boat | phone |

Nombre y apellidos _____ Fecha _____ Curso _____

“Jorge y los Constructores”

1. Marca si estos enunciados son verdaderos (V) o falsos (F).

- La Sra. Renkins le dice a Jorge que tendrá el barco arreglado esa misma tarde. V / F
- Bill le dice a Jorge que no se acerque mucho a los castores cuando trabajan. V / F
- Para Jorge, construir una presa dos veces el mismo día es muy fácil. V / F
- Bill le dice a Jorge que el agua es muy importante para los castores. V / F
- Bill no entiende por qué los castores no están contentos cuando sacan el árbol del río. V / F

2. Escoge la respuesta correcta (A, B o C).

Cuando tiene el barco reparado, Jorge vuelve al estanque pero...

- A) La vaca se ha bebido el agua del estanque.
- B) En vez del estanque hay una muralla muy rara en el arroyo.
- C) No puede jugar con su barco porque el agua está contaminada.

Jorge piensa que si los castores no quieren que él les mire...

- A) No les va a dejar su barco para jugar.
- B) No les invitará a su casa a comer.
- C) Se limitará a navegar con su barco.

La Sra. Renkins le pregunta a Jorge si...

- A) Se divierte con su barco.
- B) Los castores le han molestado mucho.
- C) Irá a cenar a su casa esa noche.

A Jorge no le importa la lluvia, ya que piensa que así...

- A) Los castores se irán a otro sitio.
- B) Su estanque será más grande.
- C) Tendrá agua aunque su presa se rompa.

Cuando Jorge juega con los castores y estos se comen su barco, él piensa que...

- A) Le pedirá a la Sra. Renkins que le haga otros.
- B) No necesita ningún barco, ya que puede jugar con los castores.
- C) Irá a la tienda y comprará muchos barcos en las rebajas.

Appendices

3. Ordena cronológicamente (del 1 al 5) estos hechos que pasan en el capítulo que acabas de ver. El número 3 te puede servir de guía para ordenar los demás.

___ Jorge y Bill observan como los castores trabajan en equipo para defender su hogar.

___ Jorge se ha olvidado el barco y vuelve a casa a buscarlo.

3 La Sra. Renkins le dice a Jorge que coja la madera que necesite.

___ Bill dice a Jorge que ha sacado fotos a los castores.

___ Jorge decide que hará una presa para él solo.

4. Marca con un círculo las palabras y expresiones que han aparecido en este vídeo.

pen

toy boat

dam

water

beavers

I'm coming

mast

creek

home

ocean

pansy

pond

Nombre y apellidos _____ Fecha _____ Curso _____

“Un mono sobre patines”

1. Marca si estos enunciados son verdaderos (V) o falsos (F).

- Ted le dice a Jorge que debería haber ordenado el salón ayer. V / F
- El dueño de la tienda le pide a Jorge que le devuelva los patines. V / F
- El dueño de la tienda le pregunta a su mujer si cree que volverán a ver a Jorge. V / F
- Jorge le deja los patines a Hundley y le dice que vuelva en diez minutos. V / F
- Hundley está convencido que de no ser por Jorge y el gato, él habría aprendido a patinar. V / F

2. Escoge la respuesta correcta (A, B o C).

Ted le dice a Jorge que no puede dejar los juguetes por el suelo ya que...

- A) Tienen visita esa misma tarde.
- B) Alguien podría hacerse daño.
- C) Las normas del edificio lo prohíben.

El dueño de la tienda dice que tienen muchos patines pero se pregunta...

- A) Por qué no han recibido los últimos que pidió.
- B) Si le han robado algunos la noche anterior.
- C) Cómo puede hacer para venderlos.

A Jorge le gustan sus patines nuevos pero se pregunta...

- A) Para qué servirán las dos cosas negras.
- B) Que dirá Ted cuando le vea patinar.
- C) Por qué no le han dado unos patines de otro color.

Las niñas piensan que Hundley está triste porque...

- A) No tiene un compañero con quien patinar.
- B) Le han fastidiado su paseo con patines.
- C) Tiene hambre y no tiene dinero para comprar comida.

Al oír el piropo de la señora del parque, Hundley piensa que...

- A) Se presentará a un concurso de belleza para perros.
- B) La señora lo ha confundido con otro perro.
- C) Puede ser famoso en toda la ciudad.

Appendices

3. Ordena cronológicamente (del 1 al 5) estos hechos que pasan en el capítulo que acabas de ver. El número 3 te puede servir de guía para ordenar los demás.

___ Jorge piensa que el perro volador no puede ser Hundley porque este nunca volaría.

___ Jorge descubre que patinar puede ser divertido.

3 El portero le dice a Hundley que lo deja al mando.

___ En la tienda, la vendedora le dice a Jorge que quizá debería practicar fuera.

___ Hundley se arrepiente de haberse puesto patines.

4. Marca con un círculo las palabras y expresiones que han aparecido en este vídeo.

laces

how to stop

thing

skates

resign

thought

wheel shoes

see

bottle

dog

pasta

trust

Nombre y apellidos _____ Fecha _____ Curso _____

“Jorge y los Renacuajos”

1. Marca si estos enunciados son verdaderos (V) o falsos (F).

- A Bill no le gusta explorar la laguna porque el agua está sucia. V / F
- Bill da a los renacuajos lechuga hervida para comer. V / F
- Jorge sabe que dentro de un mes los renacuajos se convertirán en ranas. V / F
- Bill le pregunta a Jorge si sus renacuajos le están dando problemas. V / F
- Jorge piensa que las ranas son menos divertidas que los renacuajos. V / F

2. Escoge la respuesta correcta (A, B o C).

Bill le dice a Jorge que quiere que vigile a los renacuajos de cerca ya que...

- A) Su crecimiento es asombroso.
- B) Se escapan del bote fácilmente.
- C) Son muy valiosos.

Jorge deja que los renacuajos disfruten de un baño rápido en la laguna mientras él,...

- A) Llama a Bill para que los vaya a ver.
- B) Hace los deberes.
- C) Va a buscar comida.

Al ver a Jorge poner la cabeza bajo el agua, el señor de la barca le pregunta si...

- A) Tiene sed y bebe agua del lago.
- B) Comprueba el nivel del agua.
- C) Está aprendiendo a respirar bajo el agua.

Jorge le pide a Ted que vayan al lago pero él le dice que...

- A) Irán dentro de unos días cuando florezcan los árboles.
- B) Ha quedado con sus amigos y no puede ir.
- C) Acaban de volver e irán el mes que viene.

Al ver las ranas en el suelo, Bill le dice a Jorge que...

- A) Tiene que contar que estén todas.
- B) Ahora las pueden llevar al zoo.
- C) Ha tenido una buena idea soltando a los renacuajos.

Appendices

3. Ordena cronológicamente (del 1 al 5) estos hechos que pasan en el capítulo que acabas de ver. El número 3 te puede servir de guía para ordenar los demás.

___ Jorge se lleva el bote de renacuajos a casa.

___ Bill hace fotos en los alrededores del lago y Jorge le sigue

3 Ted aparece con el coche y le dice a Jorge que es hora de volver a la ciudad .

___ Bill le muestra a Jorge la mariposa en la que se ha convertido su oruga.

___ Jorge explora la laguna con unas gafas de agua prestadas.

4. Marca con un círculo las palabras y expresiones que han aparecido en este vídeo.

creel

tadpoles

amazing

frogs

lake

eat breakfast

caterpillar

sea

hinder

bowl

lagoon

grow

Nombre y apellidos _____ Fecha _____ Curso _____

“Jorge juega al minigolf”

1. Marca si estos enunciados son verdaderos (V) o falsos (F).

- Steve no quiere que Jorge juegue a baloncesto. V / F
- Jorge sabe que es bueno al minigolf al terminar el primer hoyo. V / F
- Steve le dice a Jorge que quizás los monos no están hechos para jugar al golf. V / F
- Betsy no sabe dónde encontrar cosas chulas para construir un campo de golf. V / F
- Steve propone a Betsy y a Jorge jugar al minigolf ‘de verdad’ al día siguiente. V / F

2. Escoge la respuesta correcta (A, B o C).

Jorge pasea por la calle y se pregunta...

- A) Si va a llover.
- B) Si los números de los edificios, que van subiendo, paran alguna vez.
- C) Por qué construyen los edificios tan juntos.

Betsy le dice a Jorge que el minigolf es genial y que hay...

- A) Un castillo y un molino.
- B) Un castillo, un molino, un río y un puente.
- C) Un castillo, un dragón, un camión y un puente.

Jorge piensa que si no juega al golf,...

- A) Se va a aburrir en casa.
- B) Perderá la oportunidad de que sus amigos le enseñen a jugar.
- C) Sus amigos creerán que no se le da bien.

Ted le dice a Jorge que puede jugar con todo lo que ha cogido pero que...

- A) Cuando termine lo devuelva todo a su sitio.
- B) No lo saque del edificio.
- C) No monte en el monociclo sin casco.

Betsy le dice a Steve que ‘en su campo’...

- A) Tiene que usar un palo de color rojo.
- B) Tiene que usar un palo especial para cada hoyo.
- C) Los palos están escondidos y los tiene que buscar para poder jugar.

3. Ordena cronológicamente (del 1 al 5) estos hechos que pasan en el capítulo que acabas de ver. El número 3 te puede servir de guía para ordenar los demás.

- ___ Jorge piensa que los puntos del minigolf pueden confundir a un mono.
- ___ Steve dice que la tía Margaret quiere que bajen a cenar en una hora.
- _3_ Jorge piensa que es fácil lograr una puntuación alta en el golf.
- ___ Jorge va paseando y le entra curiosidad por un ruido.
- ___ Betsy propone jugar al minigolf en lugar del baloncesto.

4. Marca con un círculo las palabras y expresiones que han aparecido en este vídeo.

end of the game

hat

miniature golf

hole

sport

new record

great stuff

windmill

stance

number

score

peacock

Nombre y apellidos _____ Fecha _____ Curso _____

“Jorge contra el invierno”

1. Marca si estos enunciados son verdaderos (V) o falsos (F).

- Ted le dice a Jorge que no olvide los esquís cuando salga de casa en invierno. V / F
- Jorge piensa que Bill le va a pedir que cuide de los conejos. V / F
- Al ver el balón de agua helado, Jorge piensa que su plan no funciona. V / F
- Jorge piensa que el invierno es como una montaña rusa. V / F
- Ted le dice a Jorge que está nevando dentro de la casa. V / F

2. Escoge la respuesta correcta (A, B o C).

Bill le dice a Jorge que llevará a los conejos...

- A) A una granja cercana.
- B) A casa de su abuela.
- C) Al bosque para liberarlos.

Jorge encuentra interesante que la pompa se hiele pero...

- A) No es lo que él quiere.
- B) No tiene más jabón para hacer más pompas.
- C) Ted le dice que entre en casa que hace mucho frío.

Jorge se pregunta cómo puede divertirse un mono...

- A) Si no tiene amigos.
- B) Si tiene que estar en el campo y no en la ciudad.
- C) Si todas sus cosas se congelan.

Al ver que la ardilla muerde los ladrillos de hielo, Jorge piensa que...

- A) A las ardillas no les gusta jugar los bolos.
- B) Las ardilla tiene sed y por eso muerde los ladrillos de hielo.
- C) La ardilla está aburrida como él.

Jorge piensa que eso de inventar cosas sí servía para algo ya que...

- A) Podría contárselo a Bill cuando volviera.
- B) No se aburrió hasta que llegó la nieve.
- C) No pudo ir a la escuela por la nieve.

Appendices

3. Ordena cronológicamente (del 1 al 5) estos hechos que pasan en el capítulo que acabas de ver. El número 3 te puede servir de guía para ordenar los demás.

- ___ Ted vuelve a casa y trae chocolate para tomar caliente.
- ___ Ted le dice a Jorge que esperan nieve pronto.
- _3_ Ted le dice a Jorge que el día siguiente hará mucho frío.
- ___ Bill le pide un favor a Jorge.
- ___ Jorge se da cuenta que la idea de poner agua en los botes ha funcionado.

4. Marca con un círculo las palabras y expresiones que han aparecido en este vídeo.

- | | | | |
|------------|-------|---------|--------|
| love | house | bowling | forget |
| day | rain | freeze | mitten |
| having fun | rust | winter | warm |

APPENDIX F. SINGLE WORDS AND MULTIWORD UNITS CATEGORISATION

Table F.1 – Categorisation of new target words

New target words (80)							
Single words (47) (58.75%)							
Nouns (30) (37.5%)					Verbs (12) (15%)	Adjectives (4) (5%)	Adverbs (1) (1.25%)
Keys Keypads Foundation Waffle Magnet Truck	Business Stand Delivery Stain Sled Hill	Dirt Customer Employee Hive Band Hutch Contest	Beavers Dam Creek Pond Tadpoles Caterpillar	Lagoon Hole Handle Score Bowling	Thought Launch Supply Follow Licking Sting Realized Built Floating Grow Forget Trust	Muddy Spotless Homemade Warm	Sure
Multiword units (33) (41.25%)							
Collocations (14) (17.5%)		Compounds (8) (10%)		Phrasal Verbs (6) (7.5%)	Linking Adverbials (3) (3.75%)	Free word combinations (2) (2.5%)	
Shown live Go the wrong way Pull the lever Pretty impressive Shut the site down Be in a hurry Out of order Rushed right up All the way This called for research Proper way Petting a bunny Great Stuff Having fun		Warning bell Bubble-maker Homing pigeon Window display Lemon-cicles Country Air Wheel shoes Windmill		Lock up Hold up Stick out Clean up Come back Warming up	Too fast Over there How to stop	Looking forward to coming home Leave him here	

Table F.2 – Categorisation of known target words

Known target words (80)						
Single words (60) (75%)						
Nouns (51) (63.75%)					Verbs (6) (7.5%)	Adjectives (3) (3.75%)
Door	Refrigerator	Bird	Time	Lake	Skiing	Easy
Baby	Conductor	Bananas	Cars	Bowl	Draw	Amazing
Animals	Frog	Ball	Boat	Frogs	Play	Clean
Map	Lunch	Honey	Model	Number	Practice	
Food	People	Bear	Competition	Day	See	
Toy	Uniform	Bee	Home	House	Freeze	
Experiments	Monkey	Flower	Water	Winter		
Building	Boxes	Instruments	Skates	Service		
Blueberry	Balcony	Musicians	Dog	Boots		
Machine	Tree	City	Thing			
Construction		Surprises				
Multiword units (20) (25%)						
Collocations (11) (13.75%)		Compounds (5) (6.25%)	Linking Adverbial (1) (1.25%)	Speech Act Formulae (1) (1.25%)	Free word combinations (2) (2.5%)	
Great Job Hit the button Have a mission Be right back Want to see Cut in half Take a bath Make money Ice-cold lemonade Hot cocoa New record		Train Station Soccer ball Toy boat Miniature golf Snowshoes	End of the game	Merry Christmas	Help me I'm coming	

Table F.3 – Categorisation of distractors

Distractors (80)								
Single words (64) (80%)								
Nouns (49) (61.25%)					Verbs (7) (8.75 %)	Adjectives (4) (5%)	Adverbs (4) (5%)	
Snake	Sponge	Supermarket	Landscape	Laces	Piping Contrive Swim Fetch Resign Hinder Love	Gleeful Foreseeable Fair Relieved	Gently Steadily Furthermore Regardless	
Cage	Jug	Creel	Fencing	Rain				
Pilot	Clothes	Parsley	Phone	Sea				
Bicycle	Computer	Swing	Badger	Sport				
Endorsement	Ice	Wax	Ocean	Hat				
Wood	Shovel	Drums	Pen	Stance				
Gallery	Vessels	Lute	Mast	Peacock				
Telephone	Food	Briefcase	Pasta	Mitten				
Flag	Teacher	Tractor	Pansy	Rust				
Medicine	Squab	Plastic	Bottle	Service				
Multiword units (16) (20%)								
Collocations (7) (8.75%)	Compounds (4) (5%)		Phrasal Verbs (1) (1.25%)	Speech Act Formulae (1) (1.25%)		Free word combinations (3) (3.75%)		
No need to do it You are grounded Play basketball Polish your nails Eat a carrot Spare a moment Eat breakfast	Rubber duck Orange juice Water cycle Mobile phone		Wrap up	Good morning		It's three o'clock Buy sweets Buy a farm		

APPENDIX G. SINGLE WORDS AND MULTIWORDS UNITS PER EPISODE

Table G.1 – New and known target words, multiword units and distractors per episode

Episode	Target words			Distractors			
	Known		New	Known		New	
	Frequent			Related	Not related	Related	Not related
E1	Door (N) (3) Baby (N) (9)	Animals (N) (3) Map (N) (4)	Shown live (C) (1) Lock up (PV) (1) Go the wrong way (C)(1) Keys (N) (2)	Snake (N)	Good morning (SAF)	Cage (N)	Gently (Adv.)
E2	Food (N) (5) Great job (C) (1)	Toy (N) (3) Experiments (N) (3)	Launch (V)(6) Supply (V) (2) Keypads (N) (4) Pull the lever (C) (3)	Pilot (N)	Bicycle (N)	Gleeful (Adj.)	Endorsement (N)
E3	Building (N) (5) Blueberry (N) (5)	Machine (N) (2) Construction (N) (2)	Foundation (N) (2) Pretty impressive(C) (1) Waffle (N) (3) Shut the site down (C) (1)	Wood (N)	It's three o'clock (FWC)	Piping (V)	Foreseeable (Adj.)
E4	Hit the button (C) (2) Have a mission (C) (2)	Refrigerator (N) (8) Be right back (C) (1)	Magnet (N) (10) Be in a hurry (C) (2) Handle (N) (4) Hold up (PV) (1)	Gallery (N)	Telephone (N)	No need to do it (C)	Fair (Adj.)
E5	Train Station (Comp.) (3) Wants to see (C) (2)	Conductor (N) (1) Cut in half (C)(2)	Out of order (C) (2) Stick out (PV) (2) Follow (V) (2) Warning bell (Comp.)(2)	Flag (N)	Medicine (N)	Relieved (Adj.)	Steadily (Adv.)
E6	Frog (N) (3) Take a bath (C) (6)	Lunch (N) (3) Help me (FWC) (2)	Clean up (PV) Muddy (Adj.) Bubble-maker (Comp.) Truck (N)	Sponge (N)	Buy sweets (FWC)	Rubber duck (Comp.)	You are Grounded (C)
E7	People (N) (6) Soccer ball (Comp.) (8)	Make money (C) (2) Ice-cold lemonade (C) (2)	Looking forward to coming home (FWC) (2) Business (N) (2) Stand (N) (7) Licking (V) (1)	Orange juice (Comp.)	Play basketball (C)	Spare a moment (C)	Jug (N)

Appendices

Episode	Target words			Distractors			
	Known		New	Known		New	
	Frequent			Related	Not related	Related	Not related
E8	Uniform (N) (7) Monkey (N) (6)	Clean (Adj.) (5) Boxes (N) (3)	Spotless (Adj.) (2) Rushed right up (C) (2) Delivery (N) (6) Stain (N) (4)	Clothes (N)	Computer (N)	Wrap up (PV)	Contrive (V)
E9	Skiing (V) (3) Snowshoes (Comp.) (4)	Hot cocoa (C) (2) Boots (N) (2)	Leave him here (FWC) (2) Sled (N) (4) Hill (N) (7) All the way (C) (2)	Ice (N)	Water cycle (Comp.)	Shovel (N)	Vessels (N)
E10	Balcony (N) (3) Tree (N) (11)	Draw (V) (2) Bird (N) (4)	Came back (PV) (2) This called for research (C) (2) Homing pigeon(Comp.)(3) Dirt (N) (2)	Food (N)	Teacher (N)	Squab (N)	Furthermore (Adv.)
E11	Bananas (N) (5) Merry Christmas (SAF) (3)	Ball (N) (4) Hands (N) (4)	Customer (N) (4) Window display (Comp.) (3) Employee (N) (3) Lemon-cicles (Comp.) (4)	Supermarket (N)	Mobile phone (Comp.)	Parsley (N)	Regardless (Adv.)
E12	Honey (N) (9) Bear (N) (22)	Bee (N) (7) Flower (N) (4)	Hive (N) (6) Homemade (Adj.) (3) Proper way (C) (2) Sting (V) (3)	Buy a farm (FWC)	Swing (N)	Wax (N)	Polish your nails (C)
E13	Instruments (N) (6) Play (V) (4)	Musicians (N) (4) Practice (V) (3)	Warming up (PV) (2) Band (N) (6) Sure (Adv.) (2) Realized (V) (2)	Drums (N)	Eat a carrot (FWC)	Lute (N)	Briefcase (N)
E14	Easy (Adj.) (2) City (N) (2)	Surprises (N) (2) Time (N) (4)	Petting a bunny (C) (2) Country air (Comp.) (2) Hutch (N) (2) Too fast (LA) (4)	Tractor (N)	Plastic (N)	Landscape (N)	Fencing (N)
E15	Cars (N) (5) Boat (N) (26)	Model (N) (2) Competition (N) (2)	Built (V) (3) Contest (N) (3) Over there (LA) (3) Floating (V) (3)	Swim (V)	Phone (N)	Fetch (V)	Badger (N)

Appendices

Episode	Target words			Distractors			
	Known		New	Known		New	
	Frequent			Related	Not related	Related	Not related
E16	Toy boat (Comp.) (2) I'm coming (FWC) (2)	Home (N) (4) Water (N) (12)	Beavers (N) (13) Dam (N) (20) Creek (N) (3) Pond (N) (12)	Ocean (N)	Pen (N)	Mast (N)	Pansy (N)
E17	Skates (N) (8) See (V) (5)	Dog (N) (7) Thing (N) (3)	How to stop (LA) (3) Thought (V) (5) Wheel shoes (Comp.) (3) Trust (V) (2)	Pasta (N)	Bottle (N)	Laces (N)	Resign (V)
E18	Lake (N) (7) Amazing (Adj.) (3)	Bowl (N) (3) Frogs (N) (3)	Tadpoles (N) (24) Caterpillar (N) (2) Lagoon (N) (3) Grow (V) (4)	Sea (N)	Eat Breakfast (C)	Creel (N)	Hinder (V)
E19	End of the game (LA) (2) New record (C) (2)	Number (N) (5) Miniature golf (Comp.) (3)	Great stuff (C) (2) Score (N) (9) Hole (N) (15) Windmill (Comp.) (3)	Sport (N)	Hat (N)	Stance (N)	Peacock (N)
E20	Day (N) (6) House (N) (2)	Winter (N) (11) Freeze (V) (2)	Having fun (C) (2) Bowling (N) (2) Forget (V) (3) Warm (Adj.) (4)	Rain (N)	Love (V)	Mitten (N)	Rust (N)

(N) – Noun

(Adj.) – Adjective

(C) – Collocation

(PV) – Phrasal Verb

(SAF) – Speech Act Formulae

(V) - Verb

(Adv.) - Adverb

(Comp-) - Compound

(LA) – Linking Adverbial

(FWC) – Free Word Combination

APENDIX H – VOCABULARY PROFILE OF THE CURIOUS GEORGE EPISODES / RUNNING TIME OF EPISODES

Table H.1 – Vocabulary Profile of the *Curious George* episodes

Episode nº / Title		WF						Cumulative tokens (%) 1k + 2k	AWL						Off-list			
		WF		Types		Tokens			WF		Types		Tokens		Types		Tokens	
		Raw	%	Raw	%	Raw	%		Raw	%	Raw	%	Raw	%	Raw	%	Raw	%
E1 - A Zoo night	1k	166	89,7	206	85,1	549	91,1	96,5	3	1,6	3	1,24	4	0,7	14	5,79	17	2,82
	2k	16	8,6	19	7,85	33	5,5											
E2 - Curious George Rocket Ride	1k	213	86,6	252	80	709	88,3	91,7	9	3,7	9	2,86	12	1,5	29	9,21	55	6,85
	2k	24	9,8	25	7,94	27	3,4											
E3 - Curious George The Architect	1k	162	85,7	193	80,4	523	89	93,4	8	4,2	9	3,75	15	2,6	19	7,92	24	4,12
	2k	19	10,1	19	7,92	25	4,3											
E4 - Animal Magnetism	1k	186	88,6	230	83,3	684	88,5	93,4	2	1,0	2	0,72	2	0,3	20	7,25	49	6,34
	2k	22	10,5	24	8,7	38	4,9											
E5 - Curious George Station Master	1k	193	88,5	241	84,6	760	93,1	97,1	4	1,8	4	1,4	4	0,5	17	5,96	19	2,33
	2k	21	9,6	23	8,07	33	4											
E6 - Curious George Muddy Monkey	1k	174	87	213	81,6	676	88,5	97,1	1	0,5	1	0,38	1	0,1	16	6,13	21	2,75
	2k	25	12,5	31	11,9	66	8,6											
E7 - Curious George makes a Stand	1k	180	86,1	229	82,7	737	91,6	96,2	5	2,4	5	1,81	5	0,6	16	5,78	26	3,23
	2k	24	11,5	27	9,75	37	4,6											
E8 - Curious George Door Monkey	1k	171	80,3	212	76	602	85,4	93,8	6	2,8	6	2,15	14	2,0	21	7,53	30	4,26
	2k	36	16,9	40	14,3	59	8,4											
E9 - Curious George Ski Monkey	1k	172	88,7	216	83	656	90,1	95	-	-	-	-	-	-	21	8,08	37	5,08
	2k	22	11,3	23	8,85	35	4,8											
E10 - Curious George Home for Pigeons	1k	145	88,4	181	84,9	481	89,7	96	2	1,2	2	0,94	3	0,6	12	5,63	18	3,36
	2k	17	10,4	18	8,45	34	6,3											
E11 – Curious George The Grocer	1k	195	86,3	248	78,9	800	86,8	92,4	3	1,3	3	0,96	9	1,0	32	10,2	61	6,62
	2k	28	12,4	31	9,87	52	5,6											
E12 – Curious George Bee is for Bear	1k	192	82,8	239	75,9	825	87,7	93	6	2,6	6	1,90	8	0,9	34	10,7	59	6,27
	2k	34	14,7	36	11,4	49	5,2											
E13 – The All Animal Recycled band	1k	147	86,5	180	82,2	429	89,2	96,5	2	1,2	2	0,91	2	0,4	14	6,39	15	3,12
	2k	21	12,4	23	10,5	35	7,3											
E14 – Curious George Bunny Hunt	1k	180	85,3	214	82	610	89,4	95,4	4	1,9	4	1,53	4	0,6	16	6,13	27	3,96
	2k	27	12,8	27	10,3	41	6,0											
E15 – Curious George Buoy Wonder	1k	208	85,2	262	80,1	732	90,3	95,4	6	2,5	6	1,83	7	0,9	24	7,34	31	3,82
	2k	30	12,3	35	10,7	41	5,1											

Appendices

Episode n° / Title								Cumulative tokens (%) 1k + 2k	AWL						Off-list			
		WF		Types		Tokens			WF		Types		Tokens		Types		Tokens	
		Raw	%	Raw	%	Raw	%		Raw	%	Raw	%	Raw	%	Raw	%	Raw	%
E16 – Curious George and the Dam Builders	1k	168	87,5	214	85,6	569	87,8	92	2	1,0	2	0,8	2	0,3	12	4,8	50	7,72
	2k	22	11,5	22	8,8	27	4,2											
E17 – Curious George Roller Monkey	1k	169	84,9	206	79,8	652	88,5	95,4	3	1,5	3	1,16	3	0,4	17	6,59	31	4,21
	2k	27	13,6	32	12,4	51	6,9											
E18 – Curious George discovers de Poles	1k	172	81,1	214	74,1	609	83,7	90,2	5	2,4	5	1,73	5	0,7	31	10,7	67	9,2
	2k	35	16,5	39	13,5	47	6,5											
E19 – Curious George Low High Score	1k	175	89,3	219	86,6	729	88,9	94,6	1	0,5	1	0,4	1	0,1	12	4,74	43	5,24
	2k	20	10,2	21	8,3	47	5,7											
E20 – Curious George vs. Winter	1k	176	83,0	223	77,9	627	86,4	95,6	2	0,9	2	0,7	2	0,3	21	7,34	30	4,13
	2k	34	16,0	40	13,9	67	9,2											

Table H.2 – Running time of the *Curious George* episodes

Episode	E1	E2	E3	E4	E5	E6	E7	E8	E9	E10
Length	12'47"	12'20"	12'22"	12'22"	12'19"	12'17"	12'17"	12'15"	12'17"	12'49"
Episode	E11	E12	E13	E14	E15	E16	E17	E18	E19	E20
Length	12'16"	12'48"	12'18"	12'47"	12'21"	12'25"	12'17"	12'46"	12'16"	12'16"

APENDIX I - TRANSCRIPTS FROM CURIOUS GEORGE (E9 AND 19)

E9 – Curious George Rocket Ride

Narrator: George liked visiting Professor Wiseman because she did so many different things.

Narrator: This time she was helping out on a rocket launch and the Man with the Yellow Hat was going to ride the rocket.

Professor Wiseman: So, you ready to be the first average, untrained person shot into space?

Man with the Yellow Hat: I sure am.

Curious George: (chattering)

Professor Wiseman: There is only room for one, George, but I have a special seat for you right beside me so you can watch.

Curious George: (groans)

Professor Wiseman: This is Dr. Alvin Einstein and Professor Anthony Pizza.

Man with the Yellow Hat: Wow, are you related to the famous...

Alvin Einstein: No, I am not.

Anthony Pizza: Me neither.

Professor Wiseman: The International Space Station's food supply has run out.

Woman: We found a peanut. It was in the cushion of my chair.

Professor Wiseman: This man with the yellow hat will bring your food supply today.

Man: So, then it's okay to eat the peanut now?

Professor Wiseman: Yes, you may eat the peanut.

Alvin Einstein: So, this is your rocket designed by Professor Wiseman.

Anthony Pizza: It is up to you to launch the food payload at exactly the proper moment. After it detaches from the rocket, it will connect to the space station. Can we count on you?

Man with the Yellow Hat: Sure can. (becomes an astronaut)

Alvin Einstein: Professor Pizza and I are loading extra experiments you will deliver with the food.

Narrator: George wanted to see what these experiments were.

Man with the Yellow Hat: George, careful. This is an expensive, hi-tech...

Man with the Yellow Hat: raccoon?

Professor Wiseman: They will live at the space station so we can study how they adapt to life in space.

Narrator: George could see into all the containers.....except this one.

Appendices

Narrator: What was hidden in there?

Anthony Pizza: Show him the most important part.

Alvin Einstein: These keypads launch the payload. You hit these two keys on each pad at the same time.

Professor Wiseman: On all four at the same time?

Alvin Einstein: Yes, at the exact moment the rocket passes the space station.

Man with the Yellow Hat: I cannot. I only have two hands.

Alvin Einstein: The keypads were Pizza's idea.

Anthony Pizza: The raccoon was my idea. The four keypads was yours.

Alvin Einstein: Oh, here we go...

Anthony Pizza: Check your memos.

Alvin Einstein: I specifically remember when you came up with four keypads...

Curious George: (gasps) (chattering)

Professor Wiseman: Oh, we have to scrub the mission.

Man with the Yellow Hat: Or find an astronaut with four hands.

Curious George: (chatters)

Man with the Yellow Hat: Hey...

Curious George: Ha-ha!

Professor Wiseman: Don't be sad. You could've gone if you had four hands.

Curious George: (chattering)

Man with the Yellow Hat: Be a good little monkey... in space.

Professor Wiseman: Ten, nine... eight, seven, six... five, four... three, two, one. Ignition. George, you've achieved orbit.

Curious George: Ooh.

Professor Wiseman: To your left is Earth.

Curious George: Huh? Ooh...

Professor Wiseman: To your right, you're passing the space station.

Curious George: (nervous chattering)

Professor Wiseman: Do not worry. You are circling the earth, so you will pass it again when you have done one complete orbit.

Professor Wiseman: Did I mention the lower gravity in space plus the motion of the rocket make you weightless?

Alvin Einstein: He is about to pass out of communication range.

Professor Wiseman: George, we are going to be out of contact for a few minutes.

Appendices

Man with the Yellow Hat: Do not be scared, George. See you soon.

Curious George: Phew!

Narrator: There was nothing to do now but wait to push those buttons and wonder what was in that box. He'd seen how heavy these boxes were, so he gave a mighty monkey tug and found out he wasn't the only thing that weighed less in space. Toys! It was full of toys!

Man with the Yellow Hat: Poor George. I hope he is not too scared being alone and out of contact.

Narrator: George had never seen a top spin so long before. Space was a great place to play.

Curious George: Whoo!

Narrator: George wondered whether these toys would be more fun if he had someone to play with.

Man with the Yellow Hat: He is coming back into communication range.

Professor Wiseman: What's happened? Where is George?

Professor Wiseman: George, you must put the experiments away.

Alvin Einstein: Hurry! It's almost time to send them to the space station.

Anthony Pizza: At least he did not free the ants.

Curious George: Oh?

Alvin Einstein: If ants get into the back, they might eat the astronauts' food.

Curious George: Huh?

Alvin Einstein: In 28 seconds, he'll be in position to launch the payload.

Curious George: (nervous chattering)

Professor Wiseman: You have 20 seconds to get everything ready to launch.

Curious George: Ooh... Aah... Oh... Oh... Aha... (gasps)

Narrator: But George wasn't ready in time.

Anthony Pizza: He's got to do it next time around.

Professor Wiseman: George, you'll only get one more chance. Then, succeed or fail, we have to bring you home.

Both: We are really hungry!

Narrator: Hard to believe, but cleaning his room had prepared George for an important mission in space.

Curious George: (grunts)

Curious George: (chattering quietly) Ah! Ya!

Curious George: (chatters)

Man with the Yellow Hat: He looks ready.

Appendices

Professor Wiseman: George, are you ready to launch?

Curious George: (chatters)

Professor Wiseman: Excellent! I'll tell you when.

Alvin Einstein: Uh, maybe we shouldn't have sent a monkey.

Professor Wiseman: Now, George, now!

Both: He did it!

Man with the Yellow Hat: Great job, George!

Man: Thank you, George.

Curious George: (chatters)

Professor Wiseman: You're in position to return home. Pull the lever to fire propulsion rockets.

Curious George: (groans)

Professor Wiseman: George, pull the lever now, or you'll be out of position.

Curious George: (chatters)

Man with the Yellow Hat: What happens if he pulls the lever late?

Professor Wiseman: He could land anywhere top of a mountain, the North Pole.

Man with the Yellow Hat: George, pull the lever now!

Curious George: Huh? Oh! (chatters)

Narrator: So, George was a hero.

Curious George: (happy chatter)

Narrator: And he proved, just because you're a small monkey (camera shutter clicks) doesn't mean you cannot take care of everything...

Curious George: Aw...

Narrator: down to the tiniest details.

E19 – Curious George Low High Score

Narrator: It was a perfect day to walk down N. Avenue and wonder "if building numbers get higher and higher... do they ever stop?" And that is exactly what George was doing...

Narrator: ...until he got curious about a sound.

Betsy: Hi, George. Want to shoot baskets?

Steve: 23!

Betsy: When is it going to be my turn, Steve?

Steve: When I miss. If I miss. Now, where was I? Oh, yeah, 23. I'm going for a new record. 24 baskets in a row.

Steve: Oh! Now it's your turn, Betsy.

Betsy: Here, you take a turn, George.

Betsy: Whoa, maybe you can beat Steve's record.

Steve: That's one. You need 23 more to beat me.

Narrator: George was sure he could get 24... 34, maybe... 104!

Betsy: Oh, no! Wait! Come back! That's our ball!

Steve: Ha. Now what?

Betsy: How about miniature golf?

Betsy: You have never played? It is great! There is a castle and a windmill...

Steve: I will show you how you play. I'm good at this. To figure your score, add up how many times you hit the ball before it goes in the hole.

Narrator: George had one goal: to do better at golf than he had at basketball.

Narrator: By the end of the first hole, George knew he was great at this game.

Betsy: A little softer.

Steve: On hole number one, I got a five, Betsy got eight, and George 27.

Narrator: George had not come close to Steve's 23 baskets, but he already had the highest golf score!

Steve: On this hole, Betsy got four and I got three. If you get it in now, you will have nine.

Narrator: Nine? George was sure he could get that score up higher. But it took effort to miss from so close.

Betsy: That is ten. Keep trying. Eleven, twelve, thirteen, fourteen... Hit the windmill again. That is 20.

Narrator: Getting a high score at golf was easy. George was born to golf! At the end of the game, Steve totaled up the scores.

Steve: I had 35. That is my new record for the best score yet!

Appendices

Steve: Betsy's adds up to 58.

Narrator: And George 257.

Betsy: George, a high score is not good.

Steve: You want to get the ball in the hole with the least number of tries.

Betsy: In golf, the lowest score is best. I thought you knew.

Narrator: George couldn't imagine how the lowest number could be the best.

Narrator: That meant if George did not play at all, he would score zero, and be the greatest golfer in history.

Steve: Maybe monkeys just are not meant to play golf.

Narrator: Numbers can confuse a monkey. How can a big number be worth less than a small number?

Narrator: Maybe monkeys weren't meant to play golf or basketball, or to set records.

Betsy: That was just like miniature golf! I think you would have done better at miniature golf if you had known the rules.

Narrator: Sometimes girls and monkeys think exactly alike.

Betsy: We could build our very own golf course right here! I know where to get some great stuff to build a golf course with.

Man with the Yellow Hat: Sure, you can play with that stuff.

Man with the Yellow Hat: But don't ride the unicycle without a helmet!

Betsy: Wow. Great stuff. Let's get to work!

Steve: Betsy, Aunt Margaret says to be home for dinner in an hour. Holy cow!

Betsy: We made our own mini-golf course. You want to play?

Betsy: In our game, each hole uses a special club.

Steve: You will never get it through there. You kept hitting the windmill yesterday.

Betsy: A hole in one!

Steve: Great! If George can do that, I'm going to break all my golf records on this easy course.

Steve: George, thanks, but I really don't need help on this.

Betsy: Are you keeping track of your score on this easy course, Steve?

Steve: Oh, yes. Okay, on the first hole, George got a one lucky. Betsy got a six, and I got... a six.

Betsy: The club for this hole, please. Here we go. Whoa. Oh!

Betsy: Oh... Yes! There it is!

Steve: A hole in one!

Betsy: My first ever! Looks like I'm going for the record, too.

Appendices

Steve: Okay, here comes my hole in one. Aah!

Curious George: Uh-huh. Uh-huh.

Steve: A hole in two!

Betsy: We're all going for our best scores ever on this course.

Steve: They're our first scores ever on this course.

Betsy: Yeah, so? They're still the best.

Steve: What's the club for the last hole?

Betsy: No club. It is a blowhole. Get it? Whale? Blow hole? Okay... going around... and... the house... Oh, oh, oh, oh... Oh! Perfect!

Steve: At the end of the game, I have 40, Betsy has 25, and George... got a 21?!

Betsy: I told you, you would play better when you knew the rules.

Steve: I am sorry. I said monkeys should not play golf. You guys want to play real mini golf tomorrow?

Betsy: Thanks, but we like this better.

Steve: Oh. Can I play again with you tomorrow? Maybe George can help me improve my game.

APPENDIX J – ITEM TYPE AND IDEA UNITS OF *CURIOUS GEORGE* EPISODES

Table J.1 – Item type for idea units of *Curious George* episodes

Episode number	Main Ideas	Specific Details	Inferences
E1	8	4	3
E2	8	5	2
E3	8	6	1
E4	8	5	2
E5	8	5	2
E6	7	5	3
E7	7	6	2
E8	8	5	2
E9	7	7	1
E10	7	5	3
E11	6	7	2
E12	8	6	1
E13	7	7	1
E14	8	5	2
E15	9	4	2
E16	8	6	1
E17	8	5	2
E18	8	5	2
E19	7	6	2
E20	6	5	4
TOTAL	151	109	40
PERCENT	51,4	37,3	11,3%

APPENDIX K – DIFFICULTY AND DISCRIMINATION INDEX

Table K.1 – Difficulty Index of the multiple choice exercise

RELIABILITY – DIFFICULTY INDEX		
MC	L1S	L2S
E1.1a	0,17	0,5
E1.1b	0,29	0,16
E1.1c	0,46	0,5
E1.2a	0,98	0,9
E1.2b	0,98	0,73
E1.2c	0,96	0,83
E1.3a	0,91	0,77
E1.3b	0,94	0,87
E1.3c	0,94	0,9
E1.4a	0,83	0,56
E1.4b	0,87	0,83
E1.4c	0,97	0,73
E1.5a	0,97	0,87
E1.5b	0,46	0,3
E1.5c	0,01	0,43
E2.1a	0,9	0,97
E2.1b	0,17	0,45
E2.1c	0,27	0,42
E2.2a	0,97	0,97
E2.2b	0,84	0,73
E2.2c	0,87	0,76
E2.3a	0,98	0,89
E2.3b	0,90	0,73
E2.3c	0,97	0,83
E2.4a	0,56	0,49
E2.4b	0,56	0,44
E2.4c	1	0,93
E2.5a	0,97	0,93
E2.5b	0,94	0,89
E2.5c	0,97	0,89
E3.1a	1	0,97
E3.1b	0,93	1
E3.1c	0,93	1
E3.2a	1	0,97
E3.2b	1	0,97
E3.2c	1	0,97
E3.3a	0,73	0,37
E3.3b	0,93	0,93
E3.3c	0,80	0,44
E3.4a	0,83	0,62

E3.4b	0,83	1
E3.4c	0,66	0,62
E3.5a	1	0,97
E3.5b	0,52	0,41
E3.5c	0,52	0,51
E4.1a	0,84	0,9
E4.1b	0,84	0,87
E4.1c	1	0,97
E4.2a	0,71	0,3
E4.2b	1	0,92
E4.2c	0,71	0,33
E4.3a	0,43	0,23
E4.3b	0,97	0,87
E4.3c	0,46	0,36
E4.4a	0,98	0,9
E4.4b	0,98	0,83
E4.4c	1	0,9
E4.5a	0,97	0,92
E4.5b	1	1
E4.5c	0,97	0,92
E5.1a	0,83	0,87
E5.1b	0,83	0,87
E5.1c	1	1
E5.2a	0,86	0,8
E5.2b	0,90	0,9
E5.2c	0,97	0,9
E5.3a	0,86	0,7
E5.3b	0,80	0,67
E5.3c	0,93	0,97
E5.4a	0,97	0,93
E5.4b	0,97	0,83
E5.4c	0,93	0,69
E5.5a	0,80	0,2
E5.5b	0,86	0,73
E5.5c	0,93	0,4
E6.1a	1	0,97
E6.1b	0,84	0,77
E6.1c	0,84	0,73
E6.2a	0,87	0,87
E6.2b	0,83	0,57
E6.2c	0,94	0,7
E6.3a	0,02	0,77
E6.3b	0,49	0,77
E6.3c	0,97	0,97
E6.4a	1	0,97
E6.4b	0,87	0,83

Appendices

E6.4c	0,87	0,73
E6.5a	1	0,87
E6.5b	0,90	0,67
E6.5c	0,87	0,8
E7.1a	0,78	0,5
E7.1b	0,81	0,52
E7.1c	0,97	0,97
E7.2a	1	0,67
E7.2b	1	0,63
E7.2c	1	0,97
E7.3a	0,97	1
E7.3b	0,90	0,97
E7.3c	0,97	0,97
E7.4a	0,65	0,63
E7.4b	1	0,9
E7.4c	0,65	0,53
E7.5a	0,9	0,87
E7.5b	0,27	0,3
E7.5c	0,17	0,43
E8.1a	0,71	0,35
E8.1b	0,62	0,18
E8.1c	0,91	0,81
E8.2a	1	1
E8.2b	0,46	0,53
E8.2c	0,46	0,53
E8.3a	0,97	1
E8.3b	0,81	0,97
E8.3c	0,87	0,97
E8.4a	0,68	0,37
E8.4b	0,81	0,72
E8.4c	0,49	0,11
E8.5a	0,97	0,93
E8.5b	0,94	0,86
E8.5c	0,97	0,9
E9.1a	0,8	0,92
E9.1b	0,84	0,76
E9.1c	0,64	0,7
E9.2a	0,90	0,76
E9.2b	0,87	0,6
E9.2c	0,97	0,83
E9.3a	0,48	0,43
E9.3b	0,97	0,46
E9.3c	0,51	0,6
E9.4a	1	0,97
E9.4b	0,97	0,92
E9.4c	0,97	0,92

E9.5a	0,97	0,83
E9.5b	0,97	0,9
E9.5c	1	0,9
E10.1a	0,63	0,72
E10.1b	0,63	0,79
E10.1c	1	0,93
E10.2a	0,93	0,97
E10.2b	0,93	0,97
E10.2c	1	0,92
E10.3a	0,33	0,3
E10.3b	0,63	0,76
E10.3c	0,03	0,55
E10.4a	0,57	0,58
E10.4b	0,53	0,55
E10.4c	0,97	0,97
E10.5a	0,9	0,75
E10.5b	0,97	0,92
E10.5c	0,92	0,83
E11.1a	0,97	0,83
E11.1b	0,8	0,52
E11.1c	0,83	0,69
E11.2a	0,92	0,86
E11.2b	0,97	0,97
E11.2c	0,97	0,86
E11.3a	0,93	0,97
E11.3b	0,43	0,6
E11.3c	0,5	0,59
E11.4a	0,73	0,28
E11.4b	0,93	0,76
E11.4c	0,67	0,45
E11.5a	0,27	0,69
E11.5b	0,5	0,18
E11.5c	0,77	0,63
E12.1a	0,93	0,97
E12.1b	0,83	0,79
E12.1c	0,9	0,79
E12.2a	0,77	0,48
E12.2b	0,8	0,48
E12.2c	0,97	1
E12.3a	1	0,97
E12.3b	0,97	0,97
E12.3c	0,97	0,97
E12.4a	0,83	0,72
E12.4b	0,92	0,97
E12.4c	0,9	0,76
E12.5a	0,83	0,69

Appendices

E12.5b	0,92	0,9
E12.5c	0,77	0,63
E13.1a	0,8	0,57
E13.1b	0,8	0,53
E13.1c	1	0,97
E13.2a	0,9	0,73
E13.2b	0,97	1
E13.2c	0,87	0,73
E13.3a	0,87	0,76
E13.3b	0,9	0,94
E13.3c	0,97	0,93
E13.4a	0,93	0,57
E13.4b	0,93	0,9
E13.4c	0,87	0,47
E13.5a	0,87	0,73
E13.5b	0,38	0,4
E13.5c	0,5	0,6
E14.1a	0,97	0,73
E14.1b	0,93	0,9
E14.1c	0,9	0,6
E14.2a	0,97	0,83
E14.2b	1	0,97
E14.2c	0,97	0,86
E14.3a	0,97	0,97
E14.3b	1	0,73
E14.3c	0,97	0,77
E14.4a	0,97	0,97
E14.4b	0,87	0,69
E14.4c	0,84	0,69
E14.5a	0,90	0,73
E14.5b	1	0,93
E14.5c	0,90	0,86
E15.1a	0,77	0,46
E15.1b	0,80	0,53
E15.1c	0,97	0,93
E15.2a	0,74	0,66
E15.2b	0,66	0,3
E15.2c	0,90	0,52
E15.3a	0,58	0,33
E15.3b	0,51	0,4
E15.3c	0,93	0,97
E15.4a	0,54	0,53
E15.4b	0,87	0,02
E15.4c	0,41	0,7
E15.5a	0,93	0,83
E15.5b	0,80	0,6

E15.5c	0,87	0,77
E16.1a	0,97	0,97
E16.1b	0,93	0,93
E16.1c	0,97	0,97
E16.2a	0,67	0,33
E16.2b	0,93	0,97
E16.2c	0,61	0,3
E16.3a	0,97	0,7
E16.3b	0,97	0,8
E16.3c	1	0,9
E16.4a	0,84	0,6
E16.4b	0,74	0,43
E16.4c	0,90	0,83
E16.5a	0,25	0,47
E16.5b	0,35	0,36
E16.5c	0,90	0,9
E17.1a	1	0,97
E17.1b	0,97	0,97
E17.1c	0,97	0,97
E17.2a	0,97	0,8
E17.2b	0,97	0,97
E17.2c	0,93	0,74
E17.3a	0,84	0,91
E17.3b	0,89	0,97
E17.3c	0,97	0,91
E17.4a	0,97	0,68
E17.4b	0,93	0,69
E17.4c	0,97	0,97
E17.5a	0,80	0,79
E17.5b	0,87	0,73
E17.5c	0,71	0,52
E18.1a	0,87	0,72
E18.1b	0,87	0,77
E18.1c	1	0,89
E18.2a	0,9	0,79
E18.2b	1	0,97
E18.2c	0,9	0,79
E18.3a	0,83	0,49
E18.3b	0,7	0,22
E18.3c	0,87	0,73
E18.4a	0,92	1
E18.4b	0,92	0,89
E18.4c	0,87	0,89
E18.5a	0,92	0,93
E18.5b	0,97	0,89
E18.5c	0,9	0,83

Appendices

E19.1a	1	0,93
E19.1b	0,68	0,63
E19.1c	0,71	0,73
E19.2a	0,68	0,5
E19.2b	0,71	0,53
E19.2c	1	0,97
E19.3a	0,62	0,73
E19.3b	0,49	0,7
E19.3c	0,1	0,4
E19.4a	0,81	0,53
E19.4b	0,74	0,87
E19.4c	0,55	0,4
E19.5a	0,94	0,7
E19.5b	0,81	0,53
E19.5c	0,87	0,83
E20.1a	0,77	0,57
E20.1b	0,65	0,57
E20.1c	0,87	1
E20.2a	0,39	0,4
E20.2b	0,8	0,7
E20.2c	0,59	0,7
E20.3a	0,68	0,8
E20.3b	0,8	0,9
E20.3c	0,49	0,7
E20.4a	0,36	0,5
E20.4b	0,39	0,63
E20.4c	0,74	0,13
E20.5a	0,65	0,57
E20.5b	0,59	0,4
E20.5c	0,94	0,83

Table K.2 – Percentages of Difficulty Index of the multiple choice exercise

Difficulty	L1S	L2S
Very Easy	73,3%	53,6%
Easy	13,3%	23,3%
Medium	6,3%	10,3%
Difficult	4,6%	9,7%
Very difficult	2,3%	3%

Table K.3 – Difficulty Index of the True / False task

RELIABILITY – DIFFICULTY INDEX		
T/F	L1S	L2S
E1a	0,76	0,89
E1b	0,38	0,49
E1c	0,77	0,42
E1d	0,53	0,47
E1e	0,89	0,87
E2a	0,51	0,45
E2b	0,79	0,84
E2c	0,40	0,59
E2d	0,95	0,77
E2e	0,53	0,54
E3a	0,95	0,93
E3b	0,79	0,63
E3c	0,86	0,89
E3d	1	0,95
E3e	0,91	0,88
E4a	0,85	0,89
E4b	0,79	0,6
E4c	0,96	0,89
E4d	0,74	0,6
E4e	0,81	0,71
E5a	0,98	0,91
E5b	0,91	0,89
E5c	0,75	0,6
E5d	0,66	0,67
E5e	0,75	0,49
E6a	0,98	0,93
E6b	0,62	0,17
E6c	1	1
E6d	0,74	0,46
E6e	0,83	0,46
E7a	0,70	0,76
E7b	0,94	0,91
E7c	0,87	0,84
E7d	0,91	0,89
E7e	0,96	0,82
E8a	0,89	0,79
E8b	0,74	0,57
E8c	0,89	0,59
E8d	0,74	0,59
E8e	0,94	0,98
E9a	0,98	0,96

E9b	0,74	0,56
E9c	0,59	0,51
E9d	0,65	0,64
E9e	1	0,93
E10a	1	0,97
E10b	0,89	0,91
E10c	0,84	0,91
E10d	0,2	0,28
E10e	0,42	0,53
E11a	0,87	0,82
E11b	0,89	0,84
E11c	0,84	0,95
E11d	0,89	0,77
E11e	0,84	0,93
E12a	0,8	0,75
E12b	0,87	0,89
E12c	0,71	0,61
E12d	0,89	0,91
E12e	0,98	0,80
E13a	0,74	0,73
E13b	0,67	0,58
E13c	0,74	0,67
E13d	0,85	0,93
E13e	0,91	0,76
E14a	0,8	0,80
E14b	0,96	0,93
E14c	0,96	0,8
E14d	0,96	0,98
E14e	0,96	0,96
E15a	0,89	0,84
E15b	0,59	0,42
E15c	0,85	0,71
E15d	0,59	0,38
E15e	0,89	0,87
E16a	0,87	0,69
E16b	0,93	0,67
E16c	0,89	0,84
E16d	0,93	0,96
E16e	0,93	0,87
E17a	0,61	0,82
E17b	0,83	0,91
E17c	0,89	0,66
E17d	0,74	0,78
E17e	0,41	0,5
E18a	0,98	0,98
E18b	0,76	0,45

Appendices

E18c	0,89	0,84
E18d	0,87	0,91
E18e	0,85	0,57
E19a	0,79	0,82
E19b	0,72	0,78
E19c	0,94	0,82
E19d	0,85	0,82
E19e	0,94	0,89
E20a	0,96	0,96
E20b	0,96	0,84
E20c	0,87	0,87
E20d	0,45	0,49
E20e	0,60	0,51

Table K.4 – Difficulty Index percentages of the True / False task

Difficulty	L1S	L2S
Very Easy	67%	61%
Easy	25%	23%
Medium	4%	12%
Difficult	3%	1%
Very difficult	1%	3%

Table K.5 – Discrimination Index of the True / False task

DISCRIMINATION INDEX T / F		
	L1S	L2S
1a	0.06	0.07
1b	0.13	0.2
1c	0.06	0.33
1d	0.45	-0.07
1e	0.06	0.27
2a	0	0.07
2b	0.26	0.2
2c	0.26	-0.13
2d	0.06	0.13
2e	0.38	0.6
3a	0	0.27
3b	0.38	0.07
3c	0	0.2
3d	-0.06	0.27
3e	-0.06	0.27
4a	0	0.07
4b	0.06	0.47
4c	0.06	0.33
4d	0.26	0.33
4e	0.26	0.27
5a	0.06	0.2
5b	0.19	-0.07
5c	0.38	0.33
5d	0.51	0
5e	0.38	-0.07
6a	0.06	0.2
6b	0.19	0.33
6c	0	0
6d	0.77	0.13
6e	0.26	0.53
7a	0.77	0.13
7b	0.13	0.13
7c	0.13	0.27
7d	0	0.27
7e	0.13	0.33
8a	0.19	0.27
8b	0.06	0.4
8c	0	0.07
8d	0.77	0.27
8e	0.06	0
9a	0.13	-0.13
9b	0.38	0.33
9c	0.51	-0.2
9d	0.51	0.33
9e	0.06	0.2

10a	0	-0.13
10b	0	0.3
10c	-0.06	0.2
10d	0.13	0.13
10e	0	0.27
11a	0.38	0
11b	0.19	0.13
11c	0	0.07
11d	0.38	0.33
11e	-0.06	0
12a	0.38	0
12b	0.38	0.2
12c	0.38	-0.07
12d	0.26	0.33
12e	0.13	0.13
13a	-0.13	0.13
13b	0.38	0.4
13c	0.77	0.4
13d	-0.13	0.07
13e	0.26	0.27
14a	0.77	0.4
14b	0.06	0.2
14c	0.13	0.53
14d	0.13	0.07
14e	0.13	0
15a	0.38	0.07
15b	0.51	0.4
15c	0.32	0.13
15d	0.51	0.13
15e	0.26	0
16a	0.06	0.13
16b	0.13	0.13
16c	0.06	0
16d	0.13	0.13
16e	0.13	0.27
17a	-0.13	0.27
17b	0.26	0.13
17c	0.06	0.27
17d	0.51	0.33
17e	0.32	0.27
18a	0.13	0.13
18b	0.06	0.33
18c	0.19	0.2
18d	0.26	0.33
18e	0.19	0.2
19a	0.13	0.13
19b	0.38	0.2
19c	0.13	0.13
19d	0.32	0.4
19e	0.06	0.2

Appendices

20a	0	0.07
20b	0.13	0
20c	0.26	0.33
20d	0.19	0
20e	0.57	0.2

Table K.6 – Discrimination Index percentages of the True / False task

	L1S	L2S
Very Good	13%	9%
Good	16%	15%
Regular	11%	27%
Poor	60%	49%

Table K.7 – Discrimination Index of the multiple choice task

Discrimination index - MC		
	L1S	L2S
E1.1a	0.57	0.06
E1.1b	0.57	0.13
E1.1c	0	0.06
E1.2a	0.06	0.2
E1.2b	0.06	0.47
E1.2c	0.06	0.27
E1.3a	0	0.2
E1.3b	0	0.13
E1.3c	0	0.06
E1.4a	0.06	0.47
E1.4b	-0.13	0.2
E1.4c	0.06	0.27
E1.5a	0.06	0.13
E1.5b	0.38	0.27
E1.5c	0.32	0.13
E2.1a	0.06	0.06
E2.1b	0.38	0
E2.1c	0.45	0.06
E2.2a	0.06	0.06
E2.2b	0.32	0.13
E2.2c	0.26	0.06
E2.3a	0.06	0.13
E2.3b	0.13	0.33
E2.3c	0.06	0.2
E2.4a	0.64	0.06
E2.4b	0.64	0.06
E2.4c	0	0.13
E2.5a	0.06	0.13
E2.5b	0.06	0.2
E2.5c	0	0.06
E3.1a	-0.13	0.13
E3.1b	0.19	0.13
E3.1c	0.19	0.13
E3.2a	0.13	0.2
E3.2b	-0.13	0.13
E3.2c	-0.13	0.2
E3.3a	0.26	0.53
E3.3b	-0.06	0.13
E3.3c	0.19	0.53
E3.4a	-0.06	0.4
E3.4b	0.13	0.13
E3.4c	0.19	0.4
E3.5a	-0.13	0.2
E3.5b	0.13	0.47
E3.5c	0.13	0.4

E4.1a	0	0.2
E4.1b	0.19	0.13
E4.1c	0	-0.06
E4.2a	0.26	0.4
E4.2b	0	0.06
E4.2c	0.26	0.47
E4.3a	0.70	0.67
E4.3b	0.06	0.27
E4.3c	0.64	0.4
E4.4a	0.06	0.2
E4.4b	0.06	0.33
E4.4c	0	0.2
E4.5a	0.06	0.06
E4.5b	0	0
E4.5c	0.06	0.06
E5.1a	0.19	0.06
E5.1b	0.19	0.06
E5.1c	0.13	0
E5.2a	0.38	0.27
E5.2b	0.32	0.13
E5.2c	0.19	0.13
E5.3a	0.38	0.4
E5.3b	0.51	0.47
E5.3c	0.26	0.06
E5.4a	0.19	0
E5.4b	0.19	0.2
E5.4c	0.26	0.2
E5.5a	0.32	0.13
E5.5b	0.19	-0.06
E5.5c	0.26	0.2
E6.1a	0	0.06
E6.1b	0.32	0.13
E6.1c	0.32	0.2
E6.2a	0.26	0.13
E6.2b	0.38	0.6
E6.2c	0.13	0.53
E6.3a	0.57	0.27
E6.3b	0.51	0.2
E6.3c	0.06	0.06
E6.4a	0	0.06
E6.4b	0.26	0.13
E6.4c	0.26	0.2
E6.5a	0	0.2
E6.5b	0.13	0.4
E6.5c	0.13	0.2
E7.1a	0.13	0.13
E7.1b	0.06	0.2
E7.1c	0.06	-0.06
E7.2a	0	0.06
E7.2b	0	0.06

Appendices

E7.2c	0	0
E7.3a	0.06	0
E7.3b	0.06	0.06
E7.3c	0.06	0.06
E7.4a	0.26	0.06
E7.4b	0	0.13
E7.4c	0.26	0.2
E7.5a	0.06	0.13
E7.5b	-0.26	0.33
E7.5c	0.32	0.2
E8.1a	0.51	0.4
E8.1b	0.64	0.27
E8.1c	0.13	-0.2
E8.2a	0	-0.06
E8.2b	0.38	0.13
E8.2c	0.38	0.13
E8.3a	0.06	-0.06
E8.3b	0.32	-0.2
E8.3c	0.26	0.13
E8.4a	0.26	0
E8.4b	0	0.06
E8.4c	0.26	0.13
E8.5a	0.06	0
E8.5b	0.13	0.06
E8.5c	0.06	-0.06
E9.1a	0.19	0.06
E9.1b	0.19	-0.06
E9.1c	0.38	0
E9.2a	0.19	0.06
E9.2b	0.26	0.2
E9.2c	0.06	0.13
E9.3a	0.83	0.06
E9.3b	0.06	0.27
E9.3c	0.83	-0.2
E9.4a	0	0.06
E9.4b	0.06	0.06
E9.4c	0.06	0.13
E9.5a	0.06	-0.06
E9.5b	0.06	-0.06
E9.5c	0	0
E10.1a	0.32	0.06
E10.1b	0.32	0.2
E10.1c	0.13	-0.06
E10.2a	0.06	0
E10.2b	0.06	0
E10.2c	0.13	-0.06
E10.3a	0.33	-0.4
E10.3b	-0.06	-0.06
E10.3c	-0.19	-0.4
E10.4a	0.06	0.06

E10.4b	0.19	0
E10.4c	0.32	-0.13
E10.5a	0.32	0.27
E10.5b	0.19	-0.06
E10.5c	0.26	0.27
E11.1a	0.06	0.2
E11.1b	0.38	0.47
E11.1c	0.32	0.27
E11.2a	0.13	0.2
E11.2b	0.06	0
E11.2c	0.06	0.2
E11.3a	0.13	0.06
E11.3b	0.19	0
E11.3c	0.06	-0.06
E11.4a	0.19	0.06
E11.4b	0.13	0.2
E11.4c	0.32	0.27
E11.5a	0.57	0.13
E11.5b	0.32	0.13
E11.5c	0.26	0
E12.1a	0.13	0
E12.1b	0.19	-0.2
E12.1c	0.06	0.13
E12.2a	0.26	0.13
E12.2b	0.19	0.13
E12.2c	0.06	0
E12.3a	0	0.06
E12.3b	0.06	0
E12.3c	0.06	0.06
E12.4a	0.06	0.33
E12.4b	0.13	0.06
E12.4c	-0.06	0.27
E12.5a	0.32	0.27
E12.5b	0.13	0.06
E12.5c	0.45	0.33
E13.1a	0.26	0.4
E13.1b	0.32	0.4
E13.1c	-0.06	0.06
E13.2a	0.13	0.2
E13.2b	0	0
E13.2c	0.19	0.2
E13.3a	0.13	0.2
E13.3b	0.06	0.13
E13.3c	0	0.06
E13.4a	0	0.2
E13.4b	0.06	0.13
E13.4c	0.13	0.33
E13.5a	0.13	0
E13.5b	0.51	0.27
E13.5c	0.70	0.27

Appendices

E14.1a	0	0.27
E14.1b	0.06	0.2
E14.1c	0.13	0.47
E14.2a	0	0.13
E14.2b	-0.06	0.06
E14.2c	0	0.06
E14.3a	0	0.06
E14.3b	-0.06	0.27
E14.3c	0	0.33
E14.4a	0	0.06
E14.4b	0.13	0.2
E14.4c	0.19	0.27
E14.5a	0.13	0.33
E14.5b	-0.06	0.13
E14.5c	0.13	0.2
E15.1a	0.45	0.27
E15.1b	0.38	0.13
E15.1c	0.13	0.13
E15.2a	0.38	0.2
E15.2b	0.57	0.47
E15.2c	0.26	0.27
E15.3a	0.06	0.06
E15.3b	0.19	0.13
E15.3c	0.19	0.06
E15.4a	0.13	-0.13
E15.4b	0.19	0.06
E15.4c	0.26	-0.06
E15.5a	0.19	0.2
E15.5b	0.45	0.27
E15.5c	0.32	0.06
E16.1a	0	0.06
E16.1b	0.06	0.06
E16.1c	0	0
E16.2a	0.19	0.2
E16.2b	0.06	0
E16.2c	0.32	0.2
E16.3a	0	0.53
E16.3b	0	0.33
E16.3c	-0.06	0.2
E16.4a	0.32	0.4
E16.4b	0.38	0.4
E16.4c	0.06	0.06
E16.5a	0.57	0.47
E16.5b	0.38	0.06
E16.5c	0.13	0.06
E17.1a	0.06	0.13
E17.1b	0.13	0.13
E17.1c	0.13	0.06
E17.2a	0.13	0.2
E17.2b	0.13	0.13

E17.2c	0.19	0.27
E17.3a	0.26	0.2
E17.3b	0.19	0.06
E17.3c	0.13	0.2
E17.4a	0.13	0.6
E17.4b	0.19	0.6
E17.4c	0.13	0.06
E17.5a	0.13	0.27
E17.5b	0.19	0.47
E17.5c	0.26	0.67
E18.1a	0.26	0.27
E18.1b	0.26	0.27
E18.1c	0	0.06
E18.2a	0.19	0.2
E18.2b	0	0.06
E18.2c	0.19	0.2
E18.3a	0.26	0.4
E18.3b	0.51	0.6
E18.3c	0.26	0.27
E18.4a	0.13	0.06
E18.4b	0.13	0.27
E18.4c	0.26	0.27
E18.5a	0.13	0.13
E18.5b	0.06	0.27
E18.5c	0.19	0.33
E19.1a	0.06	0.06
E19.1b	0.19	0.33
E19.1c	0.13	0.13
E19.2a	0.19	0.06
E19.2b	0.13	0.06
E19.2c	0.06	0.06
E19.3a	0.38	0.2
E19.3b	-0.06	0.33
E19.3c	0.06	0.33
E19.4a	0.13	0.4
E19.4b	0.19	0.06
E19.4c	0.32	0.47
E19.5a	0.13	0.27
E19.5b	0.38	0.47
E19.5c	0.26	0.2
E20.1a	0.13	0.2
E20.1b	0.38	0.2
E20.1c	0.26	0
E20.2a	0.64	0.47
E20.2b	0.19	0.13
E20.2c	0.45	0.33
E20.3a	0.19	0.13
E20.3b	0.13	0.06
E20.3c	0.32	0.2
E20.4a	0.26	0.27

Appendices

E20.4b	0.06	0.2
E20.4c	0.19	0.06
E20.5a	0.19	-0.06
E20.5b	0.19	0.27
E20.5c	0	0.33

Table K.8 – Discrimination Index percentages of the multiple choice task

	L1S	L2S
Very Good	8.33%	13%
Good	13%	4.66%
Regular	11%	27.67%
Deficient	67.67%	54.67%

APPENDIX L – IDEA UNITS AND ITEMS CREATED FOR EPISODE 16

Table L.1 – Idea units and items for the True / False (T / F), multiple choice (MC) and sequencing exercise in episode 16

IDEA UNIT	ITEM
TRUE / FALSE (T / F) EXERCISE	
Mrs. Renkins tells George she will have the boat fixed the next morning.	La Sra. Renkins le dice a Jorge que tendrá el barco arreglado esa misma tarde.
Bill tells George not to get close to the beavers while they work or they will get upset.	Bill le dice a Jorge que no se acerque mucho a los castores cuando trabajan.
George thinks that building a dam twice in one day was hard work.	Para Jorge, construir una presa dos veces el mismo día es muy fácil.
Bill tells George that the water's important to the beavers' home.	Bill le dice a Jorge que el agua es muy importante para los castores.
Bill brings a cow to the river and moves the tree from it. He thought the beavers would be happy because he moved the tree but they are not.	Bill no entiende por qué los castores no están contentos cuando sacan el árbol del río.
MULTIPLE CHOICE (MC) EXERCISE	
When George had his boat fixed, he rushed back to his perfect pond to set sail. Now he had a boat, but no pond. And the creek was only a trickle. Where could a whole pond go?	Cuando tiene el barco reparado, Jorge vuelve al estanque pero... A) La vaca se ha bebido el agua del estanque. B) En vez del estanque hay una muralla muy rara en el arroyo. C) No puede jugar con su barco porque el agua está contaminada.
The beavers throw water at George and he thinks that if the beavers don't want him watching, he'd just sail his boat.	Jorge piensa que si los castores no quieren que él les mire... A) No les va a dejar su barco para jugar. B) No les invitará a su casa a comer. C) Se limitará a navegar con su barco.
George goes to see Mrs. Renkins and she asks him if he has been having lots of fun with the boat.	La Sra. Renkins le pregunta a Jorge si... A) Se divierte con su barco. B) Los castores le han molestado mucho. C) Irá a cenar a su casa esa noche.
It is raining and George is at home. He thinks he doesn't mind rain. It means his pond will be bigger tomorrow.	A Jorge no le importa la lluvia, ya que piensa que así... A) Los castores se irán a otro sitio. B) Su estanque será más grande. C) Tendrá agua aunque su presa se rompa.
George is playing with the beavers. He is happy but when they eat his boat he thinks he will probably need Mrs. Renkins to build him more boats.	Cuando Jorge juega con los castores y estos se comen su barco, él piensa que... A) Le pedirá a la Sra. Renkins que le haga otros. B) No necesita ningún barco, ya que puede jugar con los castores. C) Irá a la tienda y comprará muchos barcos en las rebajas.

Appendices

IDEA UNIT	ITEM
SEQUENCING EXERCISE	
<p>Bill uses a cow to remove a tree from the river and, due to that, water goes directly to the beavers' dam. George and Bill watch the beavers work together to defend their home against the raging waters.</p>	<p>Jorge y Bill observan como los castores trabajan en equipo para defender su hogar.</p>
<p>George builds a dam and gets to have a pond to play with his boat but he realises he has forgotten it at home.</p>	<p>Jorge se ha olvidado el barco y vuelve a casa a buscarlo.</p>
<p>George wants to build a dam so he thinks he needs the scrap wood from Mrs. Renkins. When she goes to see her, she tells him to take any wood she wants.</p>	<p>La Sra. Renkins le dice a Jorge que coja la madera que necesite.</p>
<p>Bill arrives to the pond and explains George the beavers are working in the woods and he tells him he even took pictures of them.</p>	<p>Bill dice a Jorge que ha sacado fotos a los castores.</p>
<p>The beavers start throwing water at George to make him He thinks that the only thing he has to do is to build a dam like they did.</p>	<p>Jorge decide que hará una presa para él solo.</p>

