

**The oral production and negotiation of meaning
of Spanish EFL children
in task-supported interaction:
The role of age and task repetition**

María de los Ángeles Hidalgo Gordo

Supervised by:
Dr. María del Pilar García Mayo
Dr. Amparo Lázaro Ibarrola

Vitoria-Gasteiz 2017



Director: Dr. María del Pilar García Mayo
Universidad del País Vasco – Euskal Herriko Unibertsitatea (UPV/EHU)
Department of English and German Philology and Translation and
Interpretation

Co-director: Dr. Amparo Lázaro Ibarrola
Universidad Pública de Navarra
Department of Philology and Education

This doctoral dissertation has been funded by a scholarship from the Spanish Ministry of Economy and Competitiveness (Formación del Personal Investigador. Reference BES-2013-063291).

Acknowledgements

I would like to express my sincere and deep gratitude to all the people that have contributed in different ways to the development and completion of the present dissertation. Special thanks to Dr. María del Pilar García Mayo and Dr. Amparo Lázaro Ibarrola, my supervisors and mentors. Without their invaluable guidance, continuous supervision, sharing of their vast knowledge, and above all, their motivation and support during these four years, this research project would not have been possible. I have been privileged to have had the opportunity to work with these two great professionals and, even more so, wonderful people. It has been a pleasure to work with you, and I would love to be able to continue doing so in the future.

The study presented in this dissertation is part of a larger research project led by Dr. García Mayo, entitled *Oral interaction among young learners of English as a foreign language: The impact of the use of negotiation and corrective feedback strategies during communicative tasks on language learning*, funded by the Spanish Ministry of Education (FFI2012-32212). This project addresses the need to carry out more research on the way young foreign language learners learn a language, specifically from an interactionist perspective. This Doctoral thesis has been possible thanks to the research grant I was awarded (BES-2013-063291), through which I became a member of the research group mentioned above. Both my grant and the research project have been supported by the Spanish Ministry of Economy and Competitiveness, which is hereby gratefully acknowledged.

Thanks Dr. Laura Gurzynski-Weiss for welcoming me at Indiana University. I really appreciate her trust and support during my stay at this institution, facilitating every aspect of my stay. I have been very fortunate to have had the opportunity to work with such an inspiring person.

I would also like to thank Dr. JangDong Seo and Michael Frisby, from Indiana University Statistical Consulting Center, without their help, and their students' at the Statistical Consulting course, the process of analysing my data would have been much more arduous.

Of course, I am also very grateful to all the participants in this study, as well as to their teachers, and the headmaster of the CEIP Azpilagaña in Pamplona. Many thanks to the

caretakers in the school who always helped me to set up the classrooms for the experiment. Without these people, this project would not be a reality.

I would also like to thank my colleagues of the Language and Speech Laboratory (LASLAB) research group at the University of the Basque Country in Vitoria. Thanks Ainara and Marta for your help and guidance when I first arrived to the laboratory four years ago, and for the nice moments we have spent together. I have really missed you this last year!

Marcel, my husband, deserves a special mention. He has been my emotional support from the beginning to the end of this process. Thank you so much for putting up with me during the rollercoaster these four years have been, and in spite of it, marrying me just a few months ago! Thanks for reading my work, for listening to my presentations, and for always giving me your best smile. You are the best partner in love and life.

To all my family, especially to my Mum and Dad, Gloria and Cristóbal, who have always believed in me and have always given me everything I needed, and much more. Porque ellos han sido mi apoyo incondicional, mi modelo a seguir y mi inspiración, y porque sin su esfuerzo no sería quien soy ni hubiese llegado a donde me encuentro hoy. Gracias por estar siempre ahí, aunque a mil kilómetros (literalmente), pero por haceros sentir siempre tan cerca. ¡Muchas gracias!

Without the help of all of the people mentioned above, I would never have been able to complete this dissertation. To all of them goes my sincere gratitude. All of them are part of this project, hence, part of my past and my future.

TABLE OF CONTENTS

TABLE OF CONTENTS	V
ABBREVIATIONS	IX
TABLES	XI
FIGURES	XIII
GRAPHS	XIII
ABSTRACT	XV
RESUMEN	XVII
INTRODUCTION	1
PART I LITERATURE REVIEW	7
CHAPTER 1 THE INTERACTION HYPOTHESIS	9
1.1 THE INTERACTION HYPOTHESIS	9
1.2 INPUT, OUTPUT AND FEEDBACK	14
1.2.1 COMPREHENSIBLE INPUT	14
1.2.2 OUTPUT	15
1.2.3 FEEDBACK	16
1.2.4 OTHER VARIABLES	17
1.3 NEGOTIATION OF MEANING	21
1.4 NEGOTIATION STRATEGIES AND RECASTS	23
1.5 L1 USE	28
1.6 LIMITATIONS TO THE ROLE OF INTERACTION IN SLA	40
1.7 CONCLUSION	41
CHAPTER 2 TASKS	43
2.1 TASK-BASED LANGUAGE TEACHING	43
2.2 WHAT IS A TASK?	46
2.3 TASK CLASSIFICATIONS	50
2.3.1 JIGSAW TASKS	57
2.4 TASK REPETITION	61
2.4.1 EFFECTS OF REPETITION ON GENERAL COMPETENCE	63
2.4.2 EFFECTS OF REPETITION ON NEGOTIATION OF MEANING	74
2.5 CONCLUSION	75
CHAPTER 3 CHILD SECOND LANGUAGE LEARNING	77
3.1 CHILD SLA RESEARCH	77
3.2 YOUNG LEARNERS' DISTINCTIVE CHARACTERISTICS	82
3.3 CHILD INTERACTION	87

3.4	MEASURING ORAL PRODUCTION DURING INTERACTION: LANGUAGE PROFICIENCY VARIABLES (CAF)	89
3.5	SECOND VS. FOREIGN LANGUAGE ACQUISITION	95
3.5.1	YLS' TASK-BASED INTERACTION IN SL LEARNING CONTEXTS	102
3.5.2	YLS' TASK-BASED INTERACTION IN FL LEARNING CONTEXTS	105
3.6	CONCLUSION	110
CHAPTER 4	CONTENT AND LANGUAGE INTEGRATED LEARNING (CLIL)	113
4.1	EMERGENCE AND SPREAD OF CLIL IN EUROPE	113
4.2	CLIL PREDECESSORS: IMMERSION PROGRAMMES AND CONTENT-BASED INSTRUCTION	119
4.2.1	CLIL, IMMERSION AND CBI, SAME OR DIFFERENT?	120
4.3	CLIL: DEFINITION AND MAIN CHARACTERISTICS	123
4.4	CLIL RESEARCH REVIEW: CLIL AND LEARNING OUTCOMES	127
4.5	CLIL IN SPAIN	133
4.5.1	CLIL IN NAVARRE	143
4.6	CONCLUSION: RESEARCH GAPS IN CLIL AND FUTURE DIRECTIONS	144
	PART II THE PRESENT STUDY	149
CHAPTER 5	THE STUDY	151
5.1	MOTIVATION OF OUR STUDY	151
5.2	RESEARCH QUESTIONS AND HYPOTHESES	153
5.3	PARTICIPANTS AND SETTING	158
5.4	MATERIALS AND PROCEDURE	160
5.4.1	THE TASK IN OUR STUDY: THE PICTURE PLACEMENT TASK	160
5.4.2	PROCEDURE	162
5.5	CODIFICATION AND DATA ANALYSIS	164
CHAPTER 6	RESULTS	183
6.1	NEGOTIATION STRATEGIES IN YEAR 3 AND YEAR 5	183
6.2	EFFECT OF PROCEDURAL TASK REPETITION AND AGE ON LEARNERS' ORAL PERFORMANCE	189
6.2.1	IMPACT ON NEGOTIATION STRATEGIES	190
6.2.2	IMPACT ON CAF MEASURES	196
6.2.2.1	IMPACT ON ORAL COMPLEXITY	197
6.2.2.2	IMPACT ON ACCURACY	202
6.2.2.3	IMPACT ON FLUENCY	207
6.2.2.4	CAF CORRELATIONS	212
6.2.3	IMPACT ON L1 USE	217
6.4	SUMMARY	225

CHAPTER 7	DISCUSSION OF MAIN FINDINGS	231
7.1	YOUNG LANGUAGE LEARNERS' NEGOTIATION OF MEANING	231
7.2	EFFECT OF PROCEDURAL TASK REPETITION ON YOUNG LEARNERS' ORAL PRODUCTION	236
7.2.1	PROCEDURAL TASK REPETITION AND NEGOTIATION OF MEANING	236
7.2.2	PROCEDURAL TASK REPETITION AND CAF	237
7.2.3	PROCEDURAL TASK REPETITION AND L1 USE	239
7.3	EFFECT OF AGE ON YOUNG LEARNERS' ORAL PRODUCTION	243
7.3.1	AGE AND NEGOTIATION OF MEANING	244
7.3.2	AGE AND CAF	245
7.3.3	AGE AND L1 USE	249
7.4	CONCLUSION	252
	PART III CONCLUSIONS AND CONTRIBUTIONS	255
CHAPTER 8	CONCLUSION: IMPLICATIONS, LIMITATIONS AND FUTURE RESEARCH	257
8.1	CONCLUSIONS	257
8.2	PEDAGOGICAL IMPLICATIONS	258
8.3	LIMITATIONS AND FUTURE RESEARCH DIRECTIONS	261
	REFERENCES	265
	APPENDIXES	287
	APPENDIX A	287
	APPENDIX B	289
	APPENDIX C	293
	APPENDIX D	296

ABBREVIATIONS

AICLE	Aprendizaje Integrado de Contenidos y Lengua
AS-unit	Analysis of Speech Unit
CAF	Complexity, Accuracy, Fluency
CBI	Content-Based Instruction
CEFR	Common European Framework of Reference for Languages
CEIL	Content and English Integrated Learning
CLIL	Content and Language Integrated Learning
CPH	Critical Period Hypothesis
CV	Correct verb form
EFL	English as a Foreign Language
EFAS	Error-free AS-unit
ESP	English for Specific Purposes
ESL	English as a Second Language
EU	European Union
FL	Foreign Language
FonF	Focus on Form
L1	Mother tongue/First language
L2/SL	Second Language
L3	Third Language
LRE	Language Related Episode
NEFAS	Non-error-free AS-unit
NoM	Negotiation of meaning
NS	Native speaker
NNS	Non-native speaker
NonCV	Incorrect verb form

PISA	Programme for International Student Assessment
PPT	Picture Placement task
SLA	Second Language Acquisition
SPL	Spanish as a Foreign Language
TBLT	Task-Based Language Teaching
TL	Target Language
TR	Task repetition
UG	Universal Grammar
YLS	Young learners

TABLES

TABLE 1 <i>COMMUNICATION TASK TYPES FOR L2 RESEARCH AND PEDAGOGY ANALYSIS BASED ON: INTERACTANT (X/Y) RELATIONSHIPS AND REQUIREMENTS IN COMMUNICATING INFORMATION (INF) TO ACHIEVE TASK GOALS (PICA ET AL., 1993: 19).</i>	52
TABLE 2 <i>SUMMARY OF THE FINDINGS ON THE EFFECTS OF TR ON LEARNERS' CAF.</i>	70
TABLE 3 <i>SUMMARY OF THE FINDINGS ON THE ROLE OF DIFFERENT TYPES OF TR ON LEARNERS' CAF.</i>	72
TABLE 4 <i>CHILDREN: THE THREE AGE GROUPS (TAKEN FROM PINTER, 2011: 2).</i>	84
TABLE 5 <i>MAIN DIFFERENCES BETWEEN ADULTS AND CHILDREN.</i>	87
TABLE 6 <i>TASK CHARACTERISTICS' INFLUENCE UPON L2 PERFORMANCE (CAF) (ADAPTED FROM SKEHAN, 2003: 5-6).</i>	91
TABLE 7 <i>CONTRASTING PREDICTIONS MADE BY SKEHAN'S AND ROBINSON'S HYPOTHESES (TAKEN FROM SKEHAN AND FOSTER, 2012: 215).</i>	92
TABLE 8 <i>DIFFERENT CAF MEASURES.</i>	94
TABLE 9 <i>CONTRASTING FOREIGN AND SECOND LANGUAGE CONTEXTS (TAKEN FROM PINTER (2011: 87)).</i>	96
TABLE 10 <i>SUMMARY OF THE FINDINGS ON THE EFFECTS OF TASK-BASED INTERACTION AMONG YLS IN SL SETTINGS.</i>	98
TABLE 11 <i>SUMMARY OF THE FINDINGS ON THE EFFECTS OF TASK-BASED INTERACTION AMONG YLS IN FL SETTINGS.</i>	100
TABLE 12 <i>BENEFITS OF CLIL METHODOLOGY.</i>	126
TABLE 13 <i>MINIMUM WEEKLY TIMETABLE OF CLIL LESSONS (ADAPTED FROM EURYDICE, 2006B: 6).</i>	134
TABLE 14 <i>PARTICIPANTS' PROFILE.</i>	159
TABLE 15 <i>PICTURE PLACEMENT TASK CHARACTERISTICS.</i>	161
TABLE 16 <i>TIME INVESTED IN EACH TASK BY THE PARTICIPANTS IN EACH GROUP.</i>	163
TABLE 17 <i>TIME DIFFERENCES ACROSS TASKS.</i>	163
TABLE 18 <i>OVERALL TIME INVESTED IN EACH TASK BY THE TWO GROUPS.</i>	164
TABLE 19 <i>CODIFIED NoM STRATEGIES.</i>	172
TABLE 20 <i>CODIFIED LI FUNCTIONS.</i>	177
TABLE 21 <i>CAF MEASURES.</i>	181
TABLE 22 <i>NoM STRATEGIES USED BY YEAR 3 AND YEAR 5 LEARNERS DURING ORAL INTERACTION AT T1.</i>	184
TABLE 23 <i>PERCENTAGES OF THE FUNCTIONS SERVED BY THE NoM STRATEGIES USED BY YEAR 3 AND YEAR 5 LEARNERS DURING ORAL INTERACTION AT T1.</i>	185
TABLE 24 <i>WITHIN- AND BETWEEN-GROUP DIFFERENCES IN THE FUNCTIONS SERVED BY NoM STRATEGIES AT T1.</i>	188
TABLE 25 <i>DESCRIPTIVE STATISTICS FOR THE FUNCTIONS SERVED BY THE NoM STRATEGIES IN THE PERFORMANCE OF THE TWO AGE GROUPS ACROSS TESTING TIMES (T1, T2, T3).</i>	191
TABLE 26 <i>MAIN EFFECTS AND INTERACTION EFFECTS OF TASK TIME AND AGE GROUP ON NoM STRATEGIES.</i>	192
TABLE 27 <i>DIFFERENCES IN THE USE OF NoM STRATEGIES ACROSS TASK TIMES.</i>	193
TABLE 28 <i>DESCRIPTIVE STATISTICS OF THE WORDS, CLAUSES AND AS-UNITS OF ALL THE PERFORMANCES FOR ALL PARTICIPANTS (N= 40).</i>	197
TABLE 29 <i>MAIN EFFECTS AND INTERACTION EFFECTS OF TASK TIME AND AGE GROUP ON WORDS, CLAUSES AND AS-UNITS.</i>	198
TABLE 30 <i>DESCRIPTIVE STATISTICS FOR THE THREE COMPLEXITY MEASURES IN THE THREE TASKS.</i>	199
TABLE 31 <i>MAIN EFFECTS AND INTERACTION EFFECTS OF TASK TIME AND AGE GROUP ON COMPLEXITY.</i>	199
TABLE 32 <i>DESCRIPTIVE STATISTICS FOR THE THREE ACCURACY MEASURES IN THE THREE TASKS.</i>	203
TABLE 33 <i>MAIN EFFECTS AND INTERACTION EFFECTS OF TASK TIME AND AGE GROUP ON ACCURACY.</i>	204
TABLE 34 <i>DESCRIPTIVE STATISTICS FOR THE TWO FLUENCY MEASURES IN THE THREE TASKS.</i>	208
TABLE 35 <i>MAIN EFFECTS AND INTERACTION EFFECTS OF TASK TIME AND AGE GROUP ON FLUENCY.</i>	208
TABLE 36 <i>TASK TIME EFFECTS ON FLUENCY.</i>	209
TABLE 37 <i>TRADE-OFF EFFECTS BETWEEN CAF MEASURES.</i>	216
TABLE 38 <i>DESCRIPTIVE STATISTICS OF LI USE ACROSS TASKS.</i>	217
TABLE 39 <i>DESCRIPTIVE STATISTICS OF LI FUNCTIONS ACROSS TASKS.</i>	219

TABLE 40 <i>MAIN EFFECTS AND INTERACTION EFFECTS OF TASK TIME AND AGE GROUP ON LI FUNCTIONS.</i>	220
TABLE 41 <i>TASK TIME EFFECTS ON LI FUNCTIONS.</i>	221
TABLE 42 <i>OFF-TASK TALK BY YEAR 3 LEARNERS.</i>	223
TABLE 43 <i>SUMMARY OF THE EFFECT OF PROCEDURAL TR ON YLS' ORAL PRODUCTION.</i>	227
TABLE 44 <i>SUMMARY OF THE EFFECTS OF PROCEDURAL TR ON THE PERFORMANCE OF THE TWO AGE GROUPS.</i>	229
TABLE 45 <i>BETWEEN-SUBJECTS STATISTICALLY SIGNIFICANT DIFFERENCES IN THE CHANGES DERIVED FROM PROCEDURAL TR.</i>	230

FIGURES

FIGURE 1 <i>MAJOR TENETS OF THE INTERACTION APPROACH (ADAPTED FROM MACKEY AND POLIO, 2009: 5).</i>	18
FIGURE 2 <i>BENEFITS OF TR AND PEER-PEER INTERACTION.</i>	107
FIGURE 3 <i>STATUS OF TLS USED FOR CLIL IN PRIMARY AND/OR GENERAL SECONDARY EDUCATION, 2010/11 (TAKEN FROM EURYDICE, 2012: 41).</i>	118
FIGURE 4 <i>THREE-DIMENSIONAL CLIL RESEARCH SPACE (TAKEN FROM DALTON-PUFFER ET AL. (2010: 10)).</i>	128
FIGURE 5 <i>SIMPLIFIED VERSION OF THE SPANISH LINGUISTIC MAP (TAKEN FROM FERNÁNDEZ FONTECHA, 2009: 5).</i>	135
FIGURE 6 <i>DIFFERENCES BETWEEN THE FUNCTIONS SERVED BY NOM STRATEGIES IN EACH GROUP.</i>	187
FIGURE 7 <i>STAGES OF THE ACQUISITION ROUTE OF NOM STRATEGIES.</i>	236
FIGURE 8 <i>YLS' NOM.</i>	252
FIGURE 9 <i>EFFECT OF PROCEDURAL TR ON YLS' ORAL PRODUCTION.</i>	253

GRAPHS

GRAPH 1 <i>CHANGES IN THE PERCENTAGES OF FUNCTIONS OF THE NOM STRATEGIES ACROSS TASKS.</i>	196
GRAPH 2 <i>RATIO OF WORDS PER AS-UNIT AT THE THREE TESTING TIMES.</i>	200
GRAPH 3 <i>RATIO OF CLAUSES PER AS-UNIT AT THE THREE TESTING TIMES.</i>	201
GRAPH 4 <i>LEXICAL D AT THE THREE TESTING TIMES.</i>	202
GRAPH 5 <i>RATIO OF ERRORS PER AS-UNIT AT THE THREE TESTING TIMES BY THE TWO GROUPS.</i>	205
GRAPH 6 <i>PERCENTAGE OF EFAS PER AS-UNIT AT THE THREE TESTING TIMES.</i>	206
GRAPH 7 <i>RATIO OF CORRECT VERBS OVER THE TOTAL NUMBER OF VERBS.</i>	207
GRAPH 8 <i>DEVELOPMENT OF THE RATIO OF WORDS PER MINUTE.</i>	210
GRAPH 9 <i>DEVELOPMENT OF THE RATIO OF LI TERMS PER AS-UNIT.</i>	211
GRAPH 10 <i>OVERALL LI ACROSS TASKS BY YEAR 3 AND YEAR 5.</i>	218
GRAPH 11 <i>PERCENTAGE OF USE OF THE LI FUNCTIONS ACROSS TASKS.</i>	222

ABSTRACT

Numerous studies have demonstrated that negotiated interaction benefits second and foreign language acquisition (Gass and Mackey, 2007; Long, 1983, 1985, 1996, Mackey, 2012). Thus, collaborative tasks have been claimed to be an effective tool for language learning, not only by studies within interactionist perspectives (Mackey, 2012), but also by those from the field of task-based language teaching (Van den Branden, Bygate and Norris, 2009). Among the variables that contribute to their effectiveness, the present study focuses on procedural repetition, which is of great interest to language instructors, who often repeat tasks more than once with more or fewer modifications. To date, researchers have mainly addressed the effects of this variable on general competence (measured in terms of complexity, fluency and accuracy) (García Mayo, Imaz Agirre and Azkarai, in press; Kim and Tracey-Ventura, 2013; Lynch and Maclean, 2000, 2001; Sample and Michel, 2014) while only few studies have analysed its impact on learners' amount of negotiation (Lázaro Ibarrola and Hidalgo, 2017b; Mackey, Kanganas and Oliver, 2007). The other pillar of our study is the specific nature of young learners' (YLS) foreign language acquisition and the aspects that influence it (specifically age and procedural task repetition). Despite the increasing attention given to YLS' language acquisition, most of the existing literature refers to studies with adults or children in second language contexts, while work addressing children in foreign language settings, a population of remarkable expansion (Cameron 2003; Pinter, 2007, 2011, 2017), are still comparatively scarce (García Mayo, 2017; García Mayo and Lázaro Ibarrola, 2015; Lázaro Ibarrola and Azpilicueta Martínez, 2015; Pinter, 2007). Thus, child foreign language acquisition, as well as the potential differences between children of different ages, is an area of study in need of further examination.

In order to fill this double research niche, the aim of the current dissertation is to explore the potential of peer-peer interaction and document the conversational exchanges of two groups of 40 YLS of English as a foreign language (EFL) (ages 8-9 and 10-11) while they resolve a picture placement task in a Content and Language Integrated Learning (CLIL) school in Spain. From an interactionist perspective, we have focused on the impact of age and procedural task repetition on the children's oral production when carrying out the task three times within a one-week interval between each data collection point. Specifically, the changes in the negotiation of meaning strategies they employ, their general performance (complexity, accuracy and fluency) and use of their first language

(L1) have been examined. Additionally, the differences between the two age groups regarding these aspects have also been addressed.

Results reveal significant differences in the use of negotiation strategies between the two groups: whereas the most common strategies in the older learners' production were those used to confirm that the message has been successfully understood, the younger participants negotiated mainly to solve communication problems. As regards the learners' general competence, older learners exhibited greater structural complexity, accuracy and fluency in their oral production. Finally, the younger learners resorted to their shared L1 most frequently. Nevertheless, as reported in previous studies, the learners' use of the L1 was not excessive, and mostly served functions that facilitate task completion and, eventually, lead to language acquisition.

The effect of procedural task repetition on the performance of these YLs is more modest than the influence of age. The most remarkable finding is the significant improvement in fluency in the output of the two groups. Structural complexity was also positively affected, although the improvement does not reach statistical significance. Two of the functions served by the negotiation strategies slightly decreased (strategies to repair and to prevent communication breakdowns), while the other two remained stable upon task repetition (strategies to confirm successful communication and those to focus on form). Procedural task repetition also lead to a drop in L1 use.

Our study emphasises the importance of considering the specific characteristics of learners' age since, as our results indicate, significant differences are present in the performance of learners of ages not too far apart. These findings widen our knowledge of the nature of young language learners' oral interaction. Positive evidence for task-supported interaction among age- and level-matched peers has also been offered, as children with a limited command of the target language have been found to be able to successfully interact and complete the task with no help from the researcher. Furthermore, the value of procedural task repetition was shown, as the learners' oral production improves upon task repetition.

In light of these results, pedagogical implications of the repetition of tasks that follow the same procedure in YLs' language classrooms will be discussed.

RESUMEN

La teoría interaccionista (Long, 1981, 1983, 1985, 1996), uno de los marcos de investigación más establecidos en el área de adquisición de segundas lenguas, establece que la interacción oral facilita en gran medida la adquisición de un idioma (Gass y Mackey, 2007; Loewen, 2005; McDonough 2005; McDonough y Mackey, 2000, 2006, 2008), especialmente cuando los hablantes negocian para entender el mensaje. Este tipo de interacción se conoce como negociación de significado y normalmente ocurre cuando hay un problema de comunicación que impulsa a los hablantes a modificar su producción inicial para alcanzar un entendimiento mutuo. Durante el proceso de negociación, los aprendices se hacen conscientes de las lagunas existentes en su conocimiento de la lengua, las cuales pueden estar relacionadas con la pronunciación, el vocabulario o la morfosintaxis. Por este motivo, un requisito importante es la presencia de un compañero activo que coopere en la negociación del significado, proceso a través del cual se estimula un posterior aprendizaje.

Puesto que las modificaciones que surgen de la interacción facilitan el aprendizaje de una lengua, otro aspecto esencial de este proceso son los ajustes conversacionales que tienen lugar durante la negociación y que estimulan estas modificaciones. Estos ajustes conversacionales o estrategias de negociación, pueden ser explícitos, a través de comentarios metalingüísticos, o implícitos, como los que consideraremos en nuestro estudio. Tradicionalmente, se han examinado los siguientes tipos de estrategias de negociación implícita: solicitudes de aclaración, confirmaciones de comprensión, verificaciones de comprensión, repeticiones y reformulaciones (Oliver, 1998, et passim). Una categorización reciente de Lázaro Ibarrola e Hidalgo (2017a) clasifica estas estrategias según la función que cumplen en la negociación: i) prevenir rupturas en la comunicación, ii) confirmar una buena comunicación, iii) reparar problemas en la comunicación, iv) enfoque en aspectos formales. Con esta nueva clasificación se evita considerar más de una vez estrategias que, aunque podrían corresponder a varios tipos, realmente cumplen una única función.

Consecuentemente, investigadores y profesores se han centrado en identificar las condiciones de aprendizaje que generan patrones de interacción, resultando el uso de tareas colaborativas un método especialmente beneficioso. Este tipo de tareas, en las que los participantes han de trabajar juntos para alcanzar un fin común, son un instrumento

muy valioso para fomentar la interacción oral y la negociación de significado, lo cual queda reflejado en los nuevos materiales de enseñanza de lenguas (García Mayo, 2007; Van den Branden, Bygate y Norris, 2009). Al trabajar con tareas colaborativas, los aprendices han de prestar atención tanto al significado como a la forma del mensaje que quieren transmitir, lo cual resulta especialmente complejo debido a la naturaleza espontánea de la producción oral. Durante el transcurso de la tarea habrá ocasiones en las que los participantes tendrán que negociar el significado, lo que les llevará a modificar su producción inicial y, generalmente, les acercará a una producción más correcta. Entre las numerosas variables posibles del estudio de las tareas colaborativas (tipo, complejidad, procedimiento), nos centraremos en la familiaridad de los aprendices con la tarea.

Debido a nuestra limitada capacidad de procesamiento (Skehan, 1998; Skehan y Foster, 2001), tendemos a priorizar la comunicación del significado (Azkarai, 2013; García Mayo, 2011; Pica, 2002; Swain y Lapkin, 2001). Este es uno de los motivos por los que la repetición de tareas se ha convertido en una práctica muy efectiva para dirigir la atención de los aprendices del contenido del mensaje a la selección de la forma correcta del lenguaje (Bygate, 1999; Mackey, 2007). La repetición de tareas, y la familiaridad con las mismas adquirida a través de la repetición, deriva en la mejora de diferentes dimensiones de la producción oral, tales como la complejidad, la corrección y la fluidez y, en general, en una mejor organización de los recursos lingüísticos. Asimismo, otros aspectos de la interacción, tales como la negociación de significado, el suministro y uso de la retroalimentación, y el uso de la primera lengua (L1) también se ven afectados (Bygate, 1996; 2001; Bygate y Samuda, 2005; Kim y Tracey-Ventura, 2013; Lynch y Maclean 2000; 2001; Patarnasorn, 2010; Pinter, 2007; Sample y Michel, 2014; Samuda y Bygate, 2008).

La investigación de la forma en la que los jóvenes aprendices de lenguas interactúan es bastante reciente. Los estudios sobre esta población son aún escasos en comparación con la extensa bibliografía que se centra en la teoría interaccionista, y lo que ha ocurrido hasta hace poco es que los resultados de estudios con adultos se han aplicado a propuestas pedagógicas para niños sin mayores modificaciones. Niños y adultos se encuentran en diferentes etapas de desarrollo, por lo que interactúan y aprenden de diferente manera (García Mayo, 2017a; García Mayo y García Lecumberri, 2003; Pinter, 2007; Singleton y Ryan, 2004). Así pues, la validez de aplicar los resultados de la investigación en la que participan adultos a niños es poco clara (Oliver, 1998). A pesar de que los aprendices

jóvenes son capaces de interactuar y entablar una conversación, haciendo uso de diferentes estrategias de negociación (García Mayo y Lázaro Ibarrola, 2015; Lázaro Ibarrola e Hidalgo, 2017a,b; Oliver, 1998, 2002), la habilidad para colaborar y comprender las necesidades del interlocutor se desarrolla con la edad, la cual que se convierte en un factor de gran influencia al trabajar con tareas colaborativas en una segunda lengua (Oliver, 2009; Pinter, 2007).

Con la presente tesis, pretendemos arrojar más luz sobre estas áreas del aprendizaje de lenguas, analizando un sector de población que, a pesar de su rápido incremento, permanece relativamente poco investigado: los jóvenes aprendices de inglés como lengua extranjera (ILE).

El estudio

Objetivos y preguntas de investigación

El principal objetivo de esta tesis es explorar el potencial de la interacción oral entre jóvenes aprendices de ILE. Estudiaremos las interacciones de alumnos de 3º y 5º de Educación Primaria (de entre 8 y 11 años) para determinar si existen diferencias relacionadas con la edad de los participantes así como los efectos de la repetición de tareas que siguen el mismo procedimiento en su producción oral. Concretamente, examinaremos el impacto de estas dos variables en la negociación de significado, la competencia general (complejidad, corrección y fluidez), y el uso de la L1. Para alcanzar estos objetivos, hemos formulado las siguientes preguntas de investigación:

Negociación de significado

1. ¿Cómo negocian los jóvenes aprendices de ILE cuando realizan una tarea oral colaborativa?
2. ¿Qué estrategias de negociación usan en sus interacciones orales con interlocutores de su misma edad y nivel de la lengua meta?
3. ¿Utilizan las mismas estrategias los niños de diferente edad?

Repetición de tareas con el mismo procedimiento

4. ¿Afecta la repetición de tareas a i) el uso de las estrategias de negociación, ii) la competencia general en el idioma, iii) el uso de la L1?

Edad

5. ¿Influye la edad en i) el uso de las estrategias de negociación, ii) la competencia general en el idioma, iii) el uso de la L1, de los jóvenes aprendices de lengua extranjera?

Para responder a nuestras preguntas de investigación, hemos analizado la interacción oral de dos grupos de jóvenes aprendices españoles de ILE mientras realizaban una tarea comunicativa bidireccional (*picture placement*) tres veces con un intervalo de una semana entre cada recogida de datos.

Metodología

Participantes

En este estudio participaron 80 jóvenes aprendices de ILE, de edades entre 8-9 y 10-11 años, de una escuela pública de educación primaria en Pamplona. El grupo de alumnos más jóvenes (N= 40) estaba en su tercer curso de educación primaria (8-9 años), y el otro grupo (N= 40) cursaba 5º de educación primaria (10-11 años). En este colegio todos los alumnos siguen un programa de Aprendizaje Integrado de Contenidos y Lenguas Extranjeras (AICLE), que es obligatorio. De este modo, hemos eliminado la posibilidad de que sólo los estudiantes más motivados o con mayor nivel en la lengua meta participen en el estudio.

Tarea y procedimiento: The picture placement game

Nuestra tarea ha sido diseñada por la investigadora y sus supervisoras, teniendo en cuenta las tareas empleadas en estudios anteriores con poblaciones similares así como los contenidos que los niños estaban trabajando en clase de inglés. Se utilizaron tres sets de materiales, uno en cada sesión.

Los alumnos trabajaron en parejas formadas por los maestros de cada grupo. Todas las parejas realizaron la tarea tres veces, en tres semanas, y en cada sesión se utilizó un póster ligeramente diferente.

Cada miembro de la pareja trabajó con un póster que ilustraba dos escenas (colegio, parque, estancias en una casa), y en el que aparecían dos niños realizando alguna actividad. Las escenas ilustradas en los pósters de cada alumno eran idénticas mientras

que los niños que aparecían en éstos eran diferentes. Cada participante tenía además cuatro fotos de niños de entre las que debía averiguar cuáles eran las que estaban en el póster de su compañero o compañera. El objetivo de estas tareas es que los aprendices interactúen en inglés para completar sus pósters de modo que ambos acabasen con las mismas fotos de niños en los mismos lugares en los pósters.

Resultados

Los resultados obtenidos demuestran que la repetición de tareas con el mismo procedimiento influye en el empleo de estrategias de negociación, la competencia general y el uso de la L1 en diferente medida. La repetición parece haber causado una disminución significativa del uso de dos de las funciones cumplidas por las estrategias de negociación (concretamente, las empleadas para prevenir problemas en la comunicación y las que se usan para repararlos). Las otras dos funciones se han mantenido estables. En cuanto al efecto de la repetición en la competencia general oral de los aprendices, hemos detectado que tanto la fluidez así como una de las medidas usadas para analizar la complejidad de la producción de los participantes mejoraron significativamente en la última repetición. Sin embargo, la corrección formal no experimentó ningún cambio significativo, lo cual se puede relacionar a un posible efecto de compensación entre este aspecto y la complejidad, como muestran los análisis de correlación entre estas dos dimensiones.

Los jóvenes aprendices de ILE en los dos grupos de edad observados son capaces de interactuar entre ellos, haciendo uso de diferentes estrategias de negociación. Sin embargo, hemos detectado diferencias claras en cuanto a las estrategias más comunes en cada grupo: mientras que los participantes más jóvenes negocian principalmente para resolver problemas con la comunicación, los mayores usan con más frecuencia estrategias que informan al interlocutor de que el mensaje ha sido correctamente recibido. El uso de estrategias para confirmar una buena comunicación es significativamente más frecuente en la producción de los alumnos de 5º de primaria, siendo esta función la más común en este grupo. Las estrategias más comúnmente usadas por los aprendices de 3º de primaria son aquellas para reparar problemas en la comunicación, aunque no hemos encontrado diferencias significativas en el empleo de estas estrategias en los dos grupos. De este modo, nuestros resultados parecen indicar que la capacidad de asistir de manera activa al interlocutor aumenta con la edad. Las estrategias de enfoque en aspectos formales y para prevenir dificultades en la comunicación son las menos comunes en ambos grupos y se

utilizan con frecuencia similar. La repetición de la tarea también afectó de manera diferente a la producción oral de los dos grupos: el uso de estrategias para prevenir problemas en la comunicación disminuyó significativamente en la producción de los aprendices mayores. Asimismo, se han observado diferencias significativas en el desarrollo del uso de estrategias que sirven esta función entre la primera y la última tarea de cada grupo. Por último, los cambios en el uso de estrategias para confirmar buena comunicación entre la segunda y la tercera tarea también fueron estadísticamente significativos.

Del mismo modo, las diferencias entre los dos grupos son también notables en lo que se refiere a la competencia general. La producción oral de los alumnos de 5° de primaria exhibe una mayor complejidad, medida en palabras y frases por AS-unit, es más correcta, en lo que corresponde al número de AS-units que no contienen errores, y más fluida (palabras por minuto y menor ocurrencia de términos de la L1). Además, la comparación del desarrollo, a través de la repetición, de estas dos medidas de complejidad también mostró diferencias significativas entre los dos grupos. Por último, aunque tanto la complejidad estructural como la fluidez mejoraron en la tercera repetición de los dos grupos, la mejora fue significativa sólo en el caso de la fluidez oral.

Coincidiendo con estudios previos, nuestros resultados muestran el uso limitado que los aprendices de idiomas hacen de su L1. Las funciones más frecuentes de la L1 en la producción de los jóvenes aprendices son claramente beneficiosas para el aprendizaje de una segunda lengua, en particular, la L1 se emplea para tratar el procedimiento de la tarea o para resolver dudas con el vocabulario necesario para completar la misma. La edad de los participantes de nuevo desempeña un papel claro: la dependencia de los aprendices más jóvenes de su L1 es significativamente mayor que la de los participantes mayores. Es además en la producción de este grupo donde tienen lugar los (escasos) ejemplos de uso de la L1 para tratar temas no relacionados con la tarea. Los dos grupos coinciden en las funciones de la L1 más frecuentes: vocabulario y discurso metacognitivo. Al analizar el uso de la L1 de los dos grupos en conjunto, estas dos funciones se ven afectadas de diferente manera: mientras que el vocabulario disminuye, el discurso metacognitivo aumenta en la última repetición. En lo que se refiere al efecto de la repetición de la tarea en cada grupo, la proporción de L1 por AS-unit no disminuye de manera significativa en ninguno de los dos grupos por separado. Sin embargo, las frecuencias de uso de las funciones que desempeña la L1 en cada tarea evolucionan de manera diferente en cada

uno de los grupos: mientras que el uso metacognitivo disminuye significativamente en la segunda tarea de los alumnos de 3^{er} curso y no sufre cambios significativos en la producción de los alumnos de 5^o, el vocabulario en L1 disminuye en la tercera tarea de estos aprendices y permanece estable en la producción de los alumnos más jóvenes. Igualmente, el desarrollo del uso de la L1 para vocabulario es diferente en cada grupo. Los participantes en ambos grupos también emplean marcadores del discurso en su L1, aunque de manera mucho menos frecuente que las otras dos funciones de la L1 y su uso no se ve afectado ni por la edad ni por la repetición.

Conclusiones

Con el presente estudio hemos intentado contribuir al conocimiento sobre cómo la interacción oral entre jóvenes aprendices y la repetición de tareas afectan al aprendizaje de lenguas. Nuestros resultados respaldan estudios previos que muestran que los jóvenes aprendices son capaces de interactuar y negociar el significado. Hemos mostrado que las tareas colaborativas son una herramienta muy valiosa para el aprendizaje de lenguas extranjeras ya que ofrecen oportunidades para interactuar en la lengua que se está aprendiendo en un contexto significativo. A pesar de su corta edad y limitado nivel de conocimiento del inglés, los jóvenes aprendices son capaces de entablar una conversación y de completar la tarea de manera autónoma en la lengua que están aprendiendo. Esto es especialmente significativo en contextos de lenguas extranjeras donde las oportunidades para usar la lengua meta fuera del aula son bastante limitadas.

Nuestros datos indican que los jóvenes aprendices negocian sobre todo para reparar problemas en la comunicación de un mensaje. Al madurar, sin embargo, se desarrolla la capacidad para asistir y tomar en cuenta las necesidades de los interlocutores (Pinter, 2007), lo que se hace evidente en el uso de estrategias de negociación que informan de que no se han producido dificultades en la comunicación. Las estrategias que cumplen esta función son precisamente las más frecuentes en la producción de los alumnos de 5^o de primaria. Estos alumnos proporcionan un apoyo mayor a sus interlocutores, al contrario que los alumnos más jóvenes, más preocupados por transmitir su propio mensaje. El uso de estrategias que cumplen estas dos funciones parece ser complementario ya que el aumento de unas estrategias a edades más avanzadas lleva consigo la disminución de las otras. Este fenómeno es visible a una menor escala a través de la repetición de tareas: mientras que las estrategias para prevenir problemas en la

comunicación siguen una tendencia descendente, las estrategias que confirman una buena comunicación aumentan gracias a la repetición.

Se aprecian más diferencias al comparar la competencia general de los dos grupos: la producción oral de los alumnos de 10-11 años es más compleja, correcta y fluida que la de los aprendices de 8-9 años. Gracias a la repetición de la tarea, ciertas dimensiones de la competencia lingüística de los alumnos más jóvenes mejoran de manera notable, en especial aspectos de la complejidad estructural, mientras que la mejora no fue tan destacada en la producción de los mayores. Esto podría deberse a que las tareas no fueron lo suficientemente exigentes y/o complejas para alumnos con un nivel más avanzado.

Por último, los resultados confirman que los jóvenes aprendices de lenguas extranjeras son capaces de completar tareas comunicativas en la lengua meta, empleando su L1 de manera moderada y con funciones que facilitan la realización de la tarea. Además, este estudio muestra que los alumnos más jóvenes hacen un mayor uso de la L1. En la producción de los alumnos de 3^{er} curso se han identificado algunos ejemplos de uso de la L1 para tratar asuntos no relacionados con la tarea, lo que evidencia una vez más un diferente nivel de desarrollo y un comportamiento más infantil que el de sus compañeros, dos cursos superiores.

Nuestros resultados apuntan a la importancia de tener en consideración las diferencias entre niños de distintas edades. Tal y como señalan estudios anteriores, la infancia es un periodo de numerosos cambios por lo que los niños de diferentes edades tienen necesidades distintas, y presentan diferentes comportamientos. Por lo tanto, ciertas actividades que tal vez sean adecuadas para alumnos de una determinada edad, pueden no resultar demasiado beneficiosas en otros grupos de edad.

En cualquier caso, los resultados obtenidos son muy satisfactorios y demuestran que la repetición de tareas con el mismo procedimiento ofrece ventajas a los jóvenes aprendices de lenguas en lo que concierne a la competencia general: tanto la fluidez como la complejidad estructural de la producción oral de los participantes mejoró en la tercera tarea. Por otro lado, también se han detectado efectos de compensación entre corrección y complejidad, aunque, en consonancia con la Trade-off Hypothesis (Skehan y Foster, 2012), los resultados obtenidos muestran que estos efectos negativos se atenúan gracias a las repeticiones. La repetición de tareas ofrece a los hablantes la posibilidad de

enfrentarse a una situación comunicativa similar más de una vez, lo que facilita una mejor y más efectiva distribución de la atención de los aprendices.

Nuestros resultados sugieren que una mejor organización de los recursos lingüísticos, lograda a través de la práctica, lleva a un menor uso de la L1. En este respecto, en la última tarea se han detectado indicadores de una posible diferenciación entre las funciones ‘relacionadas con la tarea’ y aquellas que se pueden considerar como ‘no relacionadas’. Mientras que las funciones ‘relacionadas con la tarea’ disminuyen en la última repetición (L1 vocabulario), el discurso metacognitivo y los conectores, que no serían necesarios para trabajar el contenido de la tarea, permanecen inalterados. Por otro lado, también hemos detectado que un incremento de la L1 va unido al empleo de un vocabulario más rico. De este modo, la L1 sirve como un elemento de apoyo que refuerza la seguridad de los aprendices para intentar hacer un máximo despliegue de sus conocimientos lingüísticos, por ejemplo usando un vocabulario más variado. Por estos motivos, el empleo de la L1 en el aula de idiomas no puede considerarse perjudicial para el aprendizaje sino que los profesores deberían ser conscientes de los beneficios que aporta la L1, e intentar sacar el mayor provecho de ella. Ha quedado demostrado una vez más que los aprendices recurren a su L1 como una herramienta extra necesaria y útil en determinadas etapas del proceso de aprendizaje.

Por todos estos motivos, los profesores no deberían mostrarse reticentes al uso de tareas con el mismo procedimiento y diferente contenido más de una vez en el aula de idiomas. Tal y como han mostrado diferentes estudios en este ámbito, y respaldan nuestros resultados, las oportunidades de enfrentarse más de una vez a situaciones comunicativas similares favorece el aprendizaje de una segunda lengua. Otro beneficio de esta práctica es la ventaja para los profesores de poder reutilizar y reciclar actividades de clase. Este aspecto no debería subestimarse ya que el tiempo de preparación de los maestros de educación primaria es normalmente limitado.

La presente tesis doctoral ha demostrado que la edad es un elemento crucial a considerar en el estudio de la adquisición de segundas lenguas. Nuestro objetivo principal ha sido arrojar luz sobre cómo facilitar el aprendizaje de segundas lenguas por parte de niños, centrándonos sobre todo en los beneficios que aporta la interacción oral entre pares al realizar tareas colaborativas más de una vez. Incluso dentro de la misma etapa de la infancia, y edades no muy distantes, hemos identificado importantes diferencias en la

producción oral de los dos grupos observados. Esperamos que los resultados obtenidos sirvan como guía para futuras prácticas pedagógicas y ayuden a los profesores y maestros a implementar nuevos métodos que ofrezcan más y mejores oportunidades de aprendizaje, en especial más ocasiones para interactuar usando la lengua meta, en el aula de idiomas.

Introduction

According to the Interaction Hypothesis (Long, 1983, 1985, 1996), and to studies within the interactionist framework (Gass and Mackey, 2007; Loewen, 2005; McDonough 2005; McDonough and Mackey, 2000, 2006), the positive effects of interaction are especially noticeable when negotiation of meaning occurs. Learners negotiate in order to reach mutual understanding and, during this process, they receive comprehensible input, as well as feedback on their output, which is often modified, thus providing opportunities for language learning (Mackey, 2012). Consequently, collaborative tasks that foster this type of interaction have become a valuable tool for second language acquisition, which is also reflected in new teaching materials (García Mayo, 2007; Van den Branden, Bygate and Norris, 2009).

In spite of the acknowledged benefits of interaction, some foreign language instructors are still concerned about their learners resorting to their first language (L1) instead of using the target language when working with collaborative tasks. This concern has been particularly raised when using communicative tasks with low proficiency learners in foreign language classrooms (Alegría de la Colina and García Mayo, 2009; Dicamilla and Antón, 2012; Tognini and Oliver, 2012). Nevertheless, research to date has shown that the use of the L1 is limited and serves functions that facilitate task completion and, eventually, leads to language learning (e.g. organisational purposes or to deliberate over vocabulary (scaffolding)) (Antón and Dicamilla, 1998; Azkarai, 2013; García Mayo and Hidalgo, 2017; García Mayo and Lázaro Ibarrola, 2015; Storch and Aldosari, 2010; Storch and Wigglesworth, 2003; Swain and Lapkin, 2000). All the existing literature on L1 use notwithstanding, this construct has not been sufficiently studied with young learners in foreign language settings.

As a matter of fact, in spite of the vast amount of research on language acquisition, several aspects of the way young learners acquire a foreign language have not been studied extensively. Thus, the validity of “findings from adult studies for determining pedagogical practice in child SLA” remains unclear and unproved (Oliver, 1998: 373). A very interesting aspect is children’s ability to collaborate and to understand and take into account their partner’s needs. Research indicates that these abilities develop and increase with age, which becomes an important factor of influence on interaction in a foreign language (Pinter, 2007). However, and despite the fact that it has been demonstrated that

there are differences between young and adult learners' foreign language acquisition and that age is an important variable (García Mayo, 2017a; García Mayo and García Lecumberri, 2003; Muñoz, 2007b; Singleton and Ryan, 2004), what has mainly happened so far is that adult results have been adopted without major modifications in pedagogic proposals for children. Hence, the current study aims to shed light on young learners' interactional behaviour in a foreign language setting.

Another aspect that contributes to the innovative nature of the current dissertation is the examination of the effect on young learners' performance of the repetition of a collaborative task with the same procedure but slightly different content. When working with collaborative tasks, learners have to focus both on form and meaning, which becomes especially hard because of the spontaneous nature of oral communication. Due to humans' limited processing capacity (Skehan, 1998) meaning tends to be prioritised over form (García Mayo, 2011; Pica, 2002; Swain and Lapkin, 2001), which is one of the reasons why task repetition has become a valuable way of diverting learners' attention from meaning to form (Bygate and Samuda, 2005; Pinter, 2007; Saeedi and Rahimi Kazerooni, 2014; Sample and Michel, 2014). Repetition is believed to lead to improvements in aspects of foreign language production, such as fluency, accuracy, complexity, and generally a more efficient organisation of language resources (Bygate, 1996, 2001; Bygate and Samuda, 2005; García Mayo et al., in press; Gass, Mackey, Álvarez-Torres and Fernández García, 1999; Kim and Tracey-Ventura, 2013; Pinter, 2007; Sample and Michel, 2014). Besides the different variables that contribute to its effectiveness, task repetition has also been explored because of the interest it raises for the language classroom where repeating similar or the same tasks is common practice.

The present study sheds more light into this double research niche: young learners' oral interaction in age- and level-matched dyads together with the influence of procedural task repetition on their oral production. Within the interactionist framework, we will analyse the oral interactions of two groups of young learners (n = 40) of English as a foreign language, aged between 8 and 11, at a beginner level of proficiency of the target language when performing a collaborative task. The participants attend a state school in Spain that follows a Content and Language Integrated Learning programme (CLIL) (Dalton-Puffer, 2007, 2011), which is mandatory for all students. This is also an important feature of our study since it eliminates any possibility of the 'disguised selection' argued to influence studies in CLIL settings. Some authors have claimed that only the most motivated and

those with a higher-than average proficiency in the target language are the ones who usually choose to attend these programmes when optional (Bruton, 2011a,b).

Specifically, the aim of this dissertation is to examine the nature of the negotiation of meaning as well as the different interactional strategies young learners use. Thus, we will investigate the extent to which young children negotiate for meaning and/or form when carrying out a two-way collaborative task (a picture placement game), and whether they are able to successfully complete it, interacting in the target language (English) with each other to achieve a common goal. Moreover, we will examine whether there are any differences related to age. In order to do so, the performance of two different age groups (8-9 and 10-11 years old) will be compared searching for potential differences (namely negotiation of meaning, L1 use, and the complexity, accuracy and fluency of the learners' language production). Additionally, we will analyse the influence of procedural task repetition on the above mentioned aspects of foreign language production. In other words, we will examine the influence of age and procedural task repetition on young learners' negotiation of meaning and general competence. Finally, we will inform teaching practices by exploring the potential effects of the repetition of collaborative tasks and what to expect from them. We believe the results obtained will be valuable for pedagogical practice and will help teachers to implement new methods that will offer students more occasions for oral production in the class which, eventually, will lead to increased learning opportunities. All in all, this study intends to contribute to improve the command of a foreign language (English) at primary school level.

The main findings of this study confirm young learners' ability to complete a collaborative task autonomously using the target language. These learners resort to their shared L1 moderately and for reasons that assist them with task completion. Younger children negotiate mostly in order to repair communication breakdowns. However, a change is observed in learners two years older, who show a greater concern about their interlocutor's needs, and use negotiation strategies to confirm that no communication difficulties have occurred. As in previous research addressing this population, comprehension checks and focus on form strategies are rare in the production of either group. The older learners' oral output is more complex, accurate and fluent, and contains fewer L1 terms. Procedural repetition seems to affect the general performance of the two age groups in different ways: only fluency is significantly benefited in the 10-11 year-olds group, whereas both structural complexity and fluency improved in the production

of the 8-9 years old learners. Trade-off effects are identified in the first performances, but these disappear or lessen in the last task. Finally, the use of negotiation of meaning strategies unfolds a similar pattern in the two age groups upon task repetition: most strategies decrease in the last task performance. Altogether, the repetition of the picture placement task has yielded benefits to these learners, as it has provided opportunities to interact in the target language and numerous occasions in which they need to negotiate for meaning. In addition, significant improvements in some dimensions of their general performance are evident.

The current dissertation is structured as follows: in Part I (Literature Review), Chapters 1, 2, 3 and 4 provide the background for the four main issues of relevance to our study. Chapter 1, *The Interaction Hypothesis*, describes the interactionist approach, as well as its main constructs (comprehensible input, feedback and modified output), and the process of negotiation of meaning together with the core conversational adjustments language learners use when engaged in interaction. A section that addresses the issue of L1 use and the main functions it serves in the foreign language classroom is also included in this chapter. Chapter 2, *Tasks*, reviews the Task-based Language Teaching approach, a methodology which takes the construct ‘task’ as the central unit of instruction. This will be followed by an inventory of the main features of pedagogic tasks. In this chapter the task type used in the current study will be discussed in detail. Additionally, among the different task variables that affect task performance, we will focus on the other main pillar of our study: task repetition, followed by a selection of research on this construct. Chapter 3, *Child Second Language Learning*, provides a detailed description of the core characteristics of young language learners, the population object of our study. Taking up the line of investigation presented in Chapter 1, a section is devoted to child interaction in the language classroom to complete the more general view offered in the first chapter, which includes adults and teenagers studies. Chapter 3 also provides a definition of each of the chief dimensions of second language performance (i.e. complexity, accuracy and fluency) analysed in this dissertation, accompanied by the measures used in four research studies which have addressed a similar population and followed a procedure similar to the one in our study. Finally, the differences between second language and foreign language learning will also be highlighted and illustrated with research findings. The last chapter of our literature review, Chapter 4, *Content and Language Integrated Learning (CLIL)*, is devoted to describing this methodology which is becoming prevalent in Europe

and is the teaching approach followed by the school our participants attend. This will be followed by a comparison to the two methodologies considered to be its predecessors (immersion programmes and Content Based Instruction). A selection of research studies addressing CLIL in Europe, and the effects of CLIL on learning, will also be discussed here. Finally, the main research projects on CLIL carried out in Spain will be presented. In Part II, (The present study), Chapter 5, *The study*, describes the methodology we have followed. First, we briefly review our motivation to carry out the current study. Afterwards, the main aims and research questions will be stated, followed by the hypotheses entertained. The next section will provide a description of the participants along with an account of the specific characteristics of the school. Then the procedure and the materials used will be introduced. This part ends by offering a description of the data analysis and codification. Chapter 6, *Results*, presents the results obtained to answer our research questions, which will be discussed in Chapter 7, *Discussion and main findings*, in relation to the Hypotheses posited. Finally, Part III (Conclusions and contributions), consists of one chapter only: Chapter 8, *Conclusions*, which provides the final conclusions in this dissertation and points out to its limitations. Future directions for research on young learners and pedagogical implications will also be suggested.

PART I
LITERATURE REVIEW

CHAPTER 1 THE INTERACTION HYPOTHESIS

The current chapter offers a review of the Interaction Hypothesis (Long, 1983, 1985, 1996), the theoretical perspective adopted in the present study. The first part of the chapter provides a detailed description of the Interaction Hypothesis and its main characteristics. In the second part, the main constructs of the Interaction Hypothesis (comprehensible input, modified output and feedback) are examined. Then, studies discussing the importance of negotiation of meaning (henceforth NoM) for language acquisition¹ will be presented. This will be followed by a description of the strategies that take place during NoM, a process which has been claimed to be essential for language learning by researchers within the interactionist framework.

1.1 The Interaction Hypothesis

Interaction is an essential part of communication and all human social activity. Communication in general, and interaction in particular, is collaborative and most often reciprocal as the participants in the conversation work together in order to create a meaningful exchange. While interaction has always been present in language learning processes, it was not until Long (1983) proposed the Interaction Hypothesis that this human activity started to be considered a truly potential locus for second language (L2) acquisition. Since then, the initial proposal of the Interaction Hypothesis has developed significantly and nowadays it embraces not only the traditional main tenets of interaction, but also a number of factors and processes involved in the course of interaction and, consequently, in L2 learning, for instance learners' internal capacities or the study of new constructs such as the analysis of the language related episodes (henceforth LREs) that take place during interaction.

The Interaction Hypothesis considers conversation as “[...] not only a medium of practice but also the means by which learning takes place, more specifically when it comes to the negotiation of meaning” (Gass, 2007: 234). From the interactionist perspective, language learning can emerge while participating in a conversation which involves sharing and repairing meaning, particularly in face-to-face interaction. Consequently, the opportunity

¹ It is necessary to clarify that Krashen (1982, et passim) made a distinction between the terms ‘acquisition’ and ‘learning’. Language acquisition has been used to refer to an unconscious process, similar to the way we acquire our first language. Language learning, on the other hand, has been claimed to require a conscious effort on the side of the learner. Nevertheless, the terms will be used interchangeably in the present dissertation.

to interact becomes essential to L2 acquisition (Ellis, 2003; Gass and Mackey, 2007; Keck, Iberri-Shea, Tracy-Ventura and Wa-Mbaleka, 2006; Loewen, 2005; Long, 1983, 1985; Mackey and Goo, 2007; McDonough 2005; McDonough and Mackey, 2006; 2008).

Over the last few years, the Interaction Hypothesis has started to be considered an approach to language learning and research, rather than as a hypothesis (Gass and Mackey, 2007; Mackey, 2012; Mackey, Abbuhl and Gass, 2011). This is because of the fact that interaction, although it is seen as facilitative of L2 learning, is not considered to be sufficient on its own, but a framework that can accommodate and support a variety of different processes that trigger language acquisition (Mackey, 2012).

The first version of the Interaction Hypothesis (Long, 1983) was highly influenced by Krashen's (1982) Input Hypothesis, which posits that adult L2 learning is driven by sufficient exposure to comprehensible input, so that the comprehension of language at a slightly higher level than the learners' automatically leads to acquisition. In Krashen's own words "[...] humans acquire language in only one way – by understanding messages, or by receiving comprehensible input" (Krashen, 1985: 2). Krashen's Input Hypothesis is considered to represent the first steps into linking input and acquisition (Mackey, 2007).

Long (1983) observed the conversational adjustments that occurred during interaction in the first language (L1) and realised that these were even more frequent in L2 conversation. Consequently, and in line with Krashen, Long (1983) stated that, since comprehensible input was claimed to lead to language acquisition (Krashen, 1982), the conversational adjustments that take place during interaction and make input more comprehensible will also promote language learning. Later on, Swain (1985) introduced a new perspective that focused on output and added that

“[...] learners need to be pushed to make use of their resources; they need to have their linguistic abilities stretched to the fullest; they need to reflect on their output and consider ways of modifying it to enhance comprehensibility, appropriateness and accuracy” (Swain, 1993: 160-1).

Swain's Output Hypothesis (1985, 1993) posits that only exposure to the target language (TL) is not enough for L2 learning, and that it is learners' production of language what is more likely to promote learning. Learners' output reflects what learners are actually able to produce by drawing on their emerging interlanguage when they are encouraged to

produce more target-like output. This process requires an initial communication problem, which will make the speakers modify their output. Communication breakdowns trigger the process of negotiating for meaning by indicating the speaker that there is a problem with their output which needs to be overcome, this way directing the learners' attention to form (Foster and Ohta, 2005; Pica, 1994), and promoting language acquisition. Consider Example 1.

Example 1

- Learner A:* Boy in mid.
Learner B: **What?** [Clarification request]
Learner A: Boy in mid.
Learner B: **I don't know what you saying?** [Clarification request]
Learner A: Draw a boy.
Learner B: Yeh but where?
Learner A: Not top not bottom. But the in the mid.
Learner B: In the *mid*? Mid oh **middle!** [Recast]
Learner A: Yes, there he flying a kite.
Learner B: Oh what now? Stop so I can draw.
Learner A: **A boy in the middle.** [Modified output]

(Mackey, Kanganas and Oliver, 2007: 306)

In Example 1, the communication breakdown takes place when Learner A's non-target-like utterance is not understood by Learner B, who asks for clarification. After several turns negotiating for meaning, Learner B understands the message and offers Learner A a target-like version of the original utterance, which is acknowledged and reproduced later on.

In Long's (1996) update of the Interaction Hypothesis the joint effect of input, interaction and output, as well as learners' individual capacities in the learning process were brought into focus:

“Negotiation for meaning, and especially negotiation work that triggers interactional adjustments by the NS [native speaker] or more competent interlocutor, facilitates acquisition because it connects input, internal learner capacities, particularly selective attention, and output in productive ways.” (Long, 1996: 451-2)

In a second or foreign language context, interaction quite frequently shows learners the differences between their interlanguage and the target-like form, making them modify what they initially say. Within the Interaction Approach, the general consensus is that interaction facilitates acquisition because it provides learners with opportunities for comprehensible input, output and feedback (Gass and Mackey, 2007; Swain 1985, 1993, 2005; among many others).

Related to the concepts of selective attention and modified output is the term Focus on Form (FonF). FonF “[...] consists of an occasional shift in attention to linguistic code features, [...] triggered by perceived problems with comprehension or production” (Long and Robinson, 1998: 23). NoM positively affects attention to form, which is necessary to get learners to produce more target-like utterances (Pica, 1994; 2013). Form-focused interventions take place when learners modify their output in order to make it more comprehensible, in a way that draws the speakers’ attention to L2 forms (Doughty and Williams, 1998; Long and Robinson, 1998; Pica, 2013). Moreover, FonF refers both to the meaning and the function of the forms being attended to, and therefore, deliberations over the meaning of a word are also included (Ellis, Basturkmen and Loewen, 2002). The following examples show two instances of the shift of attention that can take place in conversation. In Example 2, the teacher draws the learner’s attention to form by repeating the previous utterance and highlighting the errors. Nevertheless, the learner does not repair his/her initial output and the teacher provides the correct version by means of a recast, a form which is later on acknowledged and repeated by the learner.

Example 2

S: I think that the worm will go under the soil.

T: I *think* the worm *will* go under the soil?

S: (no response)

T: I **thought** that the worm **would** go under the soil. [Recast]

S: I **thought** that the worm **would** go under the soil. [Modified output]

(Doughty and Varela, 1998: 124)

Example 3, on the other hand, shows how learners focus their attention on the meaning of a word (*balancing*), unknown to Learner B. By asking for clarification, Learner B gets Learner A to modify his/her output and to produce more comprehensible input.

Example 3

Learner A: Where do I put the girl balancing?

Learner B: What? *Balancing*? What's that?

Learner A: You know . . . **standing on one leg and you not fall down but still standing up so balancing.**

Learner B: Oh! Like here standing on one leg on a horse, like this here on the horse.

Learner A: Yeah—that's where I going now put it.

(Mackey et al., 2007: 286)

To the best of our knowledge, two meta-analysis of research on interaction have been carried out that provide positive evidence of the beneficial relationship between interaction and L2 learning. Keck et al. (2006) analysed the findings of 14 task-based interaction studies published between 1994 and 2003, which investigated the link between task-based interaction and the acquisition of grammatical and lexical L2 features. These authors concluded that treatment groups, which were “[...] exposed to what researchers felt were ideal interaction conditions” (Keck et al., 2006: 113), outperformed both the control and comparison groups, showing that task-based interaction has a positive effect on language acquisition. Furthermore, concurring with Loschky and Bley-Vroman (1993), their results suggest that tasks which require the use of a specific language feature are more effective in promoting acquisition, leading to larger effects over time than those in which a target feature is useful but not essential. Finally, in line with Swain's (1985, 2005) Output Hypothesis, this meta-analysis suggests that opportunities for pushed output

produce larger effects on acquisition. Nevertheless, Keck et al. (2006) recommend considering these findings with great caution as some other variables may have some effect on this aspect too.

The second meta-analysis was carried out by Mackey and Goo (2007), and it is also concerned with the effectiveness of negotiated interaction in L2 learning. Their main results go along the lines of those reported by Keck et al. (2006), and confirm the facilitative role of interaction for L2 acquisition. Mackey and Goo (2007) examined 28 studies which were published between the early 90s and up to 2006, some of them already reviewed in Keck et al. (2006). According to this meta-analysis, interaction has an important beneficial effect on language acquisition, both in the short and the long term. Their results show that, although interaction promotes both lexical and grammatical development, interactional treatments proved to be more effective for lexis in the short-term whilst they seemed to be more beneficial for grammar in the long-term.

Before we move on to review the negotiation strategies speakers might use during conversational interaction, we will present the three main constructs of the Interaction Approach: comprehensible input, modified output and feedback.

1.2 Input, output and feedback

In the last twenty years, interaction research has developed in an extraordinary way. It has expanded from investigating how conversational adjustments (henceforth CAs) promote language acquisition to include a variety of interactional processes that occur in L2 learning. Among the factors and processes accounted for in the Interaction Approach are included not only the foundational three core constructs of the Interaction Hypothesis (comprehensible input, output and corrective feedback), but also aspects such as the social context of learning and learners' internal processes, as well as learners' individual cognitive differences (Mackey, 2012). In what follows, the main aspects of the Interaction Approach will be examined.

1.2.1 Comprehensible input

Building upon Krashen's (1985) proposal, it is widely acknowledged that access and exposure to the TL is necessary for L2 learning. Input has been defined as "the language that is available to a learner through any medium (from listening or reading, for example, or through gestures in the case of signed languages)" (Mackey, 2012: 9). In second

language acquisition (SLA) research, comprehensible input has been recognized as a fundamental component in the language learning process (VanPatten and Williams, 2007). If learners cannot understand the language being addressed to them, they will not be able to use that language to build their own L2 grammar. Negotiated input supplies speakers with linguistic information, as well as extra focus on how specific meanings are encoded in the L2.

However, despite the essential role of input in L2 learning, it is well-known that comprehensible input alone is not sufficient for language acquisition to take place (Mackey, 2012; Pica, 2013). The research carried out in Canadian French immersion schools has demonstrated that learners who received large amounts of comprehensible input had near-native-like comprehension but did not necessarily show near-native-like L2 production skills (Genesee, 1987; Swain, 1985; among others). These findings led Swain to propose the Output Hypothesis (1985) and directed researchers' attention to output as an integral part of the language learning process, which is the next point to be discussed.

1.2.2 Output

The second tenet to be considered here is output, which has also been claimed to be essential by SLA and a key factor of the Interaction Approach (Gass and Mackey, 2007; Mackey, 2012; Smith, 2009). Output can be defined as the spoken or written language forms produced by L2 learners.

Based on the data collected from Canadian immersion programmes, Swain (1985) proposed the Output Hypothesis which states that "the act of producing language (speaking or writing) constitutes, under certain circumstances, part of the process of second language learning" (Swain, 2005: 471). In addition to input, opportunities to actively produce the TL help learners to notice their own errors and will probably direct their attention to the relevant input and feedback from their interlocutors. According to Swain, producing output promotes fluency and automaticity, at the time that opportunities for learners to test their L2 hypotheses are provided. A process directly related to output production is the noticing of needed new forms and language features which, together with opportunities for learners to reflect on their production and think of ways to modify it, will eventually foster the generation of new linguistic knowledge and the consolidation of existing information (Swain, 2005).

Modified output is the type of output that research tends to focus on, and it refers to the result of “[...] the process of rephrasing or reformulating one’s original utterance in response to feedback or self-monitoring” (Mackey, 2012: 16). When modifying their production, learners may notice the gap between their interlanguage and the TL (Schmidt and Frota, 1986), and eventually become more aware of particular grammatical structures. Thus, “[...] the process of modifying one’s output is as important as the ultimate product” (Mackey, 2012: 17).

1.2.3 Feedback

The third construct of interest to the Interaction Approach is feedback, which refers to the information learners receive from their interlocutors about a problem in their language production during interaction (Long, 1996). Numerous studies have provided ample empirical evidence that demonstrates that feedback is beneficial for language acquisition (e.g. the meta-analysis by Keck et al., 2006 and Mackey and Goo, 2007; as well as the studies by Li, 2010; Lyster and Saito, 2010; Russell and Spada, 2006; among others). Feedback plays a key role in interaction and, consequently, in language development, by potentially directing learners’ attention to linguistic problems and promoting the noticing of mismatches between their production and the TL (Gass and Mackey, 2006; Long, 1996; 2007). In Example 4, the learner produces non-target-like output, which is followed by a target-like response on the part of the native speaker (NS). In this case, feedback is noticed by the learner who repeats the correct form in the following turn.

Example 4

Learner: When it happen?
Native Speaker: **When did it happen?** [Recast]
Learner: When did it happen?

(McDonough and Mackey, 2006: 705)

However, and in the same way as the other constructs reviewed above, the presence of feedback alone does not indicate immediate learning. Feedback can be understood in different ways by different learners, and therefore, their attention will not always be successfully directed to problems with form and/or meaning. Thus, errors pointed out by feedback may be noticed and corrected, or not. Besides, even when feedback is noticed, there is no guarantee that learning will be promoted since the developmental level of the learners also plays an important role (Mackey, 2012). Example 5 shows how Learner B

offers corrective feedback in the form of a recast which is not followed by modified output, but simply acknowledged by Learner A. See also Example 12 for more instances in which feedback is not followed by modified output.

Example 5

Learner A: No excuse me [B] did the cow have wool wool?

Learner B: **No does a cow have a tail?**

Learner A: Yeah.

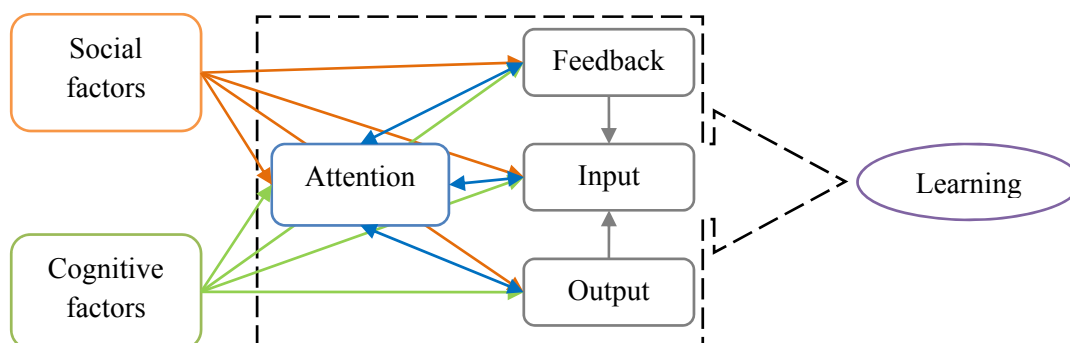
(Philp, Oliver and Mackey, 2006: 548)

Corrective feedback can be explicit or implicit. Even though the explicitness of feedback varies along a continuum, it is generally acknowledged that implicit feedback includes NoM strategies (e.g. in the form of a confirmation check or a clarification request), recasts, as well as “any type of feedback that was not intended to draw the learner’s attention to his/her erroneous production” (Li, 2010: 323). On the other hand, explicit feedback incorporates feedback types that overtly indicate that the learners’ production is not targetlike (e.g. explicit corrections). Implicit feedback is considered the most frequent in conversation whereas explicit feedback is not very common outside the language classroom because it does not follow the rules of politeness as it frequently interrupts the flow of communication (Oliver, 2009). Nevertheless, in the classroom context explicit feedback is considered desirable, as it has a greater potential to help learners notice the mismatches between their production and the target form (Gass and Mackey, 2006).

1.2.4 Other variables

It is important to bear in mind that the above mentioned constructs, considered essential to language learning by the Interaction Approach, are influenced by other aspects such as individual learners’ cognitive differences (e.g. developmental level) as well as by social factors (e.g. educational context). These two external factors are believed to determine the amount of attention learners pay to the input, feedback and output they have access to (Mackey, 1999, 2012; Swain and Lapkin, 1998). Attention is a mechanism that allows learners to administer the information they receive and cannot process at once, mediating therefore between input and learning (Mackey, 2012). In other words, by means of their attentional resources learners may focus on certain parts of input and not on others (Gass, 1997). Figure 1, adapted from Mackey and Polio (2009), illustrates the above described major tenets of the Interaction Hypothesis and the way they interact with each other.

Figure 1 Major tenets of the interaction approach (Adapted from Mackey and Polio, 2009: 5).



One more aspect that influences interaction is the context in which it happens. Although language teaching normally takes place in classrooms with relatively large groups of learners, much of the research on interaction has been carried out in laboratory conditions (see Mackey and Goo (2007) for a review). This fact has been used to question the generalisability of the findings of the Interaction Approach to regular classroom contexts (Foster, 1998; Nunan, 1991). It is obvious that these two settings are different and therefore, the specific characteristics and variables of each context should be taken into consideration.

In a laboratory, learner variables that might affect results (for instance gender, L1, or language proficiency) can be better controlled for than in a regular classroom. Besides, the research instruments used in laboratories can be more freely designed than those to be implemented in a regular language lesson (Mackey, 2012). Even the presence of the researcher is another variable to be accounted for, as learners may pay more attention to form if they feel that their performance is being assessed, and this is sometimes considered to threaten a study's ecological validity. Nevertheless, what most research studies investigate is learners' ultimate attainment, in other words, the level of command learners really have of the TL, which is precisely what show when they feel they are being observed. Moreover, the increasing evidence of the beneficial effects of learner interaction both in laboratory and in classroom settings stands against the claims questioning the applicability of interaction research findings to classroom contexts. Research to date suggests that other variables, such as task type and task familiarity, may have greater influence in the results obtained (Gass, Mackey and Ross-Feldman, 2005; Mackey, 2012).

The last variable to be introduced in this section is that of the interlocutor. Research has shown that different interlocutors influence the type and quantity of interaction. Interaction between native and non-native speakers (Long, 1983; Varonis and Gass, 1985), practitioner/researcher and language learner (Oliver, Philp and Mackey, 2008), as well as learners' individual factors such as gender (Azkarai, 2013; Azkarai and García Mayo, 2012; Ross-Feldman, 2007), age and proficiency (Li, 2010; Lyster and Saito, 2010; Oliver, 1998; 2009) have been investigated. In view of the fact that most interaction in language classrooms takes place between learners, it is important to consider the particular characteristics of learner-learner interaction.

Within the Interaction Approach, peer-peer interaction is considered a very important locus for a meaningful use of the TL, as it gives learners the opportunity to engage in authentic interaction, especially in foreign language (FL) settings where the TL is not as available to learners (Philp and Tognini, 2009; Sato, 2016). Consequently, one of the most beneficial features of tasks is the opportunity they offer for a meaningful exchange in both directions, in which learners interact and improve their knowledge of the TL².

Research has demonstrated that learners provide each other with comprehensible input, feedback and opportunities to negotiate for meaning and, therefore, to produce modified output (Adams, 2007; García Mayo and Pica, 2000; Oliver, 2009; Pica, Lincoln-Porter, Paninos and Linnell, 1996; among many others). Moreover, the literature suggests that the interaction that takes place between learners differs from the one between learners and NS (Mackey, Oliver and Leeman, 2003; Pica et al. 1996). For instance, Mackey et al. (2003) found interesting differences in interaction according to the type of interlocutor. They compared different language background (native/non-native) and age (adult/children) and reported a higher feedback provision by NSs than by non-native speakers, although significant differences were found only in the adult group. On the other hand, even though all pairings types provided opportunities for modified output, it was the adult non-natives who were more likely to offer their interlocutor opportunities to modify their output, whilst the youngest learners actually modified their output the most.

² The ideas fostered by the Interactionist approach have been applied to the language classroom through the use of communicative tasks. Different types and characteristics of pedagogical tasks, as well as their influence on SLA research will be further discussed in Chapter 2.

Additionally, the goal of negotiation also differs between learners and NSs; while NSs normally negotiate either to simplify or expand what they have previously said (Varonis and Gass, 1985), learners when negotiating might also attempt to produce a more target-like utterance or to get their interlocutors to modify theirs (McDonough, 2005; Swain and Lapkin, 2000; 2001). Sato and Lyster (2007) claimed that, although both learner-NS and learner-learner interactions provide negotiation opportunities to a similar extent, learners provide more elicitation feedback than NSs. Furthermore, learners seem to modify their output significantly more when working together, which highlights the benefits of peer interaction. These beneficial aspects of peer-peer interaction have been argued to be related to the idea that learners feel more confident and less threatened during the exchange (Sato and Lyster, 2007). From the socio-cultural framework, this feature of peer interaction is connected to another positive claim: The beneficial effect of the shared responsibility in producing a common outcome, which makes learners focus their attention on the language they use (Donato, 1994; Fernández Dobao, 2014; Swain and Lapkin, 2000, 2001).

Sato (2016) provided further evidence of the benefits of peer-peer interaction by showing how learners compensate for the gaps in their partners' interlanguage, particularly in lexical knowledge. During interaction, learners may provide a TL term their partner needs in a meaningful context, this way allowing them the possibility to expand and/or modify their vocabulary knowledge. Sato (2016) introduced another interesting variable that affects learner interaction: the learners' interaction mindset, which he defined as "a disposition toward the task and/or the interlocutor prior to and/or during the interaction" (Sato, 2016: 7). Sato claimed that learners' interaction mindsets mirror how they interact with each other and, consequently, influence the eventual benefits of interaction on L2 learning. In a similar line of thought, Storch (2016) also stated that it is the quality of the learners' engagement with a task and with language items that creates conditions for language learning rather than interaction per se. In her chapter she referred to collaborative writing tasks, but we believe the idea can be expanded to peer-peer interaction in general.

As stated above, the benefits of interaction become especially noticeable through NoM (which will predictably happen during interaction). By means of NoM, speakers receive feedback on their own performance, as well as opportunities to modify their output. Thus,

in the following section we offer an overview of the process of negotiation which will be followed by a definition of the main strategies that take place during NoM.

1.3 Negotiation of meaning

NoM refers to “[...] turns of talk in which speakers check the clarity and understanding of their own and each others’ messages, particularly at points when there seems to be a breakdown or misfire in communication” (Samuda and Bygate, 2008: 116). These breakdowns are, as mentioned above, essential to promote language acquisition and development. In Example 6, the learner’s initial utterance is not understood by the teacher who, by means of a clarification request (see next section on Conversational Adjustments), makes the student modify her/his pronunciation of the word ‘spotty’ to finally make it target-like.

Example 6

Student: It not sparty?

Teacher: What, sorry I don’t follow? [Clarification request]

Student: It not a sporty, spotty one?

(Mackey, 2012: 41)

The process of negotiation normally ‘pushes’ learners to focus on form. This happens because speakers sometimes need to provide a more target-like version of a previous utterance in order to make themselves understood, increasing input comprehensibility without limiting the access to unknown language features (García Mayo and Pica, 2000; Gass and Mackey, 2007; Long, 1996; Long and Robinson, 1998; Mackey, 2007; Mackey and Goo, 2007; Samuda, 2001; Samuda and Bygate, 2008; among many others). Moreover, during interaction learners may also reflect on their own use of language and engage in meta-talk, in what has been referred to in the literature as LREs (Kim, 2013; Swain and Lapkin, 1995, 1998). LREs include instances of “[...] talk about the language they are producing, question their language use, or correct themselves or others” (Swain and Lapkin, 1998: 326), and are claimed to promote language learning (Adams, 2007; Loewen, 2004). In Example 7 below, Learner 2 identifies the error and overtly points it out.

Example 7

Learner 1: Ok so they call each other.

Learner 2: Oh no no no no just use the past tense or use past.

Learner 1: Oh past tense yeah.

Learner 2: They called each other, ok you can just write called.

(Adams, 2007: 44)

Collaborative tasks, therefore, become crucial to the Interaction Approach since they provide the context and the information learners need in order to relate grammar to the message they want to communicate, that is, by interacting, the grammatical system and the discourse system become connected (Long, 1996). As we shall see in Chapter 2, tasks must be designed to offer possibilities for a meaningful use of the language. In other words, tasks should provide contexts for learners to use their interlanguage to communicate and to solve problems, in a process during which they may identify their own language needs. There is a large body of research that provides support for the use of collaborative tasks that promote interaction. Accordingly, tasks have become a valuable tool for language teaching, which is reflected in the design of new teaching materials (García Mayo, 2007; Mackey, 1999; Van den Branden, Bygate and Norris, 2009)³.

In order to analyse the extent to which tasks promote negotiation, research has focused on the strategies speakers use to face communication breakdowns and overcome their own linguistic deficiencies. As mentioned previously, to solve communication problems learners may offer corrective feedback in the form of NoM or recasts that show mismatches between the received input and the TL, with the aim of getting their partners to modify their output and in so doing making it more comprehensible.

As seen throughout this chapter, numerous empirical studies on interaction provide evidence of the beneficial relationship between interactional processes and language acquisition (Ellis, 2003, Keck et al. 2006; Mackey, 1999, 2007; Mackey and Goo, 2007; Mackey et al., 2011; Pica, 2013; among many others). Most of these studies have focused on communicative tasks and measured outcomes in terms of negotiation strategies, recasts or modified output (Philp et al., 2006). NoM is enhanced in collaborative tasks which, as mentioned earlier, provide a context for attending to problematic forms in which the

³ See Chapter 2 on tasks.

learners themselves can direct their attention to form (Keck et al., 2006; Mackey, 2007). This feature of interaction brings into focus the importance of the interlocutor variable, which has already been discussed in the previous section on the main tenets of the Interaction approach.

In what follows, a detailed description of two of the most frequent forms of implicit feedback used by speakers during interaction will be offered: Negotiation strategies and recasts.

1.4 Negotiation strategies and recasts

Negotiation strategies were originally defined and classified by Long (1983), and closely followed by other authors (Oliver, 1998, 2002, 2009; Pica and Doughty, 1985). We will focus on Oliver's classification, as her studies are the main reference in child L2 interaction, which is the core topic of the present dissertation. Nevertheless, our study addresses young learners (henceforth YLs) of English as a foreign language (EFL) whereas Oliver's work deals with YLs of English as a second language (ESL). This author documented the following strategies which are illustrated with her own examples (Oliver, 2009). In this section, we have also included recasts, as they are also very common during interaction, as it has been shown in different studies addressing this strategy (Ellis, 2003; Lyster and Izquierdo, 2009; Mackey, 2012; Oliver, 2009).

i) Conversational adjustments

These are strategies used in conversation to increase comprehensibility. CAs include the following three types:

Clarification requests are “strategies used by the listener to clarify what the speaker has said, including statements such as ‘I don’t understand’, wh- questions, yes/no questions, and tag questions” (Oliver, 2009: 137). Consider Example 8 below:

Example 8

NS: Just down to her shoulders?

NNS: She – she has – no ...

What? [Clarification request]

What did you say?

(Oliver, 2009: 137)

In this example, the listener does not understand what the speaker, in this case a non-native speaker (NNS) has said and asks for clarification. In other words, the listener wants the speaker to modify their previous output and to produce more comprehensible input.

Confirmation checks are defined by Oliver as “strategies used by the listener to establish that they have correctly heard and understood what has just been said. They often involve repetition accompanied by rising intonation” (Oliver, 2009: 137), as Example 9 shows:

Example 9

NNS1: In a corner.

NNS2: **Corner?** [Confirmation check]

NNS1: Yes.

(Oliver, 2009: 138)

In Example 9 above, the listener wants to confirm whether he or she has understood the previous utterance properly.

Comprehension checks are “strategies, often in the form of a question (e.g. “Do you understand?”), used by the speaker to check that the preceding utterance was understood by the listener. They may also involve self-repetition coupled with rising intonation” (Oliver, 2009: 138):

Example 10

NNS: Up there in the cupboard...

Up there.

You know what’s cupboard is? [Comprehension check]

(Oliver, 2009: 138)

In this case, the speaker is making sure that the listener understands the meaning she or he is trying to convey.

ii) Repetitions

These are the instances in which the speaker repeats a previous utterance (totally, partially or expanding it) within five speaking turns (Pica and Doughty, 1985). Furthermore, repetitions can be classified into:

Self-repetition, which “is undertaken by the speaker and may include partial, exact, and expanded repetitions of lexical items. [...] only deemed to be repetitive if they occur within five speaking turns” (Oliver, 2009: 138). See Example 11 for partial self-repetition, the first part of Example 12 for complete self-repetition and the last sentence of this for expanded self-repetition.

Example 11

NNS: Draw a boy and girl?

Boy and girl

[Partial self-repetition]

(Oliver, 2009: 138)

Example 12

NNS: Cup?

Cup?

[Complete self-repetition]

Cup is go in the left side in the – um ... [Expanded self-repetition]

(Oliver, 2009: 138)

Other-repetition, which “is done when the listener repeats, partially, exactly, or in an expanded form the lexical items used by their partner, again within five speaking turns” (Oliver, 2009: 138). See Example 13 for partial other-repetition and Example 14 for expanded other-repetition.

Example 13

NNS 1: There was a sun.

NNS 2: **Sun?**

[Partial other-repetition]

(Oliver, 2009: 138)

Example 14

- NNS 1: Up the table.
NNS 2: **Up on the eating table.** [Expanded other-repetition]
(Oliver, 2009: 138)

Oliver (1998) also explains that different negotiation strategies may overlap, e.g. a repetition can be used as a confirmation check. This type of utterances are referred to as multifunctional and in her 1998 paper she decided to classify them twice according to each of the functions they served. For instance, a repetition which is used as a confirmation check would be classified both as a repetition and as a confirmation check, as in Example 13. In this respect, Lázaro Ibarrola and Hidalgo (2017a) have recently offered a different perspective, which will be introduced later in this chapter.

iii) Recasts

This form of corrective feedback is “[...] a ‘redisplay’ of the learner’s utterance, where the structure is reformulated but where the central meaning remains unchanged” (Oliver, 2009:140). Recasts are in some cases similar to confirmation checks. See Example 15:

Example 15

- Learner A:* The sun is top of page.
Learner B: **Is at the top?** [Recast]
Learner A: Yes, is at the top.
(Mackey et al., 2007: 286)

In Example 15, Learner B seems to identify Learner A’s previous utterance as non-target-like and recasts it. Learner A appears to notice the mismatch between her/his initial utterance and her/his partner’s response and modifies the original output. The provision of a target-like version after the occurrence of a non-target-like utterance is believed to promote L2 acquisition since it helps learners notice the gap between their interlanguage and the TL (Long, 1996; McDonough and Mackey, 2006).

Although recasts have beneficial effects on language acquisition and are amongst the types of corrective feedback most frequently provided (Lyster and Izquierdo, 2009), their effectiveness at eliciting modified output has been argued to be constrained by their implicitness. Sometimes, recasts are not identified as corrective feedback and simply taken in as alternative forms to the speaker’s initial utterance or just as a follow-up to

their production (Ellis, 2003; Lyster, 1998). In addition to this, on some occasions no opportunity for repair is provided after recasts, which is usually the case when learners are engaged in meaningful interaction (Gurzynski-Weiss and Baralt, 2014; Long, 2007). As can be seen in Example 16, the recast has the form of a confirmation check and the learner, instead of reformulating his/her previous utterance to make it more target-like, focuses on the successful transmission of meaning and just confirms what his or her interlocutor has said.

Example 16

NNS: There is more grass.

[whole over] than more.

NS: **There's more grass than the tree?** [Recast]

NNS: Yep.

(Oliver, 2009: 140)

However, recasts do not always serve the function of a confirmation check and work as corrective feedback, even when there is no understanding problem (Ellis, 2003). In Example 17, the recast provides feedback on the use of the L2 preposition. Nevertheless, as in the previous example, there is not uptake, that is, the NNS does not incorporate the target-like item present in the NS's utterance.

Example 17

NNS: Put in the table.

NS: **On the table.** [Recast]

NNS: And put the knife-.

(Oliver, 2009: 140)

As mentioned above, Lázaro Ibarrola and Hidalgo (2017a) have recently proposed to classify all interaction strategies into four main groups that describe the four main purposes of the interlocutor when negotiating: prevent communication breakdowns, repair communication breakdowns, confirm successful communication and focus on form. This classification helps to understand the main functions that interactional moves serve and allows for each strategy to be classified only once, according to the function it serves rather than to the form it takes (i.e. repetitions, confirmation check). This also allows the inclusion of a more comprehensive set of strategies that might emerge from

different studies instead of limiting the inventory to the conversational adjustments and repetitions. These functions will be further illustrated in the Codification section in Chapter 5.

i) *Strategies to prevent communication breakdowns*, used by the speaker to make sure the interlocutor understands what has been said. This function is typically served by *comprehension checks* and *mere self-repetitions*.

ii) *Strategies to confirm successful communication*, used by the speaker to inform the interlocutor that the previous utterance has been understood. Within this function speakers' *acknowledgements* of understanding and *utterance completions* are the most frequent types of strategies.

iii) *Strategies to repair communication breakdowns*. This function includes those strategies that emerge once a communication breakdown has occurred. They consist of, on the one hand, (i) the strategies used by the listener to show that they have not (totally or partially) understood what the speaker said, which incorporate *clarification requests* and *confirmation checks* (usually in the form of other-repetitions). On the other hand, this function also includes (ii) the strategies used by the speaker as a reaction to the communication failure, mainly by using *self-repetitions*, which could be partially-modified to suit the learners' needs (modified output).

iv) *Strategies to focus on form*. These are used by the speaker to let the interlocutor know that the previous utterance was non-targetlike. The speaker can use either *explicit corrections* or more indirect forms (*corrective recasts*). On these occasions, 'the error might or might not have caused a communication breakdown' (Lázaro Ibarrola and Hidalgo (2017a: 98)).

In the next section, another recurrent strategy that can be found in FL child interaction will be presented: the use of the L1 in the language classroom.

1.5 L1 use

Despite increasing research providing positive evidence of the benefits of balanced L1 use in language learning (Antón and DiCamilla, 1998; DiCamilla and Antón, 2012; Storch and Aldosari, 2010; Swain and Lapkin, 2000), this issue remains a controversial topic among FL teachers. The use of the L1 has been described as a natural tool 'to compensate for lack of linguistic knowledge' (Macaro, 2005:67). The L1 has been found

to serve cognitive and social functions that are facilitative of task completion. These include helping learners with task procedure and providing key vocabulary items which eventually promote communication in the L2 and language learning. However, some practitioners seem reluctant to use communicative tasks in the fear that their learners will resort to their L1 instead of completing the task in the TL (Storch and Aldosari, 2010; Tognini and Oliver, 2012). In FL lessons for learners of the same L1 background it is likely that students resort to their L1 when engaged in communicative tasks (Storch and Aldosari, 2010; Tognini and Oliver, 2012). The facilitative role of the L1 has been reported particularly in communicative tasks and with adult low proficiency learners in FL classrooms (Alegria de la Colina and García Mayo, 2009; Antón and DiCamilla, 1998; Dicamilla and Antón, 2012; Storch and Aldosari, 2010; Swain and Lapkin, 2000; Tognini and Oliver, 2012). Thus, the prevalence of this attitude is especially surprising given that, as previous research has shown, when learners use their L1, they do so to a limited extent and for purposes that facilitate task completion, such as task management or deliberation over vocabulary (Antón and Dicamilla, 1998; Azkarai, 2013; García Mayo and Lázaro Ibarrola, 2015; Lázaro Ibarrola and Azpilicueta Martínez, 2015; Lázaro Ibarrola and Hidalgo, 2017a; Muñoz, 2007a; Storch and Aldosari, 2010; Storch and Wigglesworth, 2003; Swain and Lapkin, 2000).

In a recent study on teenage students' perceptions of L1 use in Cyprus, Neokleous (2016) showed that learners perceived their shared language as beneficial for FL learning. Specifically, the L1 was seen as a valuable resource to solve comprehension difficulties and to cultivate a positive classroom atmosphere. Additionally, the L1 afforded the learners the self-confidence to actively participate in classroom activities.

Several studies claim that L1 use varies as a function of age and proficiency, suggesting that the higher the proficiency, the lower the amount of L1 use (Storch and Aldosari, 2010; Storch and Wigglesworth, 2003; Swain and Lapkin, 2000). However, recent findings in EFL contexts suggest that this connection may be more complex than first assumed and that other variables, such as motivation, task complexity, task repetition, or instructional setting might affect L1 use (García Mayo and Lázaro Ibarrola, 2015; Lázaro Ibarrola and Azpilicueta Martínez, 2015).

Swain and Lapkin (2000) analysed L1 use in the performance of grade 8 learners in a French immersion context while carrying out two different tasks (a dictogloss and a

jigsaw task). Their findings show that L1 use was quite similar across tasks: the learners performing the jigsaw task produced 29% of the turns in the L1 (English) and the learners completing the dictogloss, 21% of the turns. Three main L1 functions were identified:

i) Moving the task along. This category includes sequencing, retrieving semantic information and understanding pieces of information and task management (p. 257). The next example portrays how the L1 is used for task management, as learners discuss how to use the tape recorder:

Example 18

D1: **Should we say it into this now?** [referring to tape recorder]

D2: **What?**

D1: **Just like, right into the thing.**

D2: **No, I don't think so.**

D1: **Let's do it. What's the name of the story again?**

(Swain and Lapkin, 2000: 259)

ii) Focusing attention. Functions such as deliberation over vocabulary and focus on form (i.e. explanations, framing and retrieving grammatical information) fall under this category. Example 19 illustrates a vocabulary search: Learner J1 uses English to ask for an L2 term he or she does not know.

Example 19

J1: Et elle est *tickelée*. **How do you say 'tickled'?**

J2: Chatouillée.

J1: OK. Chatouillée, chatouillée. **How do you say 'foot'?**

J2: Le pied.

J1: Ah, chatouillée les pieds.

(Swain and Lapkin, 2000: 259)

iii) Interpersonal interaction. This last category includes off-task talk and disagreement. According to Swain and Lapkin (2000) off-task talk is frequently done in the L1. In the next example, the L1 is used to suggest a completely unrelated activity.

Example 20

D1: **Wanna do a crossword?**

(Swain and Lapkin, 2000: 260)

In Swain and Lapkin's (2000) study, the learners mainly used their L1 to move the task along (the dictogloss group 35% of the L1 turns and the jigsaw group 43%), and within this category, task management was the most frequent function the participants' mother tongue served. When the L1 was used for focusing attention, the learners resorted to their L1 mostly to deliberate over vocabulary. The least frequent use of the L1 was for interpersonal interaction. The relationship between the amount of L1 use and the quality of students' writing was also addressed in this study. In both tasks, a relationship was found between higher L1 use and lower-rated task outcome, suggesting some sort of "[...] interaction between achievement (as measured by story quality – language and content ratings) and task with respect to the use of the L1" (Swain and Lapkin, 2000: 267). Their results also point at the greater need of lower-achieving students to use the L1. However, the authors acknowledge that the task type also exerts an influence on L1 use.

In an ESL context, Storch and Wigglesworth (2003) conducted a study with 12 pairs of age- and level-matched university students. Six of the pairs shared their L1 while the other six did not. The authors also investigated the impact of task type on L1 use and found that most of the participants (except for two pairs) kept their L1 use to a minimum. Concurring with Swain and Lapkin (2000), their results showed that L1 use differed across tasks: whereas in the joint composition favoured the L1 mainly for task management and clarification, in the reconstruction task, the main function served by the L1 was for NoM and vocabulary. Additionally, the low L1 use was related to the learners' attitude in the language classroom. Most learners were reluctant to extensively use their L1 in a second language (henceforth SL) setting because of two main reasons: they believed it could slow the pace of the class and also felt that they had to maximize their opportunities to use the TL. On the other hand, they were aware of the benefits the L1 provided when carrying out the tasks (e.g. for vocabulary deliberation, and/or to clarify some task-related).

In an EFL context, another classification of L1 functions was offered by Muñoz (2007a). She analysed the oral production of Catalan-Spanish bilingual learners, focusing specifically on cross-linguistic influence as well as on language switches. In her study, two main categories were identified: lexical transfers and code-switching.

i) Code-switching, which is defined as "a complete shift to another language for a clause or a long expression" (Muñoz, 2007a: 81), is sub-categorised into three types: explicit appeals for help, clarification requests and meta-comments.

- Appeals for help usually include using the L1 (or any other previous language) to ask about an unknown TL term. Consider example 21 below in which two different languages serve different functions. First the speaker employs a foreignising from French (*perruchet*) to overcome the lack of that specific TL term. Then continues with an appeal for assistance in Spanish in order to get the TL term from the interlocutor, which is finally followed by a borrowing from Spanish to avoid a communication breakdown.

Example 21

And I'm a **perruchet** (Fr.) ay **como se llama** hmm **bueno periquito** (Sp.).

[I have a parakeet oh what's it called mm well parakeet (in Spanish)]

(Muñoz, 2007a: 84)

- Clarification requests which, as already defined in the section devoted to negotiation strategies, are utterances in which the listener asks the speaker to clarify a preceding intervention. In the case of code-switching, only those instances provided in the L1 are to be considered, as in Example 22, in which a clarification request in Catalan is used.

Example 22

R: And what are they going to do now?

L: **Cóm** (Cat.)? [What?]

(Muñoz, 2007a: 82)

- Meta-comments “consist of comments on the communicative situation or on the learner’s inability to complete the task in English” (p. 82). Example 23 illustrates a meta-comment in which the learner expresses in Catalan his or her doubts towards the task procedure.

Example 23

Cut the ... **què he de fer? És que estic dient coses pero no ...** (Cat.)

[Cut the ... what do I have to do? I am saying things but I don't ...]

(Muñoz, 2007a: 82)

ii) Lexical transfers are described as “the use of a word from another language”. Within this category, two sub-categories are identified: borrowings and foreignisings.

- Borrowings are the use of an L1 term (or from any other previous language) in the TL discourse. In example 24, the speaker decides on using a Spanish term instead of one from the TL.

Example 24

Look the .. **la la cesta** (Sp.).

[Look in the basket]

(Muñoz, 2007a: 87)

- Foreignisings are phonological or morphological adaptations of an L1 term to the rules of the TL, as in example 25, in which the learners opts for shortening an L1 term ‘cesta’ to make it sound more similar to other TL words.

Example 25

When she open the the **cest** (Sp.)

[When she opens the basket]

(Muñoz, 2007a: 82)

In line with previous findings, Storch and Aldosari (2010) reported a modest L1 use in a study also carried out in an EFL setting (7% of the total amount of words were L1 terms and 16% out of the total number of turns contained L1 words). The participants, 15 proficiency-matched pairs of college students, performed three tasks (jigsaw, composition and text-editing). These authors identified five L1 functions: task management, discussing and generating ideas, and grammar, vocabulary and mechanics deliberations. Supporting the findings reported by Storch and Wigglesworth (2003), the authors identified a more extensive use of the L1 when dealing with more difficult tasks (in this case the editing task), mainly among low proficiency students. The L1 was mostly used for task management (41% of all L1 turns) and to deliberate over vocabulary (26%). These authors also suggest that the higher the proficiency the fewer the number of instances of L1 use. Low L1 users remained so regardless of the task type, whereas moderate or extensive L1 users tended to rely on the L1 more frequently in the text-editing tasks than in the other two types.

More recently, Azkarai and García Mayo (2015) studied the influence of task modality on the use and functions of the L1 in the oral interaction of 44 EFL Spanish university students. They distinguished five main L1 functions and, following Alegría de la Colina

and García Mayo (2009) and Storch and Aldosari (2010), elaborated the following categorization:

i) Off-task, which happens when the speakers use the L1 to talk about a topic, unrelated to the task. In Muñoz's (2007a) classification, this function was not observed but, according to her definitions, it would be considered within the code-switching categories. See example 26, which illustrates how one of the participants asks the other about a friend:

Example 26

Antonio: [...] And can make sharing a house either, either, either a great experience or a nightmare. **¿Qué sabes de Paloma?** [Have you heard from Paloma recently?]

Julio: **Pues la vi hace poco.** [I saw her recently.]

(Azkarai and García Mayo, 2015: 557)

ii) Metacognitive talk refers to the instances in which the speaker uses the L1 to talk about the task itself. This category would correspond to Muñoz's 'metacomments', in which task planning and task management are included. Example 27 illustrates this L1 function, as one of the participants asks the other whether he wants to be the one who writes in the dictogloss task, and the other one (Antonio) ignores the question and continues with the task, although he also falls back on his L1 when he does not know how to continue.

Example 27

Julio: Ok, the painting ... **¿Quieres escribir?** [Do you want to write?]

Antonio: The painting we are looking at now or **no sé ... ¿Cómo lo... ?** [I don't know... How do you ...?]

(Azkarai and García Mayo, 2015: 557)

iii) Grammar talk, which refers to using the L1 to talk about grammar, as in example 28, in which Julian explains the grammar rule for their word choice. In Muñoz's (2007a) classification this function is explicitly mentioned, but it would be considered within the code-switching categories, as meta-comment.

Example 28

Rosa: I think it's going.

Julian: Going, going! **Porque es su ... sujeto de la oración.** [Because it is ... the subject of the sentence.]

(Azkarai and García Mayo, 2015: 557)

iv) Vocabulary. Within this use, Muñoz's (2007a) borrowings as well as foreignisings would be included, as in this category the L1 is "[...] used in deliberations over word/sentence meaning, word searches and word choice" (Azkarai and García Mayo, 2015: 558). In example 29, neither of the participants knows the TL form they need, but by means of using the L1, they are able to continue with the task:

Example 29

Gema: [...] The towel is eh ... **¿Colgado?** [Hanging?]

Anita: Yes, **colgado**, [hanging] yes. Ah!

(Azkarai and García Mayo, 2015: 558)

v) Phatics are "[...] expressions to establish social contact and to express sociability rather than specific meaning" (Azkarai and García Mayo, 2015: 558). For instance expressions such as 'ok', 'well' or 'right' (example 30).

Example 30

Santiago: Ok. So, we have to write.

Virginia: To rewrite, yes. **Bueno** [Well], one. You?

The participants in this study had to complete three collaborative tasks (namely a picture differences task, a picture placement task, a dictogloss and a text editing task). Concurring with the findings of previous studies, L1 use was limited, as only 15.41% of the total amount of turns contained L1 terms. Moreover, their results suggest that task type influences L1 use. Similar to Storch and Aldosari (2010), Azkarai and García Mayo (2015) found that there is more L1 use when learners engage in collaborative speaking+writing than in speaking-only tasks.

In spite of the literature giving attention to L1 use in collaborative tasks, a comparatively small body of L2 interaction research has examined the use young FL learners make of the L1. In what follows, seven studies will be reviewed which we have deemed as the

most relevant to our study because of the population they address (Ys) and the aspect they deal with (learner-learner interaction in a FL setting):

Tognini and Oliver (2012) examined learner-learner interaction, their findings corroborating the beneficial effects of L1 use in FL contexts. The participants were primary and secondary education learners of different FLs (French and Italian) in Australia. The findings of their study showed that the learners preferred to use their L1 for management purposes (e.g., task clarification) while the TL was mainly used during form-focused exchanges (e.g., drills). In contrast, content-focused exchanges (e.g., answering questions about a text) were carried out in a mixture of the two languages. The findings were attributed to the different kind of demands each type of exchange placed on the students. According to the authors, as opposed to task management or content discussion, form-focused exchanges are predictable and require a limited command of the language, making learners feel more confident to use the L2. The authors conclude that the L1 was successfully used by learners for scaffolding and to solve difficulties with the L2.

More recently, also in a FL setting, learner-learner interaction has also been examined by Lázaro Ibarrola and Azpilicueta Martínez (2015), who have provided further support to the claims of the scarce and wise use of the L1 by language learners. The participants were 16 young EFL learners (aged 8-9) with a very low level of proficiency in the TL (English). Despite the learners' difficulties in communicating in the TL, they avoided using the L1 throughout the task (a guessing game), with only 5 instances of L1 reported which represent 0.52% of their total production. This was attributed to the participants' high levels of motivation and to the fact that they, thanks to the teaching methodology followed in the school, were used to interacting in the TL in the classroom.

In a similar context, García Mayo and Lázaro Ibarrola (2015) contrasted the oral performance of CLIL⁴ and mainstream EFL (henceforth MS) Ys (ages 8 and 11) in a primary school in Spain. Ys in CLIL programmes were found to rely on their L1 less frequently than those in traditional EFL classes (for instance, CLIL 8-year-old learners' L1 use was reported to represent 1.6% of the total sample and non-CLIL 8-year-olds, 3.6%). When comparing the two different age groups, the older children were found to resort to their L1 more frequently than their younger counterparts in the two language

⁴ See Chapter 5 on CLIL in which this teaching approach is analysed in depth.

programmes, despite their higher command of the TL (CLIL 11-year-olds resorted to their L1 what constituted 4.3% of the total production whereas the L1 use of the non-CLIL 11-years-olds constituted 9.2%). These findings were related to a potential lower level of motivation to complete the task on the part of the older learners which would lead to greater reliance on the L1. These results suggest a more complex connection between L1 use and language proficiency than the one reported by Storch and Aldosari (2010), in which factors such as motivation and task complexity may play an important role.

The effect of task repetition (procedural and exact same task) on L1 use by young EFL (age 9-10) learners has only been addressed by Azkarai and García Mayo (2016). L1 use decreased significantly in the second performance, regardless of the type of repetition. Moreover, the exact same repetition condition seems to elicit significantly more frequent L1 use than procedural repetition. Appeals for help and borrowings were the functions most commonly served by the L1 at the two testing times, under the two conditions. In the first task performance phatics were the third most common L1 function. In the second task however, the frequency of use of some L1 functions significantly changed: under the exact repetition condition the use of L1 phatics significantly decreased whereas confirmation checks and metacognitive talk increased. The procedural repetition group also employed more phatics at time 1 than at time 2 and, unlike under the exact repetition condition, the function that decreased significantly was L1 to communicate lack of knowledge.

Finally, Lázaro Ibarrola and Hidalgo (2017a) also studied the performance of 40 young EFL learners (11 years old) when carrying out a communicative task in a CLIL school. Even though a more frequent L1 use (10.49%) was reported in this study, it can still be considered low, providing further support to previous findings. The instances of L1 use found consisted mainly of unknown L1 terms and task management. However, after a deeper analysis of the participants' utterances, a large proportion of L1 (Spanish) structural transfer was identified (including questions without inversion, null subjects and nouns followed by adjectives, among others). The instances of structural transfer resulted in non-target-like output that remained unnoticed and was therefore not corrected by the learners. These findings uncover the need to focus research not only on instances of lexical transfer but also on the transfer of L1 structures.

As far as we know, only three studies have considered the impact of learning context (MS vs. CLIL) on L1 use when adopting a longitudinal perspective: García Mayo and Imaz Agirre (2017), García Mayo and Hidalgo (2017), and Pladevall-Ballester and Vraciu (2017).

García Mayo and Imaz Agirre (2017) analysed the oral performance of 27 dyads of YLs (ages 8-12) while they carried out a communicative task twice in two school years, and examined the learners' use of CAs, among which L1 use was included. Mirroring previous studies, their findings show that the learning context had a great impact on L1 use, with MS learners displaying a greater reliance on the L1 than CLIL learners at the two data collection times. Nevertheless, the amount of L1 use among the MS learners decreased which was not observed among the CLIL learners. In other words, the CLIL learners' L1 use appears to have remained stable across time. The authors claimed that this finding could have been due to the nature of the task, which may have not been sufficiently motivating for this group of learners.

García Mayo and Hidalgo (2017) also analysed the L1 use (and the functions it serves) of two groups of young EFL learners (n= 32) when performing an oral communicative task two times in two school years. The participants (age 7-10) attended two state schools in Spain, one that followed a CLIL programme and the other which taught the TL in a MS approach. This study provides further evidence of the limited use of the L1 YLs make when engaged in a collaborative task in the TL. Additionally, the results reveal a significantly higher L1 use by the MS group at the two data collection points. Moreover, corroborating previous research, the L1 is mainly used with functions that facilitate task completion, such as deliberation over vocabulary and task management. As a final point, the authors report a significant increase in the L1 use of the MS group at the second testing time. The study concludes by emphasizing the facilitative role of the L1 for FL learning, as well as the absence of an excessive use during TL interaction.

The third longitudinal study that has addressed YLs' L1 use in these two learning context is Pladevall-Ballester and Vraciu (2017). These authors examined the L1 use of 74 Catalan/Spanish bilingual learners (aged 9-12) who were asked to complete a picture-based narrative task. They collected data at four time points and controlled for the amount of L2 exposure the learners received. Thus, in order for the MS learners to have the same amount of hours of exposure (411.25 hours) as the CLIL participants, data collection

started one year earlier in the MS group. The study showed that learners in both groups produced a significantly lower number of L1 words at the last data collection point, whereas the total number of words and the total number of English words increased. On the other hand, no statistically significant differences in the amount of L1 use were observed between the two groups. Pladevall-Ballester and Vraciu attribute these findings, which contradict previous research claiming lower L1 use by CLIL learners, to the equal amount of TL exposure received by all their participants. They conclude by stating that few differences are identified in the two contexts as the participants in both groups make a limited use of the L1 which changes according to task. Finally, the L1 is used by their participants as a compensatory strategy.

Overall, the existing studies suggest that language learners do not make an abundant L1 use when performing L2 activities. When different teaching approaches have been compared (CLIL and MS), in most studies CLIL has been shown to be the setting in which learners use the L1 the least. However, the most recent study reviewed in the current section (Pladevall-Ballester and Vraciu, 2017), which also stands as the only experiment that has controlled for the number of hours of exposure to the L2, points in a different direction, as no statistical differences were found between the two educational approaches (but see the points raised by García Mayo and Hidalgo (2017)⁵). When different ages and proficiency levels are compared, the results are still inconclusive and point at the need to resort to extralinguistic reasons to predict the amount of L1 use. Moreover, when used, the L1 has been deemed a valuable tool, aiding students in task completion and language development. However, in spite of the limited amount of L1 terms found in the studies, some researchers (García Mayo and Hidalgo, 2017; Lázaro Ibarrola and Hidalgo, 2017a) have reported numerous instances of L1 structural transfers in the learners' L2 production. Therefore, we can conclude that more research is needed to determine the actual effect of the L1 on the learners' TL production, not only on the use of L1 terms, but also of the linguistic structures.

⁵ García Mayo and Hidalgo (2017) argue that the different results may be due to the different task type Pladevall-Ballester and Vraciu (2017) employed (a non-collaborative narrative task vs. a two-way communicative task) and to the data collection arrangement (MS learners starting one year earlier than the CLIL group).

1.6 Limitations to the role of interaction in SLA

In spite of the reported benefits of interaction, several studies have highlighted that the interactional tenets (comprehensible input, feedback and modified output) may produce heterogeneous effects as a function of different variables, such as learners' individual differences, learning context, task type and language area. Moreover, each of the interaction components can produce both separate effects, or work in combination, depending on external variables (Mackey, 2007). Regarding modified output, for instance, it is well known that it does not always happen in interaction, as sometimes speakers are able to understand each other without producing a grammatically correct sentence, even more so when the learners share the L1.

Similarly, within the NoM process, the role of communication strategies is not as straightforward (Ellis, 2003). These strategies can be seen as a tool for understanding L2 communication rather than for having an effect on language acquisition (Ellis, 2003). Several authors have claimed that the interactive processes that may lead to language development are broader than those typically suggested in the interactionist literature (Foster and Ohta, 2005; Gagné and Parks, 2013). Foster and Ohta (2005) stated that: “[...] interactional processes including negotiation for meaning and various kinds of peer assistance and repair are among the many ways learners gain access to the language being learned” (p. 426). Findings reveal that, although learners do not use the strategies normally associated with NoM (i.e. comprehension checks, confirmation checks and clarification checks) as frequently as we would expect, scaffolding and assistance to each other is indeed provided through the use of other strategies such as co-construction, other-correction, self-correction and encouragements to continue (see Lázaro Ibarrola and Hidalgo (2017a)).

As stated at the beginning of this chapter, current work suggests that interaction is necessary although not sufficient for L2 learning and it is now seen as a framework that accommodates different approaches to language acquisition (Mackey, 2012). As mentioned above, some studies hint at the possibility that not all interaction is facilitative of L2 learning, as interaction may also occur without true understanding and therefore, without the possibility for learning, or with true understanding but without target-like output (Ellis, 2003; Spada and Lightbown, 2008). In Example 31 below, NNS merely repeats their previous utterance in response to NS's lack of understanding. Even though

this mere self-repetition does not include any modification of the original output, thus remaining untarget-like, this time NS acknowledges their understanding by repeating part of NNS's utterance.

Example 31

NNS: I go to the cinema.

NS: Uh?

NNS: **I go to the cinema last night.**

NS: **Oh, last night.**

(Ellis, 2003: 81)

Another aspect related to the possible limitations of NoM is the fact that attention to form may be incidental, that is, it depends to a large extent on learner and interlocutor individual variables. Feedback, therefore, may be perceived in different ways and it will be noticed or not by learners, depending for instance on the type or the target of the feedback (Mackey, Gass and McDonough, 2000). Thus, the potential of feedback in negotiation is determined to a large extent by the speakers (Oliver, 2009).

These issues bring into focus the relationship between communicative effectiveness and language acquisition. Claims have been made about how they complement or compete with each other. According to some authors, there is some kind of trade-off effect between them, which would explain why it is so difficult for learners to concentrate on different aspects of the language at the same time when engaged in conversation (Foster and Skehan, 1996; Housen and Kuiken, 2009; Skehan, 1996, 1998; Wong, 2001). In stark contrast, other authors consider it to be a two-way relationship, i.e. the more negotiation takes places, the more language learners acquire and the more communicatively effective they become. Consequently, more opportunities for language acquisition will be provided (Ellis, 2003) (this aspect will be further discussed in Chapter 3). All these aspects should be taken into account when designing new studies and interpreting their findings.

1.7 Conclusion

This chapter has offered a review of the Interaction Approach and its main tenets and processes have been described. The ways in which interaction provides a rich resource for L2 acquisition have been highlighted. In spite of the strong support for the benefits of interaction, there are some limitations that need to be considered when interpreting the

findings reported by interactionists, and which leave an open door for future research in this field.

Either way, interaction is a key to L2 acquisition, and the communication of some specific meaning is what triggers the process, rather than a concern for language accuracy (Long, 1985). As shown in this review, research has demonstrated that interaction facilitates language learning by raising awareness of language forms (Gass, 1997; Schmidt, 1995; 2001), creating opportunities for learners to modify their non-targetlike utterances (Swain, 2005) and providing interactional feedback (Mackey and Goo, 2007).

Communicative tasks are a valuable tool to start the process of NoM, and have become a bridge between classroom methodology and interaction research (Bygate, Skehan and Swain, 2001; Ellis, 2003). In the following chapter, a detailed description of the concept of pedagogical task together with some of the most significant studies on the field will be presented.

CHAPTER 2 TASKS

This chapter offers a broad overview of the concept of pedagogic tasks in SLA. First, an introduction to the task-based language teaching (henceforth TBLT) approach will be provided. Secondly, some of the most widely accepted definitions of the construct ‘task’ will be offered and analysed. The main features of tasks will be presented, which will define what will be considered as ‘task’ throughout this dissertation. Taking into consideration different views on task features, various classifications will be described. Afterwards, we will focus on a particular variable which will be one of the main pillars of the current dissertation: the effect of task repetition on task performance, including its effects on general competence and NoM. Finally, after this introduction, our focus will be narrowed to the definition of the particular task we have used in our study: the picture placement task.

2.1 Task-based language teaching

There is ample evidence that tasks are considered very efficient tools for SLA (García Mayo, 2007; Leaver and Willis, 2004; Mackey et al., 2007; Nunan, 2004; Pica, 2005; Van den Branden, 2006; Van den Braden, Bygate and Norris, 2009). Accordingly, new methodologies, such as TBLT, are becoming commonplace in the western language classroom. As we will review in this section, research to date has provided reasons to believe that TBLT is a very valid approach to second and foreign teaching in a formal context. According to Van den Branden et al. (2009: 11):

[...] tasks, potentially at least, offer a uniquely powerful resource both for teaching and testing of language. In particular, they provide a locus for bringing together the various dimensions of language, social context and the mental processes of individual learners that are the key to language learning. There are theoretical grounds, and empirical evidence, for believing that tasks might be able to offer all the affordances needed for successful instructed language development, whoever the learners might be, and whatever the context.

TBLT is a teaching approach in which “tasks are the central unit of instruction: they ‘drive’ classroom activity, they define the curriculum and syllabuses and they determine modes of assessment” (Samuda and Bygate, 2008: 58). Within this methodology, tasks are considered to be essential for the language learning process as the linguistic elements

addressed in the language classroom emerge from them. Consequently, tasks are used to define the syllabus to be followed in class, as well as to assess learners' language acquisition process, which is measured in terms of task performance. Another important characteristic of TBLT is that the tasks are selected to replicate relevant and meaningful real-world situations, that is, L2 learning is promoted by a meaningful use of the language.

The increasing number of studies carried out in authentic classroom contexts, instead of under laboratory conditions, has had an important effect on the value and consideration of TBLT methodologies. These studies have contributed to shedding light on the validity of the use of tasks in language classrooms in which learners with different backgrounds and proficiency levels work together. Moreover, the possibility to design and implement a task-based syllabus that covers all of these differences has also been studied (Van den Branden, 2006). This approach is particularly linked to research focusing on how collaborative tasks can increase the opportunities for feedback and FonF (Long, 1985; Pica and Doughty, 1985; Pica, Kang and Sauro, 2006; Samuda and Bygate, 2008).

Based on the large body of research that supports the use of FonF as a very important or even essential component for interlanguage development (Doughty and Williams, 1998; Long and Robinson, 1998; Mackey et al., 2007; Skehan, Bei, Li and Wang, 2012), many pedagogic tasks are designed to promote it. FonF can facilitate the achievement of higher levels of accuracy (Ellis, 2003; Leaver and Willis, 2004) since explicit feedback helps learners recognize language forms that otherwise would have gone unnoticed (Schmidt, 1990). Later in the following section on task definitions we return to this issue.

Shehadeh (2012) pointed out the fact that much of the existing literature deals with TBLT in SL. This author argues that it is important to take into consideration the differences between SL and FL teaching and learning as the specific features of each context are likely to influence learning outcomes. This point will be discussed below in the section 'The importance of context in language learning: ESL vs. EFL'.

Still, TBLT also has a number of detractors who argue that research findings on TBLT are limited and cannot be transferred to real classroom conditions and that more evidence for the generalisability of the findings is needed (e.g. Bruton, 2002; Swan 2005).

Apart from the controversy over the pedagogical validity of TBLT, this approach has to face other challenges that hinder its implementation. There are cases in which

governments and educational authorities support the implementation of TBLT and still traditional, language and teacher-centred instruction persists. In some contexts, when TBLT is implemented, it is done in an unsystematic way, and therefore its impact on the classroom is not as noticeable as it should be (Shehadeh, 2012). Thus, it could be argued that a successful implementation of TBLT does not only depend on institutional variables but also, and perhaps more importantly, on the teachers' familiarity with this approach, their beliefs and their relationships with the students (Carless, 2004; 2012; Van den Branden, 2006), as well as on the acceptance on the part of the students to work with this methodology (Iwashita and Li, 2012). Adams and Newton (2009) classified the different variables which may hold back the implementation of TBLT into institutional, teacher and student factors.

- Institutional factors: Class size, official exams format, materials, mixed-level groups and lock-step type curriculum, with a heavy reliance on grammar, are included among these factors. Another important issue is the reluctance of part of the society to the groupings and activity types required for TBLT, which are not always considered as 'good classroom management' (Carless, 2004).
- Teacher factors: These include teachers' beliefs and theories of how a language should be taught, their own teaching experience and their interactive skills, as well as the feeling of being more in control of what happens in the class by following traditional methods (Iwashita and Li, 2012; McAllister, Nancy-Combes, Starkey-Perret, 2012) or, on the other hand, just not knowing how to work with this new approach (Chacón, 2012; Chan, 2012). Often, the successful implementation of TBLT largely depends on the teacher, while institutional factors might play a secondary role. If the "teachers lack the skills or motivation to work with tasks (and the basic belief that task-based interaction fosters language learning), no real change will take place" (Shehadeh and Coombe, 2012: XII).
- Student factors: In many cases students do not believe in the effectiveness of TBLT, in line with their parents and teachers' traditionalist ideas about education. These aspects are similar to the teacher factors, but seen from the learner's perspective, as they include the learners' own beliefs about learning, their preferences for traditional methodologies and their 'level of assertiveness'.

All these constraints however are not specific to TBLT but could also appear when trying to implement any new approach, since they come from personal views of what an appropriate methodology is.

2.2 What is a task?

Numerous definitions of tasks have been provided by different authors and theoretical approaches. In this section, we will present some of the most popular definitions and consider the common aspects they all share to try to offer a more complete view of this term.

One of the first definitions offered from a research-based perspective, is the one by Bygate (1999a) in which he underlines the role of tasks as a tool for language development, defining them as “bounded classroom activities in which learners use language communicatively to achieve an outcome, with the overall purpose of learning language” (Bygate, 1999a: 186). In this definition, the adjective ‘bounded’ indicates that tasks are delimited by a starting point (which is the input or instructions, as well as the materials) and an end (the outcome). The term ‘outcome’ here can be interpreted both as the actual communicative use of the TL, or as the achievement of the goal of the task, which can be either task completion or the development of the learners’ interlanguage (Samuda and Bygate, 2008).

In line with Bygate’s definition, Ellis (2003) also states that tasks are directed towards achieving an outcome, but he states that it may not be the actual completion of the task, but the use of the TL that will eventually lead to language acquisition. Thus, the main goal is to elicit language use. Ellis highlights the fact that tasks promote language use that resembles real-world situations, specifying that the use of several language skills and of various cognitive processes is very likely.

A task is a workplan that requires learners to process language pragmatically in order to achieve an outcome that can be evaluated in terms of whether the correct or appropriate propositional content has been conveyed. To this end, it requires them to give primary attention to meaning and to make use of their own linguistic resources, although the design of the task may predispose them to choose particular forms. A task is intended to result in language use that bears resemblance, direct or indirect, to the way language is used in the real world. Like other language activities, a task can engage productive or receptive, and oral or written skills and also various cognitive processes (Ellis, 2003: 16).

Ellis underlines the ‘interactional authenticity’ of tasks when he refers to real-world situations, that is, this type of activity needs to elicit language use that bears a resemblance to situations that happen outside the classroom (e.g. personal information exchange, problem solving or collective judgements). In contrast, Samuda and Bygate (2008) downgrade the importance of the real life component and state the importance of tasks being designed as ‘structured learning situations’, and the fact that they primarily need to promote language learning. This is significant because, as they claim, in real-world situations learning opportunities are not frequent, nor the final goal of linguistic exchanges is that of language learning.

Ellis’s (2003) choice of the term ‘workplan’ has also been criticised on the grounds that it implies that a task is basically an intention, a design to be carried out, and does not contemplate what learners may actually do (Samuda and Bygate, 2008; Seedhouse, 2005). Samuda and Bygate (2008) explain that a working definition for L2 pedagogic tasks needs to refer not only to the theoretical part of how a task should work, but also to how learners may interpret it and, particularly, to the process it entails:

A holistic activity which engages language use in order to achieve some non-linguistic outcome while meeting a linguistic challenge, with the overall aim of promoting language learning, through process or product or both (Samuda and Bygate, 2008: 69).

By proposing the term ‘linguistic challenge’, these authors underline that the main aim of pedagogic tasks is to enhance language development, making Ellis’s (2003) reference to tasks engaging ‘various cognitive processes’ more concrete. Achieving a meaningful outcome is essential, together with the process and the linguistic decisions that leads to

it, that is, the language necessary to achieve the ‘product’ or ‘outcome target’. These authors state that in order to analyse a task, it is necessary to take into account the relationships between the learners, the task and the TL, as well as the context in which it is employed. Besides, the ‘input material’ (task instructions, materials, etc.) and the different stages of a task are also essential because the actual process of carrying it out, and eventually language acquisition, develop from them. It is also important for researchers to know what the target linguistic aspects are, as well as to consider the conditions in which the task is performed, such as the time available to perform it, students’ motivation, proficiency, etc. Finally, Samuda and Bygate (2008) argue that any change in design (instruction, input, conditions, process or product) will affect learners’ language acquisition and use.

Although the main elements of tasks are quite clearly defined in the literature, there is still some controversy on what their main goal should be. While some authors state that the principal objective of tasks is to engage learners in communicative processing of language (Ellis, 2003; Leaver and Willis, 2004; Willis, 1996), others argue that tasks should reinforce or enhance certain linguistic features, and that in order to facilitate language development and acquisition, some FonF is necessary (García Mayo, 2011; Hawkes, 2012; Long and Robinson, 1998; Pica et al., 2006; Skehan et al., 2012).

Overall, the most common belief about what the goal of a task should be is a combination of two traditionally opposite views: ‘Focus on forms’ and ‘Focus on meaning’, following Long and Robinson’s (1998) terminology. Extensive research has shown that neither a mere focus on meaning and mere exposure to the L2 nor formS-based instruction are enough for learners to reach proficiency in the TL and to develop their productive skills (Doughty and Williams, 1998; Spada, 2011). Thus, a different perspective is needed. Long and Robinson (1998) proposed the FonF approach which “consists of an occasional shift of attention to linguistic code features, triggered by perceived problems with comprehension or production” (Long and Robinson, 1998: 23). According to Skehan et al. (2012), pedagogic tasks enhance the development of general communication strategies, but if they include some FonF, they will also favour the development of learners’ interlanguage system.

Researchers and practitioners should “create tasks that provide learners with opportunities to engage in meaningful interaction and to direct their attention to linguistic form”

(McDonough and Mackey, 2000:83). As Pica (2005) stated, although “many language skills can be learned through a focus on meaning, there is increasing evidence that the learning is incomplete and that grammatical imprecisions remain” (Pica, 2005: 342). Nunan (2004) emphasizes the use of grammatical forms to express meaning, underlining the close relation between meaning and form, by stating that “grammar exists to enable the language users to express different communicative meanings” (Nunan, 2004: 4). Accordingly, most of the literature to date suggests that tasks can be designed to target particular aspects of language, directing learners’ attention to specific language features (Bygate, 1999a; Dufficy, 2004; Ellis, 2005; Mackey, 1999; Samuda, 2001).

In conclusion, at the core of the existing definitions of pedagogic tasks, there is an emphasis on the following key elements:

- (i) tasks have the overall purpose of enhancing language learning (Nunan, 1991, 2004);
- (ii) they are meaningful, that is, a task must be goal oriented and have a clearly defined, non-linguistic outcome (Bygate et al., 2001; Ellis, 2003; Long and Robinson, 1998; Pica, 1993; Skehan, 1998);
- (iii) participants interact to achieve a goal (Bygate, 1999b; García Mayo, 2007; Mackey et al., 2007) and
- (iv) there is some focus on form (on the language) (Hawkes, 2012; Long and Robinson, 1998; Skehan, 2012).

All in all, tasks are beneficial for second or foreign language learning and development, among other reasons, because they require learners to work with numerous aspects of the language. As Samuda and Bygate (2008) argue, “through engaging with the task, learners are led to work with and integrate the different aspects of language for a larger purpose” (Samuda and Bygate, 2008: 8). However, since language is quite unpredictable, task use and its correct implementation becomes a complex object of analysis for both researchers and practitioners. Consequently, it is essential to understand what a task is and how it can be used in order to achieve the highest benefit for language learners.

In the following section, we will consider different variables that can affect how tasks are carried out and the outcome learners will achieve. Additionally, some of the most extended theories for task classification will be introduced.

2.3 Task classifications

The analysis of tasks facilitates their implementation and more accurate selection to achieve different pedagogic goals. Analysing tasks and their characteristics is an important starting point in order to shed light on how to implement them. Different task features, such as the access of the participants to information or having an open vs. closed solution, affect learners' performance and also L2 development (Robinson, 2011). Thus, a classification of task features and demands, as well as of their effect on L2 performance, is desirable in task-based formal language teaching.

In line with the previous definitions of tasks and the role of research, Pica (2012) stated that research should study and describe ways in which tasks can be designed in order to activate the linguistic and cognitive processes necessary for successful language development and acquisition, which, as already discussed, is the ultimate goal of tasks. Tasks have become an area of special interest for teachers and researchers, as they present a connection that allows them to complement their respective fields in a productive way. Therefore, tasks can be considered both a research tool and a learning activity since they can be adapted to help to give an answer to theoretical and practical challenges that practitioners and researchers may have to face (Pica, 2012). It is not surprising then that the study of tasks has attracted the attention of many researchers and numerous task classifications have been put forward.

One of the most popular classifications of communicative tasks was proposed by Pica, Kanagy and Falodun (1993). It is based on interactional criteria, on the direction of the 'information flow', and the outcome that is to be expected. These authors propose three categories for a communicative task typology: i) interactant relationships and requirements, ii) communication goals and iii) outcome goals. Using these constructs, they identified five task types: jigsaw, information gap, problem-solving, decision-making and opinion exchange.

- a) Jigsaw task: This task type requires a two-way information flow. As none of the participants is given all the necessary information for task completion, they need to work together in order to achieve a common goal. Jigsaw tasks are very likely to promote NoM since, as already mentioned, interaction is required.
- b) Information gap task: Although this term has been frequently used in the literature to refer to activities that promote communication, information gap tasks differ

from jigsaw tasks in that the information flow is one-way, that is, only one participant possesses the information the other participant needs to fulfil the task. In these tasks, NoM is required for task completion as the participants have to work together to reach an agreement. Nevertheless, interaction is more limited than in jigsaw tasks. Since each participant has a given role, the one that holds the information will have more opportunities to receive feedback on their output, but fewer possibilities to give feedback on unclear input. The participant who requests the information will experience the opposite.

- c) Problem-solving: In these tasks, the participants share the necessary information for task completion. Although interaction among learners is not essential to complete tasks of this type, it is likely to happen thanks to its singular goal and convergent outcome.
- d) Decision-making: Just as in problem solving tasks and in opinion exchange tasks (see below), in decision-making tasks interaction is not required. Participants have equal access to information and have to reach a convergent outcome although, in this case, there are a number of different possibilities.
- e) Opinion exchange: This task type engages learners in discussion and exchanges of ideas. As in the two previous types, participants hold all the essential information and interaction among them is not required. However, as opposed to what happens in problem-solving and decision-making, interactants are not expected to work towards a common goal. Because of this, these tasks are considered the ones offering fewer opportunities for NoM.

Table 1 summarizes the task classification described above.

Table 1 *Communication task types for L2 research and pedagogy analysis based on: Interactant (X/Y) relationships and requirements in communicating information (INF) to achieve task goals (Pica et al., 1993: 19).*

Task Type	INF			requester-supplier relationship	Interactant requirement	Goal orientation	Outcome options
	holder	requester	supplier				
Jigsaw	X & Y	X & Y	X & Y	2 way (X to Y & Y to X)	+ required	+ convergent	1
Information gap	X or Y	X or Y	X or Y	1 way > 2 way (X to Y/ Y to X)	+ required	+ convergent	1
Problem-solving	X = Y	X = Y	X = Y	2 way > 1 way (X to Y & Y to X)	- required	+ convergent	1
Decision-making	X = Y	X = Y	X = Y	2 way > 1 way (X to Y & Y to X)	- required	+ convergent	1 +
Opinion exchange	X = Y	X = Y	X = Y	2 way > 1 way (X to Y & Y to X)	- required	- convergent	1 +/-

According to Pica et al. (1993) and Pica et al. (2006), the most effective task types are jigsaws and information gap tasks since interaction among task-takers is required in order to complete the task by achieving a common goal. Decision-making and opinion exchange are considered the least effective because learners' interaction is not necessary and they offer possibilities for more than one outcome, which may lead to a decrease in negotiation among learners.

Based on Skehan's Trade-off Hypothesis (1998), which posits that human's attentional capacity is limited and therefore attending to one specific dimension of performance may take attention away from others, another taxonomy for SLA oral tasks has been provided (Skehan, 1996; 1998; 2003; Skehan and Foster, 2001). According to the Trade-off

Hypothesis, increased task difficulty, which requires more attentional resources, may lead to a poorer performance in some areas. Nevertheless, trade-off effects can be attenuated by manipulating certain aspects of a task. Hence, three main constructs for the analysis of tasks are proposed: i) Aspects that contribute to code complexity (formal factors, i.e. syntactic and lexical difficulty), ii) Aspects that relate to cognitive complexity (content), and iii) Communicative stress, that is, the pressure participants feel to achieve communication. Additionally, some factors are believed to influence task performance and learning, making tasks more difficult ('complexifying/pressuring' influences) or easier ('easing/focusing' influences). The aspects that affect performance are classified into: i) Familiarity of information, ii) Dialogic vs. monologic, iii) Degree of structure, iv) Complex outcome and v) Transformations of task material. According to Skehan, this task classification is consistent with the "the language required, the thinking required and the performance conditions for a task" (Skehan, 1998: 99).

Skehan (2001) analysed the effect of these variables on general competence (complexity, accuracy and fluency) when working with different task types. Although some limitations were reported, the results suggest that task features influence different aspects of language. Thus, the degree of structure and familiarity of information had a greater effect (the effect of familiarity was only 'slightly greater') on fluency than on complexity or accuracy, whereas complexity of outcome appeared to favour complexity, and dialogic tasks, accuracy. Regarding the transformation of task material, planning time generated greater complexity.

Leaver and Willis (2004) also took two of the main tenets of Pica et al.'s (1993) typology (i.e. information flow and outcome) and, following Skehan's complexity factors, added a third aspect: the cognitive demands a task may pose. They proposed three variables to bear in mind when analysing a task, according to the basis upon which the task is drawn up:

- Gap principle: Learners need some information to fulfil the task. To achieve the target outcome, some communicative interaction is needed. These are also known as 'jigsaw tasks' (Pica et al. 1993).
- Reaching a decision or solution: Participants interact in order to reach a decision or a solution cooperatively. These are mainly convergent tasks, as learners work together to achieve a goal, although they can also be used as divergent. Research

however, suggests that this task type produces less negotiation and is considered as less effective (Pica et al., 1993).

- Cognitive process: Willis (1996) offered a classification based on the cognitive processes necessary to perform the task. The following types are presented according to cognitive demands, from the most simple to the most challenging. However, it is important to bear in mind that by changing any variable, the level of cognitive demands may be modified.
 - Listing
 - Ordering and sorting
 - Comparing and contrasting
 - Problem solving
 - Sharing personal experiences
 - Creative tasks and projects

The last theoretical approach to task classifications we are going to present in this chapter is the one offered by Robinson (2007). Robinson's Cognition Hypothesis (2001) offers a rationale for the effect of task demands that differs from Skehan's Trade-off Hypothesis (1998). Robinson claims that tasks should be sequenced according to increasing cognitive complexity, which is argued to parallel the sequences children go through when acquiring the L1 (Robinson, 2001; 2005; 2011). In this way, trade-off effects of accuracy against complexity are avoided and, in contrast to Skehan's Hypothesis, "[...] on some dimensions of task demands [...] increasing complexity is argued to promote more accurate, grammaticized production *and* more complex, syntacticized utterances." (Robinson, 2011:14).

As many other authors do, Robinson (2007) also argues for the need for an empirical (non-intuitive) classification system, and claims that "Taxonomic descriptions can therefore guide research and application, but they must also develop to accommodate findings, and theoretical progress" (Robinson, 2007: 8). This author argues that a task classification should be *motivated by a theory* (which would prove how tasks have an effect on language performance and use, facilitating language acquisition and development), it should be *empirically researchable* (so that its effects can be predicted) and finally, *operationally feasible* (that is, ready to be used by practitioners and researches). In order to offer a task classification that meets these three requirements,

Robinson (2007) offers three kinds of analysis which he considers complement each other at different stages of L2 learning:

- i) Behaviour descriptive approach: This approach is based on the observation of what participants do when performing a task. Its aim is to identify target tasks, their subtasks and the steps needed to carry them out.
- ii) Information-theoretic approach: By means of this analysis, tasks can be classified according to the information processing stages and the cognitive processes engaged in mediating input and output for the successful performance of the task.
- iii) Ability requirements approach: This analysis would result in a task classification based on the human cognitive abilities the tasks require. Thus, learners' variables come into play.

Taking into consideration the approaches described above, and following his Cognition Hypothesis (Robinson, 2001; 2003; 2005), Robinson (2007) proposes three criteria for the classification of pedagogical tasks:

- a) Interactional criterion: The type or degree of participation of learners in a task will lead to different task variables such as: open/fixed solution; one-way/two-ways; +/--need to reach an agreement; number of participants or degree of negotiation. Participant variables also need to be taken into account, including students' proficiency level, gender, familiarity with the other learner, shared knowledge, learner role within the task, and knowledge of how to interact in the L2.
- b) Cognitive criterion: The task features related to this criterion are defined considering the information-theoretic analysis. Robinson interprets cognitive here as the learner's attention focus. Accordingly, tasks can be classified taking into account two different aspects that affect task complexity:
 - i. Resource-directing variables, which make cognitive/conceptual demands. They affect learners' focus of attention and facilitate language acquisition and development. These variables include the "here-and-now" vs. "there-and-then" reference, few vs. many different elements, information transmission vs. reasoning, or 1st person vs. 2nd or 3rd person perspective.
 - ii. Resource-dispersing variables that make performative/procedural demands, which affect the amount of attention learners will be able to devote to a specific aspect, that is, they facilitate access to the

interlanguage learners already possess. These demands include variables such as planning time, provision of background knowledge to perform the task, amount of demands and the existence of a necessary sequence for the task to be successfully performed.

- c) Ability-determinant criterion: Individual learners' characteristics vary greatly and, therefore, Robinson argues that the criteria described above are for 'groups of learners' and that they need to be combined with the learner variable in order to optimise results and learning outcomes. Within this criterion, affective variables related to the interactional criteria, such as high or low task-specific motivation, openness to experience, ability to control emotion or willingness to communicate are included. Variables related to task difficulty, like switching between goals in multiple or dual-task performance or the learners' working and reasoning capacities also exert an influence.

As noted throughout this section, task typologies often overlap as a function of the approach authors follow. Taking the above reviewed aspects into account, four main parameters for task design have been identified in the literature (Ellis, 2003; Leaver and Willis, 2004; Samuda and Bygate, 2008). These four parameters take into consideration two features each, and are commonly expressed as 'either-or' distinctions.

- Open vs. Closed tasks: Closed tasks are those with only one possible correct solution. They are very specific and often have a tight structure (e.g. spot-the-differences). Closed tasks are believed to enhance interaction and elicit more NoM (Ellis, 2003; Long, 1991; Pica et. al, 1993; Pica et. al, 2006). On the other hand, open tasks are more likely to give opportunities for more language use, and lead to longer interventions (Skehan, 1998).
- One-way vs. Two-way tasks (also known as reciprocal or non-reciprocal (Ellis, 2003)): In two-way or reciprocal tasks, all the participants have essential information to communicate and to receive, whereas in one-way or non-reciprocal tasks, only one of the participants controls the information. Interaction studies suggest that two-way tasks are more efficient in promoting interaction (Ellis, 2003; Pica et al., 2006).
- Focused vs. Unfocused tasks: Focused tasks are those designed to promote communicative language use, emphasising and working on a particular language feature. Unfocused tasks however, are not designed with the objective of

promoting any specific linguistic form, thus the learners are free to choose any language form to communicate (Ellis, 2003; García Mayo, 2017b).

- Real-world target tasks vs. Pedagogic tasks: Real-world or authentic tasks are those that may be carried out in a real-life situation (Long and Crookes, 1992). From these, pedagogic tasks are derived and worked on in the classroom. More recently, the term ‘real-world’ is being used to make reference to the spontaneous language used during interaction in the classroom, which can also be used in real-life (Ellis, 2003), making these two parameters less clear.

The classifications presented above define different aspects of pedagogic tasks, which are very important when designing materials. Different task features produce specific types of interaction or ‘negotiated talk’ and are more or less effective on different aspects of acquisition. In addition to this, it is essential to bear in mind the importance of the learner variable, which cannot be fully predicted (Robinson, 2007; Samuda and Bygate, 2008), and the fact that, due to the great variety of task types, contexts, participants, etc. a firm conclusion has not been reached yet as to how all these characteristics relate to the language acquisition process. In the following section, the task we have worked with in the current dissertation, namely the picture placement task, will be described in detail, together with a revision of other research studies which have also used different jigsaw type tasks.

2.3.1 Jigsaw tasks

Pair and group work have become increasingly popular in language teaching. Particularly, thanks to the communicative approach, group work is regarded as a means to grant learners a greater amount of practice in the TL (Gagné and Parks, 2013; Long and Porter, 1985; Pica, 2002; Pica et al., 1996).

As mentioned above, the most efficient tasks are those in which participants must exchange information necessary to achieve a common unique goal (Ellis, 2003; Pica, 2013; Pica et al., 1993; Pica et al. 2006). These conditions make mutual understanding essential and will most likely lead to modifications in the learners’ output through NoM. This way, learners will focus on how they formulate their message in order to be understood, and not only on what they want to say (Long, 1981; Pica et al., 2006; Varonis and Gass, 1985). Information gap and jigsaw tasks are amongst the collaborative tasks

that research has proven as most effective for language learning in terms of interaction and, therefore, NoM (Pica, 2013; Pica et al., 1996; Pica et al., 2006). Thus, the implementation of this task type is becoming more and more frequent in FL material, particularly at elementary schools (Butler and Zeng, 2014). Moreover, these task types have been adjusted to work effectively both as learning activities as well as research tools (Pica, 2005; Pica et al. 2006). Taking all this into account, we have chosen a jigsaw task for our study, specifically a picture placement task.

Picture placement tasks (henceforth PPT) are two-way communicative tasks (although they can also be considered one-way repeated (Azkarai, 2013)) as the two members of the pair hold essential information to achieve the task goal, and both of them act as information holders, requesters and suppliers. As noted above (cf. 2.3), in jigsaw tasks participants have to interact in order to achieve a closed common outcome. This task type has been shown to promote modified interaction among learners, providing opportunities for NoM and drawing their attention to form (Pica et al., 2006; Pinter, 2007; Swain and Lapkin, 2001).

Jigsaw tasks have been used in a number of studies within the interactionist framework to investigate the role they play in the provision of opportunities to use the L2 in ways that benefit language acquisition. In what follows, we will present some studies that include this type of task in their research.

Mackey and Oliver (2002) used a PPT, together with other communicative tasks (namely spot-the-differences, story completion and picture sequencing task) to test the effect of interactional feedback on ESL YLs. The authors stated that these tasks were used because they provide a context for the target structures (question forms) as well as opportunities for negotiated interaction, which are two of the main characteristics of pedagogical tasks. In this study, the results obtained from the different tasks were merged so the potential differences across tasks were not accounted for. Along the lines of the results reported on in research with adults, this study provides further positive evidence of the beneficial role of interactional feedback on SL acquisition of child learners. Moreover, it is worth noticing that children benefited more rapidly from interactional feedback than the adults in their experiment.

Oliver (2002) studied the conversational interaction of young ESL learners (n = 192, age 8-13) using two communication tasks, a picture description (one-way) and a jigsaw (two-

way). She examined the effects of learner variables (i.e. native/non-native speaker (NS/NNS), proficiency, age and gender) on the NoM strategies used by children. As in Mackey and Oliver (2002), the author combined the results obtained from the two tasks alleging non significant differences of these two task types on the amount of negotiation. Supporting previous research findings, her results indicate that the individual characteristics of YLs influence interaction. Regarding the NS/NNS variable, results show a greater amount of negotiation for the NNS–NNS pairs in relation to the NS–NNS pairs, which in turn used less negotiation strategies than the NS–NS dyads. Although TL proficiency had an effect on the negotiation that took place among the participants, it was the least native-like dyads the ones that negotiated the most, as opposed to the expected pattern of High-Low dyads engaging in more negotiation. These were followed by the High-High dyads. Finally, when analysing the last two variables (i.e. age and gender) no significant effect on the negotiation strategies used by YLs was found, in contrast to results obtained with adult learners previously.

Pinter (2006, 2007) employed a spot-the-differences task, similar to the one we have used in the present dissertation. As already explained, this type of task requires the transaction of information between the participants. Pinter's (2006, 2007) conceptualization of tasks is based on Skehan's (1998) definition, and emphasizes the main focus on meaning, as well as the need for a communication gap to be filled in. These 'conflict points' make attention to detail essential to successfully complete the task. Because of this, NoM is expected as learners need to make themselves understood, check understanding of their counterpart, and deal with possible misunderstandings. Moreover, Pinter includes the similarities with real-world activities as a feature of the task, as well as the importance of task completion over other performance outcomes. In the first study (Pinter, 2006), the oral performance of 5 pairs of Hungarian EFL adults and 10 pairs of YLs was compared and the results revealed that, although children and adults use similar interaction strategies, YLs do so at a lesser degree. Moreover, children and adults were found to understand and complete the same task in different ways, a very important fact to consider when working with these two different populations. In Pinter (2007), the same spot-the-differences task was used to examine the effect of task repetition, using different versions of the task in each of the sessions. However, the differences were not stark, as all of them represented the same scenes, but with a different organization of how the elements were presented. This study will be further reviewed in the next section. Pinter (2006, 2007)

concludes that this task type proved to be beneficial, besides the positive effect on learners' general performance, because it provides learners with opportunities to interact using the TL without teacher intervention.

PPTs have also been used in studies with adult learners. For instance, Ross-Feldman (2007) included this task type when investigating the role of gender on task-based interactions and the opportunities for language development they promoted, operationalised as LREs. She analysed the production of 64 ESL university students, whose L1 was Spanish, performing three different collaborative tasks (a picture differences task, a PPT and a picture story task) in mixed- and matched-gender pairs. Partly concurring with Oliver's (2002) results regarding this variable in YLs' interaction, Ross-Feldman's study showed that the gender composition of the pairs did not significantly influence the incidence of LREs initiated by the participants (with the exception of on the picture story task, in which male–male dyads produced fewer LREs). With regard to the relationship between task and language learning, her results reveal that tasks have an effect on the incidence of LREs. The picture story task seems to lead to a higher amount of LREs than the other two tasks. The PPT was found to be the second task type in which learners initiated more LREs, and the picture differences the task that promoted fewer LREs. Finally, she suggests that tasks that incorporate a writing component (i.e. picture story) might create more language learning opportunities than oral-only tasks.

Motivated by Ross-Feldman's (2007) study (as reviewed above) Azkarai and García Mayo (2012) analysed the effects of different collaborative tasks (namely a picture differences task, a PPT, a picture story task and a dictogloss) on the oral performance of Basque-Spanish EFL university students. Their results concur with Ross-Feldman's (2007) study with no significant differences between dyad types being found. Nevertheless, and again in line with Ross-Feldman (2007), Azkarai and García Mayo (2012) found that task type had a clear effect on the promotion of interaction, and consequently, on language learning. Mirroring previous findings, the tasks that required the production of a written text (i.e. the picture story task and the dictogloss) generated more LREs than the picture differences and the PPT, in which the final outcome was only oral. As seen in the section on L1 use, Azkarai and García Mayo (2015) in the same context as in their 2012 study, observed that task modality also influenced learners' L1

use, reporting that collaborative speaking+writing trigger a more frequent use of the L1 than when learners engage in speaking-only tasks.

Research has shown that, in general, the type of task used has important effects on the participants' performance. PPTs (together with the other collaborative tasks mentioned above) have been used in research mainly to investigate learners' interaction and negotiation (Lázaro Ibarrola and Hidalgo, 2017a; Oliver, 2002); to analyse learner variables (Oliver, 2002; Ross-Feldman, 2007; Azkarai and García Mayo, 2012); as contexts for the application of a treatment (i.e. corrective feedback) (Mackey and Oliver, 2002), and learners' L1 use (Azkarai and García Mayo, 2015; García Mayo and Hidalgo, 2017). However, to our knowledge, PPTs have not been used to investigate the role of repetition in the NoM YLs might engage in when performing this type of task, which is one of the objectives of the present dissertation.

2.4 Task repetition

The literature to date suggests that different task features produce different effects on learners' interactional patterns and learning outcomes (Kim, 2009; Mackey et al., 2007; Nuevo, 2006; Révész, 2011). Among the numerous possible variables (e.g. cognitive complexity, planning time or participant role), in the present dissertation we will concentrate on task repetition (henceforth TR). Following Bygate and Samuda's (2005) definition, TR consists of "repetitions of the same or slightly altered tasks – whether whole tasks, or parts of a task" (Bygate and Samuda, 2005: 43).

The general claim about the effects of TR is that it helps learners to produce improved output. By repeating a task, learners become familiar with the content of the task, with its procedure, or both (Kim and Tracy-Ventura, 2013). This familiarity, gained through the repetition of the task, allows learners to take some aspects of the task for granted, which will facilitate task performance. Task familiarity is believed to have a positive effect on language learning as it seems to be beneficial for the learner's ability to focus on form, as well as to promote a better organization of the learners' linguistic resources (Mackey et al., 2007, Saeedi and Rahimi Kazerooni, 2014). When working with oral collaborative tasks, learners must focus both on form and meaning, which requires many of the learners' attentional resources. Sometimes this is difficult due to the spontaneous nature of oral communication and to humans' limited processing capacity (Skehan, 1998;

Skehan and Foster, 2001). When these processing difficulties arise, language learners generally prioritise meaning over form, particularly in information-gap and jigsaw tasks (Azkarai, 2013; García Mayo, 2011; Pica, 2002; Swain and Lapkin, 2001). Because of all this, the repetition of pedagogic tasks has become a valuable tool to grant learners the benefits of familiarity, and therefore to help them to shift their attention from meaning to form.

“[...] familiarity gives us the time and awareness to shift attention from message content to the selection and monitoring of appropriate language. By enabling a shift of attention, learners may be helped to integrate the competing demands of fluency, accuracy and complexity” (Bygate, 1999b: 41).

TR, and the subsequent task familiarity learners gain from it, affects different aspects of L2 production, such as fluency, accuracy, lexical precision and complexity. It has also been found to influence aspects of the interactional process, such as amount of NoM and provision and use of feedback, with TR generally leading to a more efficient organisation of the language resources (Bygate, 1996; 2001; Bygate and Samuda, 2005; Kim and Tracey-Ventura, 2013; Lynch and Maclean 2000; 2001; Patarnasorn, 2010; Pinter, 2007; Sample and Michel, 2014; Samuda and Bygate, 2008).

Lynch and Maclean (2000; 2001) speculate that these gains are partly attributed to the fact that learners feel more relaxed when the task is familiar and they are able to shift their attention from what they want to say to how they are going to say it. This aspect of repetition is related to communicative stress, the pressure learners feel to accomplish communication, as Skehan (1998) pointed out in his tasks typology. Later, and also based on Skehan's (1998) Trade-off Hypothesis, Ellis (2003) attributed the benefits of TR to interlanguage restructuring, which provides more attentional resources available to focus on form.

TR can be considered as preparedness: the first task performance serves as pre-task planning (Ellis, 2005), providing learners the opportunity to start to organise content and useful grammar and vocabulary in real time (Bygate and Samuda, 2005). This occurs because during this session learners tend to “prioritize conceptualization over formulation and articulation, it is closely associated with complexity of their language production” (Saeedi and Rahimi Kazerooni, 2014: 11). Therefore, this session allows learners to work

with the material and language they will need in their following performances, facilitating a better ‘packaging’ of their ideas and directing their attention from meaning to form, resulting in more fluency, accuracy and/or complexity (Ahmadian, 2012; Bygate, 1999a,b; 2001; 2006; Bygate and Samuda, 2005; Ellis, 2005; Samuda and Bygate, 2008). From a cognitive or information-processing perspective, learners’ attention capacity needs to control more processes the first time they perform a task, whereas when they repeat it, some of these processes have previously been monitored. Hence, some of the meaning and models needed will already be available so that not as many attentional resources are required as during the first time a task is carried out (Muranoi, 2007).

The literature reviewed above suggests that a great deal of attention has been placed on the repetition of oral tasks, whereas the written modality appears to have been disregarded. Nevertheless, as Manchón (2014) states, this is changing and the interest in the benefits of the repetition of written tasks is increasing considerably. The nature of writing has unique features which add a new perspective to the potential of TR for language learning. An important feature of written tasks is the greater availability of time during production, which allows learners to better concentrate on the language forms, and to devote more linguistic resources to what is being produced, as well as to the feedback received. Nevertheless, this field within TR, though essential as it may be, lies beyond the purposes of the present study.

Several studies have suggested that when TR is combined with learner interaction, learning opportunities increase and the effects of repetition can be reflected in the learners’ overall performance as well as in their use of interaction strategies and/or their ability to focus on form (Kim and Tracey-Ventura, 2013; Lynch and Maclean, 2001). In what follows, studies examining the effects of TR will be presented. Generally, these studies have analysed the influence of TR on two general aspects: effects on general competence (i.e. complexity, accuracy and fluency) and effects on NoM. Nevertheless, research addressing the latter is comparatively scarce.

2.4.1 Effects of repetition on general competence

In order to test the claims presented above about the benefits of TR, several authors have explored how TR affects L2 oral performance by measuring complexity, accuracy and fluency (CAF). These studies have mainly dealt with learners repeating monologic oral

tasks (i.e. tasks that do not require learner-learner interaction), using sets of picture prompts or videos (Ahmadian and Tavakoli, 2011; Bygate, 1996, 2001; Bygate and Samuda, 2005; Saeedi and Rahimi Kazerooni, 2014; Wang, 2009), although some studies working in dialogic contexts have also been reported (Lynch and Maclean, 2000, 2001; Hawkes, 2012). In what follows, sixteen empirical studies will be presented organised into i) studies testing the effects of TR on CAF and ii) studies dealing with the role of different aspects of TR on CAF.

i) Studies testing the effects of TR on learners' CAF.

In a small scale pilot study Bygate (1996) tested the effect of repetition on the performance of an ESL learner narrating a video extract immediately after watching it and repeating it two days after. The results suggest that the learners' oral production improved the second time she performed the task in terms of accuracy, fluency complexity. These findings were partly confirmed in a larger scale study by this author in 2001, when he found that repeating a narrative task led to gains in terms of complexity and fluency. In this latter study, Bygate (2001) administered two tasks (an interview and a narrative) over a 10-week period. However, the results also showed that the benefits of TR did not carry over to a new context, i.e. other task types.

Later on, Bygate and Samuda (2005) analysed the effect of repetition on 'framing', which is considered part of the complexity of the participant's discourse, and refers to "any language additional to the narrative content" (Bygate and Samuda, 2005: 47). The participants were 14 students of English of different levels of proficiency and different language backgrounds at a British university. In this study, the participants had to tell the story of a video they had previously seen. This task was administered twice and the results showed that the learners produced more elaborate language the second time they narrated the story than when they did it for the first time. Thus, although TR did not seem to significantly influence individual oral fluency, it improved the speakers' discourse complexity, allowing them to report the story in a more consistent way, as opposed to the 'disjointed reports of events' observed in the first performance. Moreover, Bygate and Samuda (2005) stated that "repeated encounters do not involve the learner in doing the 'same' thing, but rather in working differently on the same material" (p. 67).

Wang (2009) provided further evidence of the benefits of TR. He compared on-line planning, strategic planning and TR and his results suggest that repetition was the only

condition that had a significant effect on complexity, accuracy and fluency, being accuracy and fluency the most affected areas.

Ahmadian and Tavakoli (2010) examined the oral production of 18-21 years old Iranian intermediate EFL learners when narrating a story after watching a silent video. Similar to previous studies, these authors also reported positive effects of repetition on complexity and fluency (Bygate, 1996, 2001; Gass et al., 1999; Lynch and Maclean, 2000, 2001). Additionally, their results suggest that TR together with an additional treatment, such as careful online planning, leads to greater improvements in all CAF areas.

In a similar context, Saeedi and Rahimi Kazerooni (2014) examined the influence of the repetition of two different types of narrative tasks (loosely vs. tightly structured) on CAF in the oral performance of a group of 60 Iranian (undergraduate) adult EFL students. Their results suggest that narrative type influences the effect of TR on the learners' performance. The repetition of a tightly structured narrative appears to be more beneficial, as it leads to significant gains in the learners' complexity, accuracy and fluency. When working with a tightly structured narrative, speakers seem to have more attentional capacities to monitor these formal aspects, instead of devoting the time to unraveling the argument of the story. On the other hand, repeating a loosely structured narrative also benefits learners' oral complexity and fluency, but it does not make a significant difference in their accuracy.

So far, the above reviewed studies analyse monologic tasks. Nevertheless, the effects of collaborative TR on CAF have also been studied. Lynch and Maclean's (2000, 2001) findings provided support to Bygate's (1996) suggestion that TR with different partners enhances its benefits since "different people will do tasks in different ways and a variety of partners could provide valuable learning opportunities" (Bygate, 1996: 145). In their study, learners improved without teacher intervention, interlocutors pushing each other for a more accurate and consistent performance. These authors studied immediate TR (as opposed to repetition over days/weeks intervals) with 14 adults performing a poster carousel task in an English for specific purposes (ESP) context. They found that in the repeat task, learners paid more attention to language, performing more accurately and fluently, concurring partly with Bygate (1996, 2001) and Bygate and Samuda (2005) (see above). On a final note, Lynch and Maclean (2001) claimed that learners did not perceive

the repetition of the task as ‘repetitious’, a possibility that may prevent practitioners from using TR in the classroom.

Also arguing in favour of TR as a tool to direct learners’ attention towards form, Hawkes (2012) claims that a form-focus post-task activity is necessary for students to shift their attention from meaning to form when engaged in communicative tasks. In his study, 14-15 year-old students of English in a junior high school in Japan performed three tasks (an opinion exchange task, a describe and draw task, and a timed conversation) with a form-focus session on the explicit language they needed, before TR. The results show that learners’ accuracy, use of the TL and confidence to perform the task improved with the repeat performance. The author argues that “TBLT with no focus on form post-task session may not be enough for interlanguage acquisition” (Hawkes, 2012: 335).

Two recent studies have addressed the effect of TR on young EFL learners’ CAF, using the framework of Skehan’s Trade-off hypothesis (1998). These are very interesting for our current study as the setting (EFL) as well as the population (9-10-year-old YLs) are very similar. The participants in Sample and Michel’s (2014) small-scale exploratory study were six young Chinese EFL learners who completed an oral spot-the-differences task three times. Their findings show how TR seems to benefit task performance, although each of the CAF components is not equally influenced. Fluency, for instance, is the only dimension that improves significantly across tasks. Moreover, their results suggest that the first and the second time the task is repeated, some trade-off effects are observed: in the first repetition, the speakers who produced more complex structures also made more errors. In the second repetition, fluency seemed to be favoured over lexical complexity, as the participants who used more elaborate lexicon were less fluent. Nevertheless, these trade-off effects seem to disappear once the learners become familiar with the task, thanks to TR.

Bret Blasco (2014) analysed the effect of TR on CAF in two different oral tasks (an interview and a narrative task). The data were collected at four points over a period of two academic years. Her findings mirror Sample and Michel’s (2014) and show how the three main dimensions of general L2 performance (CAF) do not develop simultaneously in either of the two tasks the learners carried out. A significant increase in fluency through TR was reported, together with mixed findings with regard to complexity and accuracy. Her results suggest trade-off effects between complexity and accuracy in regard to the

interview, whereas in the case of the narrative task, the improvement in fluency and complexity across tasks are more pronounced. Accuracy is reported to be the least affected dimension by time, task and instructional setting.

ii) Studies on the role of different types of TR on learners' CAF.

Gass et al. (1999) analysed the effects of different aspects of TR (specifically, procedural repetition, which consists in repeating the same task type but with different content, and exact same repetition) on the output of 103 undergraduate students of Spanish at an American university. The participants had to tell a story after watching a short video extract. A positive relationship between content repetition and the learners' overall proficiency, lexical sophistication and morphosyntax was reported, resulting in a more accurate use of certain linguistic aspects. However, their results also suggest that these benefits may not generalise to tasks with new content.

Pinter (2007) analysed the oral performance of one pair of EFL Hungarian YLs (age 11) with a low command of the TL. Her study addressed the effect of procedural repetition on peer-peer interaction. The participants repeated a spot-the-differences task three times over a period of three weeks following the same procedure but with slightly different content every time. Considering the alleged egocentricity of children, Pinter (2007) focused on how TR promoted the attention learners paid to each other's needs. Her findings suggest that children this age, even at a very low level of competence, benefit from peer-peer interaction, reporting a greater self-confidence and fluency by the learners' last performance. The participants were interviewed after the last performance, and the children acknowledged the advantages of TR and peer interaction. For instance, they felt more confident and relaxed and noticed the gaps in their previous performances. This author concluded that the repetition of this type of task provides learners with plenty of opportunities to use the TL and that YLs benefit from it without or before teacher intervention thus representing a very valuable tool for the language classroom.

Bygate (2009) provided further evidence of the effect of exact same repetition on complexity and fluency. In his study, 48 university students were divided into groups which performed two different task types over a 10-week period. Both groups carried out a narrative and an interview task in the first session and repeated it 10 weeks later (content TR). During those 10 weeks, the each group performed the same task type three times, either a narrative or an interview task. The results are twofold: whereas the repetition of

a specific task after 10 weeks led to more complex language in the two tasks (interview and narrative), fluency was affected only in the interview task, which decreased as more pauses were identified. As for the repetition of the same task type (procedural TR), the results revealed an effect on both fluency and accuracy. The two groups were less fluent (i.e. more pauses) and their language more complex in the last performance.

Slightly different results were obtained by Patanasorn (2010). This author worked with 92 Thai EFL undergraduate students and compared exact same task repetition, procedural repetition and content repetition. The results suggest that while procedural repetition seems to encourage more accuracy in the use of the simple past, the content repetition group improved in terms of fluency. In contrast, no changes were observed in the exact same task repetition group.

Kim and Tracey-Ventura (2013) also examined the benefits of these two different types of TR (i.e. procedural repetition and exact same task repetition) on learners' L2 oral performance. The participants were 36 13 year-old Korean female students who had been learning English for at least 4 years at school. They used the same pre-test and post-tests with both groups, but one of the groups repeated the exact same task three times whereas the other performed three tasks following the same procedure with a different content. Their results suggest that both types of TR are beneficial for language acquisition and produce few differences in learners' oral performance. Procedural repetition led to a greater syntactic development in the first post-test, although not in the second, whereas the exact same task repetition group produced fewer clauses on both tests. While little difference was found in the overall accuracy of the two groups, both improved significantly in their use of the task-induced linguistic feature (past simple). Interestingly, in the last post-test fluency decreased in both groups, suggesting "possible trade-off effects among syntactic complexity, accuracy and fluency measured by speech rate and amount of reformulation" (Kim and Tracy-Ventura, 2013: 839). This may be due to the fact that during interactive tasks learners produce many interactive features, such as negotiation strategies. They concluded that their results do not support any method over the other, since no significant differences were found between the two groups' performance.

García Mayo et al (in press) compared the effect of procedural repetition and exact same task repetition on the oral performance of 60 pairs of Spanish young EFL learners (aged

8-9 and 9-10) in a CLIL school. The participants performed a spot-the-differences task, which was repeated three months later. A positive effect of procedural repetition on the younger learners' fluency was identified, whereas in the performance of the group of older YLs a significant increase in accuracy took place. Exact same task repetition, on the other hand, did not seem to have any significant effect on the general performance of the older learners and the only significant change in the younger learners group was a decrease in oral complexity.

More recently, Lázaro Ibarrola and Hidalgo (2017b) addressed the effect of procedural repetition on the oral production of a group of 20 11-year-olds learning English in a CLIL school in Spain while carrying out an interactive task in pairs. The participants had to resolve the same task type (a PPT) with different content three times over a three-week period. These authors reported that the only remarkable effect of this type of repetition was a slight improvement in accuracy, observed in terms of a lower number of errors per clause. A significant improvement in fluency took place in the second task performance but, even though the last performance was the most fluid, the final improvement did not reach statistical significance. Complexity, on the other hand, remained clearly stable throughout the three tasks.

In summary, although some trade-off effects have been reported, research to date points to the beneficial effects of TR on learners' general oral performance (e.g. CAF) (Ahmadian and Tavakoli, 2010; Patanasorn, 2010, among many others). The overall results suggest that manipulating the language we work with facilitates its later use (Ahmadian, 2012; Bygate, 2001, 2006; Ellis, 2005), giving the speakers better accessibility to their linguistic resources and helping them to avoid errors they made the first time they performed the task. However, the great diversity of interacting variables in the above reported studies (context, age, level, task, time span between repetition, etc.) makes it impossible to specify the concrete gains that TR will yield in a given situation. In order to provide a clearer view of the different studies reviewed in these two sections, tables 2 and 3 offer a summarised picture of the main aspects of each of these research papers.

Table 2 Summary of the findings on the effects of TR on learners' CAF.

Authors	Setting	Task	Number of repetitions & time span	Goal	Results
Bygate (1996)	ESL 1 participant	Oral story re-telling task. Monologic.	Twice 3-day interval	To test the effects of TR on learners' oral production (CAF).	(+) CAF. (+) reuse of linguistic constructions. (+) metalinguistic comments.
Bygate (2001)	ESL university students	Oral story re-telling task and interview. Monologic.	Twice 10-week interval	To test the effects of TR on learners' oral production (CAF).	(+) complexity. (+) fluency. (-) benefits did not generalise with new content.
Lynch & Maclean (2000, 2001)	ESP adult learners	Poster carousel. Presentation to different peers. Dialogic.	6 times Same session	To analyse the effect of immediate TR on learners oral production (CAF).	(+) accuracy. (+) fluency.
Bygate & Samuda (2005)	ESL university students	Oral story re-telling task. Monologic.	Twice 10-week interval	To test the effects of TR on discourse framing.	(+) complexity. (+) fluency. (+) framing.
Wang (2009)	EFL	Oral story re-telling task. Monologic.	Twice Same session	To compare the effects of on- line planning, strategic planning and TR on CAF.	TR was the one that had a significant effect on CAF.
Ahmadian & Tavakoli (2010)	EFL university students	Oral story re-telling task. Monologic.	Twice 1-week interval	To test the effects of on-line planning and TR on learners' oral production (CAF).	(+) complexity. (+) fluency. Repetition & careful online planning → (+) CAF.

Hawkes (2012)	EFL high school students	Opinion exchange, describe and draw task and a timed conversation. Dialogic.	Twice Same session	To test the effectiveness of a form-focus post-task session before the repeat task.	(+) accuracy. (+) overall performance. (+) confidence.
Saeedi & Rahimi Kazerooni (2014)	EFL university students	Oral story re-telling task. Monologic.	Twice 1-week interval	To examine the effects of repeating 2 narrative types (loosely structured narrative vs. tightly structured narrative) on learners' oral production (CAF).	Loosely structured → (+) complexity & fluency. (=) accuracy. Tightly structured → (+) CAF.
Sample & Michel (2014)	EFL YLs	Spot-the-differences. Dialogic.	3 times 1-week interval	To examine the relations between CAF & the effect TR exerts on them.	(+) fluency. (+) overall performance. Trade-off effects diminish with familiarity.
Bret Blasco (2014)	EFL YLs	Interview (dialogic). Narrative task (monologic).	4 times 5/6-month interval	To examine the relations between CAF & the effect TR exerts on them.	(+) fluency & complexity. (=) accuracy. Trade-off effects reported.

Table 3 Summary of the findings on the role of different types of TR on learners' CAF.

Authors	Setting	Task	Number of repetitions & time span	Goal	Results
Gass et al. (1999)	SPL university students	Oral story re-telling. Monologic.	4 times 2/3-day interval 1 week interval (final task)	To test the effect of content repetition as a tool to allocate more attention to form on learners' output.	(+) overall proficiency. (+) lexical sophistication & morphosyntax. (+) accuracy. (-) benefits did not generalise with new content.
Pinter (2007)	EFL YLs	Spot-the-differences. Dialogic.	3 times 3-week interval	To analyse the effect of task repetition on YLs' peer-peer oral performance.	YLs assisted each other throughout the tasks. Repetition promoted: (+) fluency. (+) relax & confidence.
Bygate (2009)	ESL university students	Narrative. Interview. Monologic.	5 times 2-week interval	To test the effects of content and procedural TR on learners' oral production (CAF).	Content TR: (+) complexity. (-) fluency (interview) Procedural TR: (+) complexity. (-) fluency.
Patanasorn (2010)	EFL university students	Collaborative tasks. Oral story re-telling.	3 times 1-day interval Immediate Post-test 1 week later Delayed Post-test 2 weeks later	To compare the effect of procedural repetition, content repetition and exact same task repetition.	Procedural TR: (+) accuracy. Content TR: (+) fluency. Exact same TR: (=) overall performance.

Kim & Tracey-Ventura (2013)	EFL high school students	Information exchange tasks. Dialogic.	3 times 1-day interval Post-test 1 week later Post-test 2 weeks later	To compare the effect of procedural repetition and exact same task repetition.	Both methods reported as beneficial (+) task-induced linguistic feature. Last post-test: (-) fluency: possible trade-off effects among CAF.
García Mayo et al. (in press)	EFL YLs	Spot-the-differences. Dialogic.	2 times 3-month interval	To compare the effect of procedural repetition and exact same task repetition.	Exact same TR: Age 8-9 (-) complexity. Procedural TR: Age 8-9 (+) fluency. Age 9-10 (+) accuracy.
Lázaro Ibarrola & Hidalgo (2017b)	EFL YLs	PPT. Dialogic.	3 times 3-week interval	To measure the effect of TR on the oral production of YLs in a CLIL program.	(-) repetitions → Learners face fewer difficulties in understanding each other & resolving the task. (+) accuracy. (+) fluency T1-T2

2.4.2 Effects of repetition on negotiation of meaning

The benefits of NoM have been recognized by many authors and, consequently, an interest has developed in how task design can stimulate it (Gass and Varonis, 1985; Long, 1981; Pica and Doughty, 1985, Pica et al., 2006). As seen in Chapter 1, there is empirical evidence of the positive effect of NoM, a process in which learners experience the need to modify their output, on language acquisition (Pica et al. 1996). However, research has provided inconclusive findings regarding the effect of task design on interactional features (e.g. negotiation strategies), and only few studies have been carried out considering the effect of TR on the use of negotiation strategies, leaving much room for discussion on this issue (García Mayo and Imaz Agirre, 2016; Lázaro Ibarrola and Hidalgo, 2017b; Mackey et al., 2007).

One of the few studies dealing with this variable is the one by Mackey et al. (2007), who examined the effect of different types of repetition on the amount of NoM and provision and use of feedback. The participants in their study, 40 7-8 years old ESL learners, performed several oral communicative tasks in pairs. The learners performed two practice tasks one week before the first data collection time and then, two weeks later, the repetition tasks were carried out. The findings show that task familiarity had an effect on both the use of interaction strategies and the provision and use of feedback. Unfamiliar tasks (unfamiliar procedure and content) generated more CAs whereas procedurally familiar tasks produced more opportunities to use feedback as well as more actual use of feedback. On the other hand, familiar tasks (both content and procedure) led to more actual use of feedback. That is to say, there was more negotiation and feedback during unfamiliar tasks while learners attended more to the feedback when the task was familiar.

García Mayo and Imaz Agirre (2016) analysed the effect of TR on YLs' NoM. The authors examined the oral interaction of 60 pairs of YLs from two age groups (8-9 and 9-10 years old) while completing two communicative tasks at two testing times under two different TR conditions. At the first testing time, all the participants completed a spot-the-differences task. At the second data collection point, 21 pairs repeated the same task (exact same repetition), 16 pairs performed the same task type, but with different content (procedural repetition) and 23 pairs performed a different task (a guessing game). Their data revealed that the differences in the learners' use of NoM strategies between task 1 and task 2 did not reach statistical significance under any of the TR conditions examined.

The authors attribute this finding to a possible lack of interest on the part of the learners as the tasks may have been too easy for them.

Lázaro Ibarrola and Hidalgo (2017b) in their study, analysed the participants' oral interactions to search for any effects of repetition not only on the learners' general performance, but also in their use of interaction strategies. Their results show that, apart from accuracy, the only aspect which seems to be clearly affected by repetition is the amount of repetitions to solve communication breakdowns, which decreases significantly on the third performance, supporting Mackey et al.'s (2007) claim that more negotiation takes place during unfamiliar tasks.

However, more research evidence is needed to explore to what extent TR affects the use of negotiation strategies by YLs, which is what we intend to shed light on in our study.

2.5 Conclusion

In this chapter an overview of the groundwork for the implementation of TBLT has been provided, together with a review of the main issues that surround this approach. As seen in the analysis of the definitions and their main characteristics, tasks are a very valuable tool for second and foreign language learning. After assessing the existing classifications of tasks, jigsaw and information gap tasks are considered the most effective for language acquisition, and the ones that trigger the most NoM (Pica, 2013; Pica et al., 1996; Pica et al., 2006). Jigsaw tasks have proved to be very beneficial to promote interaction, specially the PPT. This is one of the reasons why we have selected this type of task for our study. In order to evaluate the effectiveness of this task type, a revision of its specific aspects has been provided. Moreover, some studies which have used this task have been reviewed.

The construct TR has also been addressed, as it is one of the variables that affect learners' interactional patterns and CAF, and one of the foundations of the present dissertation. Although general positive results have been reported, some slightly different conclusions have been reached. The other aspect our study is mostly concerned with, the effect of repetition on NoM has, to our knowledge, not been sufficiently analysed yet, and contradictory results have been reported (see for instance Gass et al. (1999) and Mackey et al. (2007)). Nevertheless, the different findings may be due to different variables such as age, FL vs. SL context, task type, or level of proficiency in the TL. Thus, the present

study aims to fill the gap in research on YLs' oral production in EFL settings as well as the impact of TR on the general performance of the members of this cohort.

CHAPTER 3 CHILD SECOND LANGUAGE LEARNING

As stated in Chapter 1 in relation to the Interaction Approach, individual learner differences affect different aspects of language learning. Consequently, age, being one of these differences, is expected to have significant effects on SLA (Mackey et al., 2003). The main aim of the present chapter is to reflect on the unique nature of a population which has received increasing attention in recent years: young L2 learners (Cameron, 2001, 2003; García Mayo, 2017a; Muñoz, 2006, 2014; Nikolov and Mihaljevic-Djigunović, 2006, 2011; Pinter, 2007, 2011). We seek to give a detailed description of the main features of this population and to point out the key differences between young and older learners. An overview of the research on child SLA and age-related aspects of L2 acquisition will also be provided. Finally, we will review recent findings on child interaction in second and foreign language settings. The main issues of interaction among young language learners will be illustrated with relevant empirical findings and some implications for theory and pedagogy that might derive from this evidence will be drawn.

3.1 Child SLA research

To date, and despite the acknowledged differences between adult and child learners, there is a general lack of focus on YLs SLA (Lázaro Ibarrola and Azpilicueta Martínez, 2015; Philp et al., 2006; among others). Most SLA studies involve adult learners or high school students and their results are still frequently applied as the basis for teaching methodologies, whilst research pertaining to children remains comparatively scant (Butler and Zeng, 2014; Gagné and Parks, 2013; García Mayo, 2017a, García Mayo and Lázaro Ibarrola, 2015; Mackey and Oliver, 2002; Pinter, 2007).

An exception to this has been the pioneering work carried out by Oliver (1998, 2002, 2009) within the Interaction Approach and YLs in ESL settings as well as the work done in Canadian immersion programmes. Nonetheless, it is a fact that the potential differences between learners of different ages are receiving increasing attention in recent years (Butler and Zeng, 2014; Mackey, 2012; Nikolov, 2009; Nikolov and Mihaljevic-Djigunović, 2011) and studies dealing with YLs are becoming more numerous. Furthermore, the contexts in which research is being carried out are more varied, including not only SL settings but also FL learning contexts (García Mayo, 2017; García Mayo and García Lecumberri, 2003; Muñoz, 2006).

In the literature, there is a growing body of empirical findings that shows how age has an effect on L2 acquisition (Birdsong, 2005a; Mackey et al., 2003; Pinter, 2011, 2017; Muñoz and Singleton, 2011; among many others). It has long been hypothesized that YLs have an advantage over adults when learning an L2 and that they tend to achieve higher levels of proficiency in the long term (Nikolov, 2009; Nikolov and Mihaljevic-Djigunović, 2006), which has often been taken as the rationale for the introduction of early start second or foreign language learning programmes (Cameron, 2001; Muñoz, 2006). Nevertheless, this widespread belief that ‘the younger the better’ for SLA, is often grounded in the observation of language acquisition by children in naturalistic settings (immigrants) and also in the study of SLA by YLs in immersion programmes (e.g. Canada) (Muñoz, 2006; Pinter, 2017). However, most research in EFL school settings has shown that the initial advantages younger learners may have tend to disappear by the age of 16 (Pinter, 2017). These learning contexts cannot be compared, as in FL settings learners have limited access to the TL outside the classroom and reduced opportunities to interact compared to learners in SL settings (Muñoz, 2014; Philp and Tognini, 2009). Thus, studies in FL settings suggest that the instructional context, rather than age, may constitute the determining factor underlying the success of these programs (Muñoz, 2006; Murphy, 2014).

Researchers in FL settings have found that YLs develop their interlanguage at a slower rate than adolescents or adults (Muñoz, 2006; 2007b; Nikolov, 2009; Nikolov and Mihaljevic-Djigunović, 2006; 2011). In line with this, Mackey et al. (2003) stated that “[...] regarding the rate of acquisition, older learners tend to demonstrate initial advantages, especially in grammar” (p. 41). This appears to happen thanks to the access to strategies and analytical learning methods adults have which YLs lack (Cameron, 2001; Lightbown and Spada, 2006). In a nutshell, while YLs may do better in ultimate attainment in comparison to adult learners, their learning process seems to be slower and needs larger amounts of exposure to the TL. Thus, it is essential to bear in mind the minimal conditions that YLs need to reach native-like proficiency (e.g. amount of exposure to the TL) when analysing YLs’ second or foreign language acquisition (Dimroth, 2008).

In spite of the existing evidence (as outlined above), the observation of the hypothetical advantages YLs seem to have when acquiring an L2 has triggered numerous studies, some of which have proposed a critical period for SLA (DeKeyser, 2000, 2003; Lenneberg,

1967). The Critical Period Hypothesis (henceforth CPH), posited by Lenneberg (1967), claims that a language needs to be learnt before the end of that critical period (which has been understood by many as before puberty, although whose exact temporal boundaries are still under discussion (Singleton, 2005)) in order to achieve native-like proficiency. According to the CPH, age effects are believed to be related to biological changes that affect the human capacity for language acquisition, and maturation goes along with a decrease in the language learning potential. Some authors have related these age effects on language acquisition to the concept of Universal Grammar (UG), considered to be the innate knowledge of a set of abstract grammatical rules that facilitates language learning (Chomsky, 1986), and which is believed to fade with maturation. Since adult learners are claimed to lose access to UG, they need to rely on other strategies, such as general problem-solving procedures (Bley-Vroman, 2009), which are supposed to be less suited and efficient for language acquisition (Dimroth, 2008).

It is important to note, however, that a growing body of research has challenged the CPH on the basis of accumulating empirical evidence against it. One of the observed contradictions that undermines the validity of the CPH are the different onset and offset ages that have been proposed, as well as the suggestion of a continuous decline on language outcomes both before and after maturation (Birdsong, 2005a; Singleton, 2005), as opposed to the initial view of a precise boundary after the critical period. Furthermore, there is a range of different views on the areas affected by the CPH (Singleton, 2005). For instance, some authors have stated that some adult learners have in fact been able to attain native-like proficiency in at least some tasks, which makes the strong version of the CPH no longer sustainable (Birdsong, 2005a; Nikolov and Mihaljevic-Djigunović, 2006). Finally, the conceptualization of ‘nativelikeness’ has not been clearly defined yet (see Birdsong (2005b) for a discussion of this concept), which makes it difficult to determine whether a learner has achieved native-like proficiency or not. Thus, the lack of consensus regarding the effect of age and maturational development on SLA has greatly undermined the reliability of the CPH (Birdsong, 2005a; Singleton, 2005).

Such evidence is essential when we look into ‘the younger the better’ as the reason for earlier start programmes for FL learning. As mentioned above, the implementation of these programmes in FL settings is often based on the assumption that children acquire higher levels of proficiency (but forgetting their slower rate of acquisition). Besides,

‘Early Start’ programmes have also taken as support the expectation of a FL acquisition similar to that achieved in SL immersion situations (Muñoz, 2014; Pinter, 2017). As will be discussed below, the specific characteristics of these two different settings (e.g. type and amount of input) have a strong influence on language learning. As Muñoz (2006) points out regarding immersion settings:

[...] empirical research in those contexts has shown that individuals who begin to learn a second language very early in life generally attain higher levels of proficiency than those who start at a later stage. However, an inferential leap is made in the assumption that learning age will have the same effect on students of a foreign language, where they are exposed to only one speaker of that language (the teacher, who is not usually a native speaker), in only one setting (the classroom) and only during very limited amounts of time (Muñoz, 2006: vii).

Another significant factor is that, when different starting ages (in FL contexts) have been compared, results have suggested that late starters seem to have some advantages over early starters when the amount of exposure to TL input has been controlled for (see García Mayo and García Lecumberri (2003) and Muñoz (2006, 2007b) in the Spanish context).

The chapters in the volume edited by García Mayo and García Lecumberri (2003) deal with the age factor from a FL acquisition perspective (rather than a SL) and present the TL (English) as a third language being learnt in two bilingual communities in Spain (The Basque Country and Catalonia). The results obtained by the research carried out in these communities indicate that an early formal instruction of the FL does not “contribute to better results as regards to proficiency in that language” (García Mayo and García Lecumberri, 2003: ix). The studies conducted in the Basque Country explore the effect of age, among other factors, on English acquisition in Basque bilingual schools. The findings suggest that, while YLs may have better attitudes and motivation towards the TL, older learners seem to progress more quickly, reaching better linguistic outcomes. The researchers attributed these results to the older learners’ more developed cognitive maturity, and a longer amount of exposure to the TL. The last two contributions in the volume present results obtained in Catalonia, which again reported that learners’ cognitive maturity and TL exposure yielded more beneficial effects on language acquisition than an early start in FL instruction.

The volume edited by Muñoz (2006) presents results from the Barcelona Age Factor (BAF) Project (1995-2002). The focus of this research project was to examine the effect of different onset ages (8, 11, 14 and +18 years old) on instructed EFL learning. This research project was more concerned with the rate of acquisition than with ultimate attainment. The different studies provide evidence of how late starters outperform early starters in most of the measures and during different times of data collection and report how older learners progress faster. In addition, their results suggest that late starters may benefit from ‘explicit teaching processes in the classroom’. On the other hand, earlier starters’ language development might not be favoured by the ‘implicit learning mechanisms’ they have access to due to the limited amount of TL input they are exposed to in the classroom (p. 33). Muñoz’s (2006) study underlines that the acquisition of areas that are learnt implicitly and under a great exposure to the TL seem to be promoted by an early onset age, whereas older learners display a greater rate of acquisition, thanks to their higher cognitive development and their explicit learning mechanisms. In sum, in FL settings, where opportunities for implicit learning and language use are minimal, early start learners will not benefit from their ‘potential advantage’ over older learners (p. 34). More recently, Muñoz (2014) provides further evidence for the effect of onset age and TL exposure on FL learning and concludes that the benefits purported by research on early start in SL settings do not seem to apply to FL contexts. What is more, her results seem to confirm that the amount and quality of TL exposure, which is obviously much lower in FL settings, have a greater effect on language learning than starting age.

Taking everything into consideration, the general consensus is that YLs and adults, both in FL and in SL settings, learn languages in different ways (Muñoz, 2006, 2007b; Nikolov, 2009; Nikolov and Mihaljevic Djigunovic, 2006). All the evidence presented above does not reject the existence of age effects on SLA, but indicates that there is not “[...] a sharply bounded interval in human development that is followed by a biologically determined incapacity to attain native like levels of proficiency in a second language” (Dimroth, 2008: 52), and implies that other age-related aspects must exert their influence (Birdsong, 2005a). Pinter (2011, 2017), concurring with García Mayo and García Lecumberri (2003) and Muñoz (2006, 2014), claims that, although age plays an important role in language acquisition, other variables such as opportunities to use the TL, learner motivation and quality of the formal instruction, also seem to have an effect on this process. In what follows, we will focus on the unique features of young language learners.

3.2 Young learners' distinctive characteristics

The age range within which learners are considered 'young' varies depending on the context in which this population is considered (Nikolov and Mihaljevic-Djigunović, 2011). In Europe, the setting of the current study, the member states agreed to differentiate between 'very young learners' (pre-school children between three to six years old) and YLs: "[...] primary-school pupils between seven to twelve [...], although in certain contexts even 14-year-olds are included in the YLs' group" (Nikolov and Mihaljevic-Djigunović, 2011: 96).

As already mentioned, the way children and adults learn a language is different. The main aspect that differentiates them is that, unlike adults, children are still developing in various ways. As Nicholas and Lightbown (2008) point out:

For young learners, language acquisition involves cognitive, social, and physical engagement over long periods during which many changes take place in the developing child (Nicholas and Lightbown, 2008: 46).

In other words, childhood is a period of many and frequent changes which, following Berk (2006), can be divided into four different stages: early childhood (ages 2-7); middle childhood (7-11), early adolescence (12-14) and later adolescence (15 years and older). The existence of these different stages draws attention to the importance of considering also the differences between younger children and older children. Thus, whilst some activities may be beneficial for a specific age group, others will not result in much improvement in their language development (Muñoz, 2007b; Nicholas and Lightbown, 2008; Pinter, 2006).

Philp, Mackey and Oliver (2008) provide a more detailed insight of the characteristics of children in the middle childhood stage, which is the age range we will study in the present dissertation. According to these authors, children at this age become more logical and able to categorise and organise objects, as well as capable of considering different aspects of a problem and thinking from different perspectives. During middle childhood, children can also take turns, talk about the same topic for longer than younger children, and are aware of the pragmatics of speech acts. With regard to language development, children at this age already possess a highly developed L1, which continues increasing in vocabulary and grammar complexity, whilst their metalinguistic awareness and their

ability to read and write are also developing. The last aspect to be referred to here is the fact that their context and their social experiences are different from those in early childhood (e.g. school, peer relations). All in all, during this stage, children's social and cognitive abilities are developing fast and undergoing important changes (Philp et al., 2008).

Taking these particular features of YLs into consideration, Muñoz (2007b) reviews age-related differences which are noticeable in the process of L2 acquisition and suggests that different learning activities are necessary for different ages. Apart from the already mentioned different cognitive developmental stage of children, their language aptitudes as well as their learning strategies are different. For instance, while adults rely more on their analytical capacity for language learning, YLs are more memory-oriented (e.g. use of language chunks). Thus, simpler activities related to the 'here and now' and to 'doing things' are considered suitable for younger children, whilst older children benefit more from activities that require more complex cognitive processes. Classroom activities should therefore be designed to make the most of the learners' different age-specific skills. Accordingly, assessment formats and procedure also need to meet YLs' cognitive and developmental stage (Butler and Zeng, 2014). Not taking this into account, that is, not matching the learners' cognitive maturity to the demands of a given task may lead to the failure (or to a not completely satisfactory outcome) of teaching practices (Muñoz, 2007b).

As Philp et al. (2008) briefly introduced in their description of the main aspects of middle childhood, another important factor that influences child SLA is the socio-cultural context. This context is different for younger learners, older young learners and adults and it has an effect on the way learners interact with their interlocutors (Nicholas and Lightbown, 2008; Philp et al., 2008). Different factors, such as the educational experiences that accompany maturation, influence the way learners interact, and consequently, the way they learn a language (Nicholas and Lightbown, 2008). Other variables are the interlocutors' age and/or status (e.g. teacher, peer, parents), as well as the learners' independence in their interactions. Pinter (2011) provides a division of the age groups within childhood which, in general terms, coincides with the ones presented above by Nikolov and Mihaljevic-Djigunović (2011), Berk (2006) and Nicholas and Lightbown (2008). As can be seen in Table 4, Pinter includes a brief description of the

different instructional contexts and capacities of these groups, providing further evidence for the existing differences within childhood.

Table 4 *Children: the three age groups (taken from Pinter, 2011: 2).*

Education	Age	Capacities
Preschool	3-5	No formal learning experiences. No literacy skills. Large differences among children with regard to readiness for school.
Primary school years	6-12	Primary/elementary schooling. Often divided into lower primary and upper primary years.
Early adolescence	13 onwards	Change of schooling to secondary or high school at around 11 or 12, but in some countries not until 14.

The changes in the socio-cultural context go together with YLs' different cognitive developmental stage, which, as already mentioned, is also reflected in their interactions (Philp et al., 2008). YLs, apart from being quite frequently still developing their L1 (Cameron, 2001), are also learning how to interact with other people. This particular behaviour is observed in how YLs are less constricted by social conventions and, therefore, by task conditions (Nicholas and Lightbown, 2008; Philp et al. 2008). Children are more likely to openly disagree with their partners, change the topic of the conversation, or even try to cheat when carrying out a task. Example 32 illustrates the participants' childish behaviour, which takes them away from the purpose of the task into a completely different discussion.

Example 32

- S1: Do you have any kangaroos on your paper?
S2: You looking
S1: No I didn't
S2: Not doing like that...see...you know it...because you look my one
S1: No
S2: I don't know what to do...stupid thing... [This one] is stupid.
S1: xx xx
S2: Not doing like that...just put it out... not doing like that
S1: xx xx
S2: Don't look
S1: I'm not looking...I'm looking...
S2: Just look there then
S1: OK
S2: I don't know what to do
S1: I'm not looking
S2: I don't know what to do
S1: Yeah what...See you want to do like...and you want to look like that
S2: I don't
S1: Ok...I look your one too...you can look my one
S2: (Makes frustrated, growl noise.) Not good

(Oliver, Philp and Mackey, 2008: 144)

The previous dialogue depicts a side-sequence to the main flow of the conversation, in which S1 is caught when looking at S2's picture instead of trying to perform the task orally, as they were supposed to. The two participants are arguing about how they should carry out the task and S2, after being caught, openly shows his or her frustration with the task. This type of behaviour is not likely to happen when adults perform a collaborative task. This interactional feature is related to another characteristic of child learners: their purported egocentric nature. YLs are believed to be less prone to care about their interlocutors' needs than about their own will to communicate something (Oliver, 1998, 2009). This has been typically argued in the interaction field based on the low rate of comprehension checks that children produce. In fact, it is believed that "[...] the ability

to understand each other's needs increases with age and is developed throughout the elementary school years and beyond" (Butler and Zeng, 2014: 50).

Nevertheless, a recent study by Lázaro Ibarrola and Hidalgo (2017a) suggests that the characterization of YLs as egocentric or not interested in their interlocutors' meaning when engaged in interaction might not be entirely adequate. In line with previous studies, these authors reported a very low rate of comprehension checks during young EFL learners. However, the inclusion of a more comprehensive analysis of their students' interactions revealed a different set of strategies that showed YLs' concern about each others' messages and how they assisted one another (e.g. by completing each other's utterances, acknowledging comprehension of their partner's output and using self-repetitions).

Finally, although we have outlined the main characteristics YLs have in common, it is important to note that, like among adult learners, there are individual differences among children, which have to be taken into account. Nikolov and Mihaljevic-Djigunović (2011) enumerate not only attitudes, motivation, learning strategies and aptitude, but also YLs' language anxiety. By the same token, the different levels of development children may have in particular aspects of their own L1, will also influence the way YLs acquire those areas of the L2 (Cameron, 2001).

As noted throughout this section, there are important differences between adult learners and YLs, which arise from children's different linguistic, psychological and social developmental stage (Cameron, 2001). The effect of age on language acquisition is believed to be mainly related to the socio-psychological variables described above as well as to learner-external factors such as the type and amount of input the learners are exposed to and their experience with the TL (Birdsong, 2005a; Mackey et al., 2003; Philp and Tognini, 2009; Pinter, 2011). The particular characteristics of YLs need to be taken into consideration when teaching this population, as well as when designing materials to be used in the language classroom: for an activity to be successful, it needs to be adequate for the learners' needs (Cameron, 2001; Muñoz, 2007b, 2014; Nicholas and Lightbown, 2008; Pinter, 2006, 2011). To illustrate the main characteristics of YLs described above, Table 5 summarises the main differences between this population and adults.

Table 5 *Main differences between adults and children.*

Different cognitive developmental stage.
Different socio-cultural context.
Childhood is a period of many changes, within which different stages can be identified.
Frequently still developing cognitively, socially and physically, as well as acquiring their L1 and learning to interact with others.
Less constricted by social conventions.
As with adult learners, individual differences must be considered.

Taking all this into account, and to further illustrate how children learn a second or foreign language by linking it to another of the main topics of the present study (the interactionist approach), some of the most relevant studies on child interaction will be presented in the following section.

3.3 Child interaction

As discussed in Chapter 1 on the Interaction Approach, research has demonstrated the beneficial effects of interaction on L2 acquisition (García Mayo and Alcón Soler, 2013; Mackey, 2007; Mackey, Abbuhl and Gass, 2012; Pica, 2013). Accordingly, a wealth of studies has analysed the different learning conditions that promote interaction (e.g. NS-NNS pairs, different task types or proficiency levels). Nevertheless, as we stated at the beginning of the current chapter, child SLA literature remains relatively scant when compared with that focused on adults and, consequently, there are not many studies dealing with child interaction (except for the already mentioned work of Oliver (1998, 2000, 2009) and García Mayo and her colleagues (Alcón Soler and García Mayo, 2008; García Mayo and Lázaro Ibarrola, 2015; among many others). This fact brings us back to the previous section of the chapter, where we already commented on the differences between young and adult learners' SLA (Mackey et al., 2003; Oliver, 2002) and the frequent, although inadequate practice of establishing parallelism between adult SLA and YLs' SLA (Butler and Zeng, 2014; Collins and Muñoz, 2016; Lázaro Ibarrola and Azpilicueta Martínez, 2015).

A growing amount of research has shown that YLs benefit linguistically from opportunities to interact, both with NS and fellow language learners whether adults or children (Oliver, 1998, 2009; Philp and Duchesne, 2008). As already discussed, interaction fosters opportunities for NoM, providing learners with comprehensible input and feedback on their production, as well as with occasions in which they need to modify their output in order to make themselves understood⁶ (Mackey et al., 2007; Oliver, 1998, 2009; Oliver and Mackey, 2003).

Some pedagogical collaborative tasks (such as information gap tasks or jigsaws) have become a very valuable tool to promote interaction in the language classroom. These types of tasks provide a context for interaction, in which NoM will most likely take place and, consequently, all the processes it triggers (i.e. production of comprehensible input and provision of feedback, as well as output modification). This way, children consolidate their emerging interlanguage by using the opportunities communicative tasks provide to experiment with the language they are learning. Example 33 illustrates the process of negotiation and output modification. Learner B fails to understand his or her interlocutors' initial utterance and starts a routine of NoM. By means of a clarification request (What?), this learner asks learner K to modify a previous utterance and pushes him/her to produce a more target-like one.

Example 33

K: One. How many girls can you see feeling [flying]?

B: **What?** (Clarification request)

K: How many girls you can see fly fling [flying]?

(Oliver et al., 2008: 132)

In what follows, we will offer an introduction to the aspects which are usually addressed in applied linguistic research when analysing L2 general performance: complexity, accuracy and fluency (CAF for short). As the most relevant studies which have addressed language learners' CAF in oral L2 production have already been discussed in Chapter 2, this section will focus on the measures of CAF used in four studies addressing TR and YLs. These specific CAF aspects of L2 oral production are the ones which have been

⁶ Nevertheless, some authors point at the possibility of peers acting as 'gatekeepers', that is, as elements that may prevent the other learner from learning (see Philp and Duchesne, 2008).

analysed in the current study as descriptors of our participants' performance and indicators of the learners' L2 development.

3.4 Measuring oral production during interaction: language proficiency variables (CAF)

L2 proficiency is a construct that has been claimed to be comprised by three main components: complexity, accuracy and fluency. The analysis of these three areas of general performance has become a very frequent way of measuring L2 development in SLA research studies. However, it was not until the 80s that a distinction was made between "fluent versus accurate L2 usage" (Housen and Kuiken, 2009: 461). In the mid-nineties, the last component of the CAF triad, complexity, was incorporated in Skehan's proficiency model (1996, 1998) in which CAF were considered the three main aspects of L2 proficiency for the first time.

As stated above, the analysis of CAF has received the attention of many researchers who have taken these as the main dimensions of L2 proficiency and performance (Housen, Kuiken and Vedder, 2012). CAF have been analysed to evaluate other factors such as the influence of task type, learners' age, type of instruction or learning context (e.g. Bret Blasco, 2014; Lynch and Maclean, 2000, 2001; Sample and Michel, 2014). Furthermore, these three areas are believed to be connected to cognitive processes that entail major changes in the learners' L2 interlanguage. According to Housen et al. (2012), complexity is associated with the internalization of new L2 elements, as it implies the development of more sophisticated and elaborate L2 systems. Accuracy on the other hand, corresponds to the learners' development of their ability to modify and restructure their L2 knowledge, by building a more targetlike interlanguage. Finally, a higher fluency is achieved through the 'consolidation and automatisisation' of L2 knowledge (Housen et al., 2012: 3).

Nevertheless, the study of CAF still has to face a number of challenges in order to become an uncontroversial research field. For instance, as seen in Chapter 2, some inconsistent results have been found across different studies, which on the other hand, sometimes cannot even be compared. According to Housen et al. (2012), this is due to the variety of definitions and measurements that different studies apply to the analysis of CAF. Consequently, common measures that allow for the comparability of the findings are essential. In order to shed light on this research gap, these authors claim that a unified

common definition of the terms is needed, as well as a clear and detailed description of the CAF measures used in the studies that analyse these L2 performance dimensions.

Housen et al. (2012) define accuracy as “[...] the extent to which an L2 learner’s performance (and the L2 system that underlies this performance) deviates from a norm (i.e. usually the native speaker)” (p. 4). These deviations are normally considered ‘errors’. When reviewing this aspect, the authors argue for the inclusion of appropriateness and acceptability of the learners’ output, and not to take into consideration only the standard TL norms. Linguistic complexity, on the other hand, is conceptualised as the “intrinsic formal or semantic-functional properties of L2 elements (e.g. forms, meanings, and form-meaning mappings) or [...] properties of (sub-)systems of L2 elements” (p. 4). When considering the last aspect of this triad, fluency, Housen et al. (2012) follow Skehan (2003), and define it as composed of three submeasures: i) Speed fluency (speech rate in Skehan (2003)), which refers to “the rate and density of the linguistic units produced” (p. 5). ii) Breakdown fluency, which accounts for the amount, the length and the location of pauses. iii) Repair fluency, in which false starts, misformulations, self-corrections and repetitions are considered.

Another challenge CAF research must attend to is the connection and interdependency of CAF elements and the cognitive mechanisms of L2 learning (Housen et al., 2012). This aspect has been addressed by two competing models: Skehan’s (2009) Trade-off Hypothesis, earlier known as Limited Attentional Capacity Model (Skehan, 1998) and Robinson’s (2001, 2005) Multiple Resources Attentional Model or Cognition Hypothesis.

As briefly described in Chapter 2, the Trade-off hypothesis is based on the assumption that human attentional resources are limited, and therefore, when task demands surpass the available resources, different aspects of L2 performance (CAF) may compete for them, making learners prioritise some over others, i.e. attending to one CAF area may take some of the attentional capacities off another (Skehan and Foster, 2001). At best, fluency might go together with complexity or accuracy, but not with both (Skehan, 2003). Thus, only the language dimensions receiving enough attentional resources will attain optimal performance, whereas those which are not allocated enough attention will not. Nonetheless, according to this model, the characteristics of either the communicative situation or the task will influence the allotment of the learners’ attentional resources on one aspect or the other. Task characteristics and conditions have selective effects, which

may influence more than one CAF dimension, thus attenuating the trade-off impact. Consequently, complexity and accuracy may both benefit simultaneously under specific circumstances, such as the combination of certain task characteristics and conditions (e.g. planning time, familiar vs. unfamiliar information, information flow) (Skehan and Foster, 2012). Because this slight modification of the original hypothesis, Skehan's model is now termed as the Extended Trade-off Hypothesis, as it now incorporates a wider range of influential task features (Skehan and Foster, 2012). Table 6 illustrates how different task characteristics promote different dimensions of L2 general performance.

Table 6 *Task characteristics' influence upon L2 performance (CAF) (Adapted from Skehan, 2003: 5-6).*

Task Characteristic	Influence upon performance
Structured tasks (i.e. clear time line or macro- structure)	Clearly greater fluency, tendency towards greater accuracy.
Familiar information	Greater fluency and greater accuracy.
Outcomes requiring justifications	Justifications lead to markedly greater complexity of language.
Interactive vs. monologic tasks	Interactive tasks produce markedly more accuracy and complexity, monologic tasks more fluency.

On the other hand, the Cognition Hypothesis (Robinson, 2001, 2005) posits that attentional resources are not limited and, consequently, they can be allocated on different aspects at the same time, making it possible for learners to improve different dimensions of CAF simultaneously (Robinson, 2001). Moreover, the Cognition Hypothesis assumes a correlation between complexity and accuracy, as both of them are regarded as determined by task demands. As seen in Chapter 2, according to this model tasks should be sequenced on the basis of increasing cognitive demands, so that they gradually approach the requirements of real-world tasks. The sequencing from simple to more complex pedagogical tasks promotes the automatization and more efficient organisation of the constituents of *complex task performance* (Robinson, 2003). As a result, an increase in task demands will foster greater complexity and accuracy. A brief summary of the main tenets of these two models is presented in Table 7.

Table 7 *Contrasting predictions made by Skehan's and Robinson's Hypotheses (taken from Skehan and Foster, 2012: 215).*

Trade-off Hypothesis	Cognition Hypothesis
When attentional resources are limited, there will be competing priorities in performance.	Task complexity leads to increased complexity and accuracy simultaneously.
Task characteristics can have selective influences which modify the effects of trade-off.	Language complexity and accuracy should correlate, and be mediated by difficulty of task.

As can be noted from the revision offered above, both the Extended Trade-off Hypothesis and the Cognition Hypothesis predict that both complexity and accuracy can be increased at the same time, although each of them offers different reasons. The first one suggests a “combination of task characteristics and task conditions” (Skehan and Foster, 2012: 215), whereas for the latter task complexity is the driving feature. Nevertheless, none of these two models has been proven to be more valid than the other (Housen et al., 2012).

Chapter 2 featured some of the most significant studies addressing the effect of TR on learners' L2 general performance, operationalised as CAF, which have shown how repetition facilitates L2 acquisition. However, most of the studies have focused on adult and adolescent populations, and only five recent ones have studied how TR influences YLs' L2 oral performance. Out of those five, only four offer a detailed description of the CAF measures employed in their analysis (Bret Blasco, 2014; García Mayo et al., in press; Lázaro Ibarrola and Hidalgo, 2017b; and Sample and Michel, 2014). As we have already reviewed their findings, in what follows we will narrow our focus to the measures these studies used in order to analyse the CAF of their participants' oral performance. Table 8 illustrates the differences in a more visual way.

Sample and Michel (2014), following Housen and Kuiken (2009), coded for measures of structural and lexical complexity, specific and global measures of accuracy, and fluency. In order to measure structural complexity, they considered the total number of clauses as well as the number of words per AS-units. Lexical complexity, on the other hand, was analysed using Guiraud's Index, which is calculated by dividing the number of different words (types) by the square root of the total number of words (tokens) ($\text{Types}/\sqrt{\text{Tokens}}$) and D value (type-token ratio against increasing token size). Accuracy was measured

analyzing the different categories of errors per AS-units (e.g. number of error-free clauses/AS-units, agreement errors/AS-units, article errors/AS-units and other errors/AS-units). Finally, fluency was measured considering the filled pauses/minute, the words/minute (per pair) and the time the participants needed to complete the task (per pair) (Sample and Michel, 2014: 29-30).

In order to measure her participants' L2 oral skills, Bret Blasco (2014) used slightly different CAF measures to the ones employed by Sample and Michel (2014). To analyse propositional complexity, Bret Blasco (2014) computed the total number of *units* produced at each data collection point. The unit considered in her study was designed for her research, and defined as “[...] a context-dependent meaningful utterance, grammatical or ungrammatical, which conveys one piece of information or idea.” (p. 125). Linguistic complexity was operationalised as i) syntactic complexity, examined as the percentage of coordinate and subordinate units, and ii) lexical complexity, which considers the percentages of nouns, lexical verbs and adjectives over the total production. Grammatical accuracy comprised i) global accuracy, i.e. the percentage of error-free units, and ii) accuracy of a specific grammatical feature, i.e. percentage of correct lexical verb forms. Finally, fluency was analysed as speed fluency, considering speech rate in words per minute. Moreover, Bret Blasco (2014) also took into account the percentage of L1 words over the total number of words as indicator of breakdown (dis)fluency.

García Mayo et al. (in press) also considered structural and lexical complexity, and concurred with the previous authors in the choice of some of these measures. Structural complexity was calculated as complexity by coordination: the percentage of clauses per C-unit (Bret Blasco, 2014) and lexical complexity was calculated with the D value, as in Sample and Michel (2014). Accuracy was measured as the percentage of error-free clauses per C-units. Two measures were employed to examine fluency: the rate of syllables per second and the rate of meaningful syllables per minute.

Finally, Lázaro Ibarrola and Hidalgo (2017b) measured complexity by coding the total number of words, clauses and AS-units. They calculated the ratio of words per AS-unit and the ratio of clauses per AS-unit. Clauses were the base unit for accuracy, which was analysed computing the ratio of error-free clauses per total number of clauses as well as the number of errors per clause. Fluency was equalled to speech rate and, unlike the two previous studies, calculated as the number of syllables per minute.

Table 8 *Different CAF measures.*

		Sample & Michel (2014)	Bret Blasco (2014)	García Mayo et al. (in press)	Lázaro Ibarrola & Hidalgo (2017b)
Complexity	Structural	Words/ AS-units	%coordinate measures %subordinate measures	Clauses/ C-units	Words/ AS-units Clauses/ AS-units
	Lexical	Guiraud's index D value	%nouns/ words %lexical verbs/words %adjectives/ words	D value	—
Accuracy	Global	Error-free clauses/AS-unit	Error-free units/ unit	Error-free clauses/ C-unit	Error-free clauses/clauses Errors/clause
	Specific	Agreement error/ AS-units Article error/ AS-units Other error/ AS-units	%correct lexical verbs	—	—
Fluency		Filled pauses/ minute Words/minute Time to complete the task	Words/ minute %L1 words	Syllables/ Second Meaningful syllables/ minute	Syllables/ minute

As can be seen from the three studies presented above (carried out very closely in time), there are difficulties comparing results, as different measures are used to analyse CAF. This may also be a reason why different, sometimes contradictory findings have been reported as regard these aspects of L2 performance (Robinson, 2007).

In summary, although some researchers have already attempted to provide a clear and unified definition of the three main dimensions of L2 performance (CAF), much work is still needed in to reach common measures of CAF that allow for a better comparability of the different studies, and eventually obtain more consistent results. Nevertheless, we

understand that the variety of measures used in the different studies may also depend on the type of task used and/or the learners' output. The same applies to the results obtained, as different task types trigger differences in the learners' oral production, that is to say, there are tasks that may not promote structural complexity or fluency.

In this section we have provided one of the most accepted definitions of the terms as well as offered a brief introduction to the two main frameworks which have triggered a great amount of research on CAF (The Extended Trade-off Hypothesis and the Cognition Hypothesis). Finally, the CAF measures employed by three different papers addressing YLs' oral L2 performance have been reviewed.

So far, in the present chapter we have been talking about research on YLs and their specific characteristics in general, without actually taking into account the setting in which learning takes place. Language learning differs in important ways depending on the context in which it happens (Cameron, 2001; Muñoz, 2006) and the status of the TL within the community in which it is learnt.

The next section considers two typical contexts for language acquisition: SL and FL learning contexts will be reviewed. In the last section of the current chapter, studies on child interaction involving oral communication tasks carried out in these two different language settings will be presented. It must be noted that most of the studies presented in these subsections are concerned with YLs of English as a second or foreign language. This is mainly due to the current dominant role of English globally and the increasing number of FL programmes for children all over the world (Collins and Muñoz, 2016; Enever, 2011, 2016; Pinter, 2011, 2017).

3.5 Second vs. foreign language acquisition

Although the internal (i.e., cognitive processes) and external (i.e., positive/negative input and output) factors that regulate language learning are present in both SL and FL contexts, there are important differences between the two settings (Alcón Soler and García Mayo, 2008). Conventionally, in an FL setting the TL is not the L1 in the country or region in question and is therefore not generally used as a means of communication by the local community, the media or the government or as the medium of instruction in schools. Thus, FL settings are characterized by low amounts of exposure to the TL and limited access to it outside the language classroom. FL acquisition therefore typically occurs

when a language is learnt mainly to communicate with foreigners, and usually in the learner's own country. On the other hand, in a SL context the TL is used by most of the population of a region, or plays an important role in the life of that particular area and its inhabitants (government, education, media), although it may not be the L1 of many of the people who use it (Shehadeh, 2012). Consequently, SL learners normally come from a variety of L1 backgrounds and receive larger amounts of exposure of the TL, in contrast to the limited exposure and relative homogeneity of L1s in FL learning settings (Loewen, 2015). For instance, learning English is described as FL learning in Spain, France or Germany. Conversely, it would be considered a SL in the United States or Great Britain, where English is the L1 of the wider population.

FL contexts have been defined as 'low input contexts', as opposed to 'high input' SL settings, where learners are exposed to large amounts of input (Pinter, 2011). Nowadays, it is interesting to point out that, if we stick to the case of English, EFL learners have relatively easy access to the TL by means of television, the internet, video games, music, etc. Still, the amount of exposure or the opportunities to interact in the TL language EFL learners have will never equal that of ESL learners (Philp and Tognini, 2009), as people in FL contexts do not use English for everyday life communicative purposes (Cameron, 2001). Table 9 summarizes the main characteristics of FL and SL learning.

Table 9 *Contrasting foreign and second language contexts (taken from Pinter (2011: 87)).*

Foreign language	Second language
Low level of input: Typically 1-3 hours a week timetabled lessons.	High(er) level input: More than just a timetabled lesson.
No/restricted opportunities outside class to use the target language.	Regular opportunities to use the target language outside class.
Focus is on language as a formal system and as a subject.	Focus is on content and language integrated across the curriculum.

Connecting the specific features of these two language contexts to the field of language learning, one of the main differences between FL and SL learners is that, since in FL contexts learners do not have as many opportunities as SL learners to use the TL, these learners may not feel the same need to learn it (Alcón Soler and García Mayo, 2008). Another aspect, just as important, is the fact that in a FL context most learners share their L1. The existence of a predominant language, which is not the one being taught, makes the TL not essential for real communication. One consequence of the previously

mentioned aspects is the possibility that the TL may become an object of study which lacks connection with the real world and, consequently, it may be separated from its real communicative function. Finally, on the part of the FL teacher there are also some difficulties to overcome, such as the need to motivate learners to use the TL and to make them aware of its usefulness.

As noted in this short section, there are important differences between SL and FL settings which have an effect on the way a TL is learnt. The specific characteristics of these two language learning contexts, above all the low level of input and the restricted opportunities to use the TL outside the FL classroom, provide further evidence for the inadequacy of transferring the reported benefits for language acquisition in SL contexts to FL settings. For instance, the case of immersion programmes we referred to at the beginning of this chapter which should not be considered as evidence supporting the implementation of early start programmes in FL contexts without further empirical investigation (Muñoz, 2014). In the next two subsections, studies on YLs interaction both in FL and SL settings will be presented. Tables 10 and 11 list the studies that will be discussed and briefly present information about the age and number of participants, the TL, the task used in the study, the research goal and the results obtained.

Table 10 Summary of the findings on the effects of task-based interaction among YLs in SL settings.

Authors	Age & L1	Task & TL	Goal	Results
Van den Branden (1997)	11-12 (N=48) Mainly Moroccan	One-way picture description task. Dutch	To determine whether YLs negotiate for meaning, and see if their negotiation varies depending on the interlocutor (peer/researcher).	YLs negotiate for meaning. Negotiation of form in researcher-learner dyad. (-) Complexity & grammaticality.
Oliver (1998)	8-13 (N=96) Variety of L1s	One-way picture description task. Jigsaw. English	To determine whether YLs can negotiate for meaning, the strategies they use and if these differ from adults'.	YLs negotiate for meaning: YLs use a variety of strategies, although in a lesser proportion.
Oliver (2002)	8-13 (N=192) Variety of L1s	One-way picture description task. Jigsaw. English	To study the effects of corrective feedback on YLs' oral performance.	YLs seem to benefit more rapidly from interactional corrective feedback than adults.
Oliver (2009)	5-7 (N=32) Variety of L1s	One-way picture description task. Jigsaw. English	To determine whether very young learners negotiate for meaning and use corrective feedback in a similar way to older YLs.	Very YLs negotiate for meaning and provide feedback: Similar amount of NoM strategies. Younger YLs (+) self-centred.
Mackey & Oliver (2002)	8-12 (N=22) Variety of L1s	Spot-the-differences. Story completion. PPT. Picture sequencing English	To test the effects of interactional feedback on YLs' L2 development.	Interactional feedback benefits YLs SLA. It seems to have a more rapid effect than on adults.

Mackey, Oliver & Leeman (2003)	8-12 (N=96) Variety of L1s	One-way picture description task. Jigsaw. English	To investigate the effects of the interlocutor type on feedback provision and actual use in task-based interaction.	Adults provide more opportunities for modified output whilst YL dyads produce more modified output.
Mackey, Kanganas & Oliver (2007)	7-8 (N=40) Variety of L1s	Picture description. PPT. English	To examine the effect of task familiarity on interactional feedback.	Unfamiliar tasks → (+) negotiation & feedback Familiar tasks → (+) actual use of feedback
Oliver, Philp & Mackey (2008)	5-7/11-12 (N=22/20) Variety of L1s	5 information gap tasks. English	To analyse the effect of different types of teacher guidance.	Teacher guidance in general promoted interaction. On-task guidance: (+) benefits for older children: (+) modified output.
Mackey & Silver (2005)	6-9 (N=26) Chinese	Communicative tasks: meet-your-partner, Spot-the-differences, PPT, Picture sequencing. English	To test whether task-based interactional feedback promotes ESL for children in a multilingual context.	Interactional feedback benefits question formation, and SLA.
Gagné & Parks (2013)	Grade 6 (N=29) French	Numbered Heads Together Jigsaw, Round-robin. English	To analyse whether YLs scaffold each other when performing cooperative tasks.	YLs scaffold each other providing a variety of strategies. (!) NoM strategies used rarely.

Table 11 Summary of the findings on the effects of task-based interaction among YLs in FL settings.

Authors	Age & L1	Task & TL	Goal	Results
Pinter (2006)	10 (N=10 adults & 10 YLs) Hungarian	Spot-the- differences. English	To compare task-related strategies of adults and YLs.	YLs & adults complete tasks differently. Adults: (+) focus on details; YLs keep solving of misunderstandings to a minimum.
Pinter (2007)	10 (N=2) Hungarian	Spot-the- differences. English	To analyse the effect of TR on YLs' peer-peer oral performance.	YLs assist each other throughout the tasks. Repetition promoted: (+) fluency. (+) relax & confidence.
Bagheri, Rahimi & Riasati (2012)	Age: n.a. N= n.a. Persian	Spot-the- differences English	To analyse the effect of TR on YL's peer-peer oral interaction.	YLs assist each other. (+) efficient & confident task completion. Provided opportunities to use the TL.
Butler & Zeng (2014)	9-10/11-12 (N=64) Chinese	Information gap & decision making. English	To report on the benefits/limitations of task-based assessment on YLs by identifying developmental differences in different age groups.	Younger learners use a smaller variety of interactional functions. Task-based assessment seems to be less useful with this age group.
Lázaro Ibarrola & Azpilicueta Martínez(2015)	7-8 (N=16) Spanish	Guessing game. English	To document YLs' interaction strategies.	YLs in EFL settings do negotiate for meaning, but less than ESL YLs. YLs' L1 use is scarce.
García Mayo & Lázaro Ibarrola (2015)	8-9/10-11 (N=80) Spanish	Spot-the- differences. English	To analyse the effect of CLIL and EFL on NoM, as well as to compare different ages.	CLIL: (+) CAs & (-) L1. Older YLs: (-) CAs & (+) L1.

Azkarai & Imaz Agirre (2016)	9-12 (N=72) Spanish	Guessing game & PPT. English	To analyse the effect of CLIL and MS on conversational strategies, as well as to compare different ages and tasks.	YLS in EFL settings do negotiate for meaning, but less than ESL YLS. Younger learners: (+) NoM MS: (+) NoM.
García Mayo & Imaz Agirre (2017)	8-9/10-11 (N=54) Spanish	PPT. English Longitudinal	To examine the effect of the teaching approach (MS & CLIL) and of time.	Younger MS learners: (+) CAs. Age: CAs decreasing trend in both groups.
Lázaro Ibarrola & Hidalgo (2017a)	11 (N=40) Spanish	PPT. English	To examine YLS use of interactional strategies when performing an oral communicative task.	Most frequent strategy: Repetitions. Clarifications req > Confirmation checks. Comprehension checks very scant. Little L1 use. (!) Abundant structural transfers.
Lázaro Ibarrola & Hidalgo (2017b)	11 (N=20) Spanish	PPT. English	To measure the effect of TR on the oral production of YLS in a CLIL program.	(-) Repetition (+) Accuracy

3.5.1 YLs' task-based interaction in SL learning contexts

In this section we will provide a review of research on interaction relevant to SLA undertaken with children. Taking the findings of the studies reviewed into account, we will conclude by describing the main characteristics of interaction among YLs in SL contexts.

Studies by Oliver (1998, 2002, 2009), show how ESL children are able to engage in conversation cooperatively, using similar strategies to those identified in adult studies (coinciding also in the most frequent one, i.e. repetitions) during interaction. Nevertheless, this author also reported some differences between the way YLs and adults interact. In one of Oliver's earliest works (1998), she studied the oral interactions of 96 age- and gender-matched pairs of young ESL learners (aged 8 to 13) while carrying out two communicative tasks. The results of this study provide evidence on how YLs interact and use CAs when negotiating for meaning. Some differences with regard to the way adults interact are related to the smaller amount of negotiation strategies appearing during child interaction, and the fact that YLs use very few comprehension checks. This latter characteristic is attributed to the developmental stage of the participants. Oliver explains this finding on the basis of the egotistic nature of children who, as already mentioned, seem to be more prone to expressing their own meaning than to focusing on their conversational partner's needs.

Oliver (2002) analysed the effect of nativeness and proficiency on the interactions of learners the same age as in the previous study (i.e. 8-13 years old) when performing two communicative tasks. The findings suggest that these two factors have an effect on the amount of NoM that takes place. With regard to proficiency, the overall tendency is that the lower the proficiency level of the learners, the more the children negotiate for meaning. When looking into the effect of nativeness, the results mirror those obtained in adult studies and go in line with the results for proficiency level, that is, NNS-NNS pairs seem to produce the greatest number of negotiation strategies, whilst NS-NS dyads negotiate the least.

Also in a SL context, Van den Branden's (1997) provided further positive evidence on how 11-12 year-old learners of Dutch in a Flemish school negotiate the meaning and content of their production and are able to modify their output in order to make it more comprehensible. He also found that YLs are prone to recycle the new language they have

acquired. Nevertheless, the negotiations did not show significant improvements on the complexity nor the grammaticality of the learners' output, as the participants did not seem to focus much on language form and only negotiated when working with the researcher.

The study by Mackey and Oliver (2002) already been reviewed in Chapter 2, is concerned with ESL learners and focuses on the effects of corrective feedback on the oral performance of YLs and adults. This study presents more supporting evidence for the benefits of interactional feedback on language development for children, and the authors suggest that YLs benefit more rapidly from it than adults do. This confirms once more the importance of not generalizing adult findings to YLs' language acquisition without empirical evidence.

Sharing the concern of the previous study with corrective feedback, Mackey et al. (2003) compare children and adult interactions in the TL (English) to analyse the effect of age on the provision and incorporation of implicit negative feedback. Their results suggest that, although no significant difference on the amount of feedback was found, adults seem to provide their interlocutors with more opportunities for modified output whilst NNS/NNS YL dyads actually produce more modified output.

Another interesting study with YLs in an ESL context is Mackey et al. (2007), already referred to in Chapter 2. The authors draw attention to a different aspect and manipulated familiarity with the content and with the procedure of several communicative tasks. They analysed the effect of familiarity on the amount of NoM and provision and use of feedback provided by 7-8 year-old ESL learners. As already discussed, their main finding was that there seems to be more NoM and feedback provision during unfamiliar tasks, whereas when the task is familiar, learners actually seem to attend more to the feedback received.

Oliver et al. (2008) focus on the effect of the interlocutor. In this study, the interaction of two groups of YLs (aged 5-7 and 11-12) when performing tasks in the language classrooms was analysed. The researchers examined the effect of the teachers' input under three different conditions: task instructions, task instructions with models and task instructions plus within-task guidance. Their findings suggest that the three conditions promote interaction and consequently, are beneficial for language learning. Nevertheless, they also report that older children appear to benefit most from on-task guidance, since this age group seems to produce more modified output under this condition. The authors

conclude highlighting the importance of the effect of both the teacher's role and learners' age on the outcome of task-based interaction.

In a multilingual and multicultural setting such as Singapore⁷, Mackey and Silver's (2005) study provided further positive evidence of the benefits of interactional feedback for language development. They analysed the oral interaction of 26 Chinese YLs of English, aged 6 to 9, while they performed different communicative tasks with native English adults. They had two groups under two conditions: an 'Interaction and feedback group' and an 'Interaction control group'. The experimental group received interactional feedback on their non-targetlike questions whilst the control group did not. In their study, the group receiving interactional feedback improved accuracy in terms of question formation. With respect to the context in which this study was carried out, Mackey and Silver (2005) note that in non-Western educational contexts, interactional activities among peers are not usually considered as worthwhile or even appropriate. Some practitioners have even stated their concerns about communicative methods and consider their efficacy as limited. Therefore, the attitude of the learners may be different to that of western students, and it may influence the effect of interactive tasks. Thus, in order to test the effectiveness of the Interaction Approach in these settings, more empirical research is needed (Mackey and Silver, 2005).

More recently and again in an ESL setting, Oliver's (1998) findings were further supported by Oliver (2009) when she compared very young ESL learners (age 5-7) to 'older' young ESL learners (8-13). She found that younger learners are also able to negotiate for meaning and use the same negotiation strategies as older YLs. Nevertheless, a lower percentage of use of comprehension checks and other-repetition was reported. According to this author, these results suggest that the younger the speakers are, the less concerned they are about others' needs (Oliver, 2009).

The most recent study we are going to review here is Gagné and Parks (2013) on YLs' ability to provide assistance to their peers. The authors considered whether young ESL learners were able to provide linguistic scaffolding to each other when performing collaborative oral tasks. A variety of scaffolding strategies were reported (requests for assistance and other-correction being the most frequent ones), although "[...] the

⁷ There is an ongoing discussion on whether English should be considered a FL or a SL in Singapore. However, as this lies outside the scope of the present dissertation, I will not go into detail about it here.

strategies typically associated with negotiation of meaning within an interactionist perspective were rarely used” (Gagné and Parks, 2013: 1). Interestingly, the authors stated that clarification requests were used in 9% of the instances whilst comprehension checks were not used at all, partially supporting the results reported in earlier ESL interaction studies (e.g. Oliver, 1998).

As noted above, research on YLs SLA has revealed that children do negotiate for meaning, and are able to accomplish their role as interlocutors. Moreover, children have been shown to use a variety of negotiation strategies in their interactions (Mackey et al., 2007; Mackey and Oliver, 2002; Oliver, 1998, 2009). However, strategies typically associated with adult interaction (comprehension checks) are rarely used by YLs, or used far less than by adults (Gagné and Parks, 2013; Oliver, 1998; et passim). Another interesting point to be observed here is the fact that most studies have been carried out in ESL contexts and only one of the studies reported considers the learning of a different SL (i.e. Dutch). English language teaching at primary education levels is steadily increasing in a wide range of contexts and English is being consolidated as part of the primary curriculum (Enever, 2016). Consequently, a significant amount of the research is concerned with this language. Of course, there is research outside the ESL realm. Nevertheless, this is an issue that lies outside the scope of the present investigation.

3.5.2 YLs’ task-based interaction in FL learning contexts

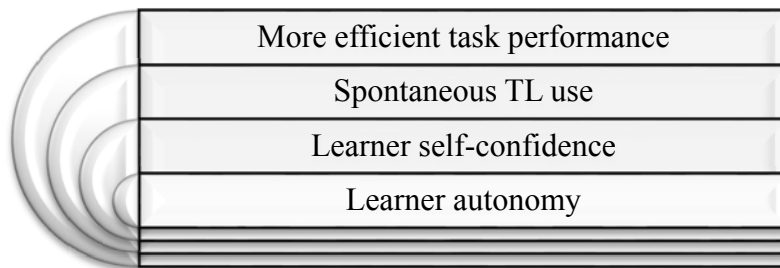
In the final section of this chapter, recent findings in child interaction in FL settings will be presented in comparison to the previously discussed research on SL contexts. Although still not as abundant as SL literature, a growing body of research on child oral interaction in FL settings has developed in recent years, providing quite positive results which partially support the findings reported in SL studies. In general, the main finding in interaction research with FL children is that YLs in these settings are also capable of interacting and negotiating for meaning, although they do this at a lesser extent than SL children. YLs in FL settings are a very interesting population, since foreign languages (above all English), are being taught in schools all over the world to an enormous number of learners and this number still continues to expand (Cameron, 2003; Collins and Muñoz, 2016; Enever, 2016; García Mayo and Lázaro Ibarrola, 2015; Lázaro Ibarrola and Azpilicueta Martínez, 2015; Nikolov and Mihaljević-Djigunović, 2006, 2011; Pinter,

2007). In what follows, a review of relevant literature of the effects of interaction on child FL learners will be presented.

In the section devoted to TR in Chapter 2, two studies addressing YLs have been reviewed. Pinter (2006) compared the oral interaction of 10 pairs of 10 year-olds and 10 pairs of adult Hungarian EFL learners with a very low level of command of the TL while completing a spot-the-differences task. She found that YLs interact in a similar way to adults (they are able to collaborate with each other as well as to clarify their messages), although children do so in a smaller proportion. The author maintains that children and adults understand and complete tasks in different ways, a fact that needs to be taken into account by practitioners and researchers; while adult learners seem to have more efficient task-related strategies at their disposal, children do not follow an apparent order and use less time and language to complete the task. Pinter (2007, 2011) analysed the oral performance of a pair of EFL YLs in the same context and with the same characteristics as the ones described for the previous study (i.e. Pinter, 2006). The learners completed the task (with slight variations of the content) three times over a period of three weeks. The results that show TR, together with peer interaction, is beneficial for YLs, even with a low command in the TL. Through TR, learners gained confidence to use the TL and became more fluent, at the same time that they had the opportunity to notice mismatches between their production and the TL.

Bagheri, Rahimi and Riasati (2012) provide additional support for the benefits of TR and peer-peer interaction within the same age group at a limited level of competence. These authors analysed the effect of TR on the interactions of Iranian YLs with a low level of proficiency in English. They carried out a spot-the-differences task three times over a three-week period. Their results mirror Pinter's (2007) and show that, by the last performance, the learners carried out the task more efficiently and with more confidence. The tasks also provide an opportunity to enjoy the use of the TL in a spontaneous way and the autonomy to work with a peer without teacher intervention. The authors concluded that YLs peer-peer interaction "[...] offer(s) multiple benefits to learners and practicing with similar tasks is an effective way of encouraging these positive changes to take place" (Bagheri et al., 2012: 951). The reported benefits of TR and peer-peer interaction, regardless of CAF, are summarised in Figure 2.

Figure 2 *Benefits of TR and peer-peer interaction.*



Another very interesting construct which, to our knowledge, has not received much attention is how teachers assess YLs task-supported interaction (Nikolov and Mihaljevic-Djigunović, 2011). Butler and Zeng (2014) analysed the developmental differences in YL interactions when engaged in task-based language assessment in order to identify the benefits of this type of evaluation. The participants in this study were two groups of YLs of EFL, fourth-grade 9-10 year-olds and sixth-grade 11-12 year-olds in a school in China. The main differences found between the two groups were that: “[...] the fourth-grade dyads showed less mutual topic development, used formulaic turn-taking more frequently, and had a harder time taking their partners’ perspectives during tasks.” (Butler and Zeng, 2014: 45). The authors concluded that, because of the younger learners’ limited use of interactional functions, task-based assessment may not be a very valid tool for this age group.

Having considered some of the studies in an international context, we now look at the Spanish context, where the current study is set. In the last years, a number of studies on child interaction have been carried out by García Mayo and her colleagues. These researchers have conducted a four-year research project in Spain (FFI2012-32212), and have analysed the oral interactions of YLs (aged 7 to 12) in three different schools in the Basque Country and Navarre. The schools followed two different English teaching approaches: MS and CLIL (see Chapter 4 for a detailed description of this pedagogical approach which is becoming prevalent in Europe (Enever, 2011)). Specifically, the researchers have focused on YLs’ negotiation for meaning and form, their use of interactional feedback and of their L1, as well as on the effect of the teaching methodology and age. In what follows, some of the results obtained by the members of this research group are listed.

Lázaro Ibarrola and Azpilicueta Martínez (2015) observed the performance of 7-8-year-old EFL learners with a very low level of proficiency in the TL (English) while carrying out a guessing game. Supporting the findings reported in Oliver's studies, their results show that EFL YLs use the same negotiation strategies as young ESL learners (except for comprehension checks), although in a smaller proportion. Their study provides valuable evidence of how young EFL learners, despite their low command of English, are able to interact in the TL and do so avoiding the use of their L1 throughout the task.

The effect of the methodological approach (MS and CLIL) on the NoM of 20 pairs of young EFL learners in Spain has been investigated by García Mayo and Lázaro Ibarrola (2015). Additionally, the potential influence of the participants' age (namely 8-9 and 10-11) was also addressed. Their findings concur with those reported in ESL contexts as well as with those reported by Lázaro Ibarrola and Azpilicueta Martínez (2015), and illustrate how EFL children can negotiate for meaning and use different negotiation strategies. The authors observed that the CLIL groups used more CAs and repetitions and relied less frequently on the L1 than the MS groups. This was attributed to the fact that CLIL students are more used to communicating in the TL, as well as to their slightly higher command of the language. When comparing the different ages, the older groups were found to use fewer CAs and more the L1 than their younger counterparts. This may be due to the developmental level they were approaching, as well as to their ability to participate and follow a conversation without much NoM. However, as regards the connection between proficiency and L1 use, García Mayo and Lázaro Ibarrola (2015) suggest a more complex explanation: a possible lack of interest or motivation may be the cause, although this question is left open to further investigation.

Azkarai and Imaz Agirre (2016) carried out a follow-up study to that of García Mayo and Lázaro Ibarrola (2015). Their participants (n=72) were the same YLs as in García Mayo and Lázaro Ibarrola (2015) only a year later. Azkarai and Imaz Agirre (2016) also focused on the conversational strategies (CAs, repetitions and L1 use) YLs of different ages (9 and 12 specifically) use in two different learning contexts, CLIL and MS. Nevertheless, these authors employed two different tasks (i.e. guessing game and PPT). Reflecting the results of previous studies, these children negotiated for meaning and used a variety of conversational strategies to overcome communication breakdowns. Also in line with previous research, the most frequent strategies were clarification requests, confirmation checks and repetitions, whereas comprehension checks were barely used (García Mayo

and Lázaro Ibarrola, 2015; Lázaro Ibarrola and Azpilicueta Martínez, 2015; Oliver, 1998, 2000, 2009). Concurring with García Mayo and Lázaro Ibarrola (2015), the younger learners (age 9) employed significantly more NoM strategies in the two instructional settings. The authors attribute this to the possibility that the task was too easy for the older learners, thus requiring fewer negotiation moves. Their results regarding the differences between the two instructional settings differ, however, from those obtained by García Mayo and Lázaro Ibarrola (2015). In Azkarai and Imaz Agirre (2016), MS learners used certain conversational strategies significantly more than CLIL learners. Since the participants in their study were a year older, the authors suggest that the CLIL learners may already have gained the necessary skills to fulfil the task without much difficulty, whereas the MS learners had further developed their ability to negotiate for meaning and still experienced a greater need to overcome difficulties. Finally, as regards the effect of the task type, the most frequent conversational strategy used in the guessing game was clarification requests, whilst no conclusive results were reported when analysing the PPT. It may be concluded that task type, as well as age and instructional settings exert an important influence in the conversational strategies language learners use.

As already reviewed in Chapter 2 on tasks, Lázaro Ibarrola and Hidalgo (2017b), also worked with YLs in a CLIL school in Spain and, using the same PPT as García Mayo and Lázaro Ibarrola (2015), analysed the effect of procedural repetition on the oral performance of a group of YLs (n= 20). A decrease in the use of repetitions was reported in the last performance, as well as a small improvement in the participants' accuracy. In their study, and providing further support for the findings reported in previous FL research, the amount of negotiation strategies was comparatively smaller than in SL studies and the L1 use of these participants was also scant (García Mayo and Lázaro Ibarrola, 2015; Lázaro Ibarrola and Azpilicueta Martínez, 2015).

As mentioned in Chapter 1, to the best of our knowledge, García Mayo and Imaz Agirre (2017), whose findings regarding L1 use have been previously reviewed, is the only longitudinal study that has examined NoM in an EFL context comparing the influence of the pedagogical approach (CLIL vs. MS). Their results suggest that younger learners (age 8-9) in MS employ these strategies significantly more frequently than their CLIL counterparts, concurring with Azkarai and Imaz Agirre's (2016) findings in their cross-sectional study. However, one year later, at the second data collection time, no statistically

significant differences between the groups were found. The effect of testing time changed as a function of the learning setting: whilst the use of conversational strategies significantly dropped among the CLIL learners, the decrease in the MS learners was not statistically significant.

Finally, the study already presented in the section on L1 in Chapter 1 by Lázaro Ibarrola and Hidalgo (2017a) addressed young EFL learners' ability to carry out a PPT and their use of interactional strategies. As with previous research, their findings show how YLs in FL settings are able to use negotiation strategies, although less frequently than adults and ESL YLs. Also in line with previous studies, corrective feedback was barely used, whereas repetitions were the most frequent strategy and clarification requests the most used CA, followed by confirmation checks (Gagné and Parks, 2013; Mackey et al., 2007; Oliver, 1998, 2002, 2009). Although comprehension checks were rare, other strategies were identified that show the participants' interest in their partner's task performance: utterance completion, acknowledgements and mere self-repetitions. This finding suggests that YLs are in fact concerned about their partners' comprehension and that, as the authors state, "the lack of comprehension checks cannot be interpreted exclusively as a sign of egocentricity or disinterest in their partner's meaning" (Lázaro Ibarrola and Hidalgo, 2017a: 98).

3.6 Conclusion

A growing body of empirical finding provides solid evidence for the important role the age of language learners plays in SLA. YLs and adults learn a language in different ways. This is because of the different cognitive developmental stage YLs and adults are at, which causes differences in the way their interlanguage develops. For instance, YLs are frequently still developing their L1, which influences how they acquire some L2 aspects. Further, YLs seem to develop their interlanguage more slowly than adult learners, and therefore they need to receive larger amounts of TL input. Depending on their age, learners have access to different language learning strategies: YLs appear to benefit from implicit learning mechanism that develop under great amounts of exposure to the TL, whilst adult learners have access to explicit language learning mechanisms which grant them a higher rate of acquisition (Muñoz, 2006). On top of the clear differences between children and adults, research has also considered the importance of the different stages within childhood (Berk, 2006; Nicholas and Lightbown, 2008; Philp et al., 2008; Pinter,

2011), which should be taken into account when designing and implementing classroom activities in order to achieve the highest linguistic outcomes. As for the CPH, the literature indicates that, although childhood is indeed a sensitive period for language learning, no conclusive evidence exists to support the existence of such clearly delimited period for language learning. Additionally, as mentioned above, this period (i.e. childhood) is influenced by several external factors.

Apart from the cognitive differences, socio-cultural contexts also play an essential role in child SLA acquisition. Instances of factors affecting children's language acquisition are their experiences at school and with adults, as well as their relationships with other children. External factors such as the quality of formal instruction and the language learning context, i.e. SL or FL settings, have an important effect on the way children learn a language (Muñoz, 2014). Other variables affecting the way YLs acquire an L2 can be found within the individual differences among children, such as motivation, aptitude and learning strategies (Nikolov and Mihaljevic-Djigunović, 2011).

As seen throughout this chapter, empirical evidence suggests that YLs in both SL and FL contexts benefit from interaction in general and they are able to negotiate, fulfilling their role as interlocutors with peers, adults, NS or NNS. Research has also shown how the influence of age and interlocutor type is in fact noticeable (Oliver, 1998, 2002; Mackey et al., 2003; Van den Branden, 1997). When compared with adult learners, YLs make use of all previously identified negotiation strategies, but they rarely use comprehension checks.

Contrasting YLs' performance in FL and SL settings, the main finding in interaction research has been that, although young FL learners are also capable of interacting and negotiating for meaning, they negotiate to a lesser extent than their SL counterparts (García Mayo and Lázaro Ibarrola, 2015; Lázaro Ibarrola and Azpilicueta Martínez, 2015). Another big difference is that children learning a FL, as they normally share the L1, sometimes rely on it as a communication strategy (García Mayo and Lázaro Ibarrola, 2015; Lázaro Ibarrola and Azpilicueta Martínez, 2015; Lázaro Ibarrola and Hidalgo, 2017a; among many others).

However, and in spite of the reported benefits of interaction for YLs, there are some researchers who, anchored in the unique features of children (i.e. different social and cognitive developmental level from adults), have questioned the validity of interaction,

and the positive effects it has on YLs' L2 acquisition. For instance, considering that collaboration is essential in interactive tasks, speakers' awareness of their interlocutors' needs is also necessary. As has been acknowledged, although YLs are able to interact and participate in a conversation, their ability to cope with the demands of interaction (e.g. understanding others' needs) seems to increase with age (Butler and Zeng, 2014; Pinter, 2006). Moreover, Oliver (1998) stated that "[...] primary school children focus on constructing their own meaning and less on facilitating their partners' construction of meaning" (Oliver, 1998: 379). This may imply that YLs might not benefit from the opportunities provided by interactive tasks as much as adults do (but see Lázaro Ibarrola and Hidalgo (2017a)). Another limitation is that YLs lack a focus on form because of their cognitive developmental stage, which may prevent them from benefiting from the advantages of this type of interaction as much as adults do, above all in terms of accuracy (Van den Branden, 1997). Along the same line of thought, Lyster (2001, 2004) states that, although YLs may be able to interact, they are not able to notice the implicit corrections that appear in feedback the way adult learners do and that YLs seem to need more explicit directions and error centred instructions (although see Oliver and Mackey (2003)).

Finally, one more aspect to take into account when dealing with YLs' interaction is the fact that childhood is not a homogeneous period. Consequently, a specific type of interactive task that promotes SL or FL development in children at the middle-childhood period, might not be appropriate for younger children (Lázaro Ibarrola and Azpilicueta Martínez, 2015; Muñoz, 2007b, 2014; Pinter, 2006).

All things considered, despite the limitations and based on the increasing positive evidence reported about the relationship between interaction and child SLA, the use of communicative tasks is recommended by many SLA researchers in the second and foreign language classroom (García Mayo and Lázaro Ibarrola, 2015; Lázaro Ibarrola and Azpilicueta Martínez, 2015; Lázaro Ibarrola and Hidalgo, 2017a,b; Mackey, 2012; Oliver, 1998, 2009; Oliver and Mackey, 2003; Pica, 2013; Pinter, 2007; among many others). Consequently, it is essential that the unique characteristics of these language learners and their specific needs are taken into consideration in order to design and implement an approach to teaching YLs that makes the most of their skills and their possibilities to learn an L2 (Nicholas and Lightbown, 2008).

CHAPTER 4

CONTENT AND LANGUAGE INTEGRATED LEARNING (CLIL)

As noted throughout the current dissertation, second and foreign languages are learnt in a variety of circumstances (e.g. degree of exposure, teaching approaches) (Llinares, 2015). However, most of the literature deals with the acquisition of an SL by adult learners. Moreover, little research on YLs with low levels of proficiency in the TL has been carried out from the interaction approach perspective, and even fewer studies deal with contexts in which a FL is both the object and the vehicle of instruction, that is, in what is commonly known in Europe as CLIL programmes. In order to shed light into this research gap, in our study we have worked with learners who attend a school where CLIL is fully implemented. Therefore, we believe that a review of CLIL is necessary in order to gain a better understanding of our participants' learning context.

Taking this into account, this chapter provides an overview of the research conducted in CLIL settings, addressing the learners' acquisition of the TL and content knowledge. First, we focus on the emergence and spread of CLIL programmes across Europe. After that, the predecessors of European CLIL will be introduced: Canadian immersion programmes and American Content-Based Instruction (CBI). Based on the similarities and differences between these approaches, a definition of CLIL will be provided, which will be supported by some of the most relevant research findings on learning outcomes in CLIL programmes. This will be followed by the main features of CLIL in Spain, together with an overview of the regions in which CLIL research has been prolific. Finally, some of the reported limitations of this teaching approach will be presented.

4.1 Emergence and spread of CLIL in Europe

The implementation of CLIL programmes has undergone a fast spread all over Europe⁸, which is the context in which the current research study is set. The rapid increase of CLIL programmes has been driven both by the European educational authorities as well as by parents and teachers, who were at the same time motivated by widespread language

⁸ In other contexts, such as the Latin American and Chinese contexts, similar processes can be observed (see Lim and Low, 2009). Unfortunately, the analysis of those settings goes beyond the scope of our investigation.

learning beliefs, such as the already discussed “the younger, the better” (Dalton-Puffer, 2011; Dalton-Puffer, Nikula and Smit, 2010).

Since the European society is becoming more international, it requires ever better educated citizens to be part of an international workforce, with a high command of FLs. Dalton-Puffer (2011) sees CLIL in Europe as “[...] a way to transcend the perceived weaknesses of traditional foreign language teaching” (p.185). Hence, the rapid spread of CLIL is driven by two converging forces: i) *Reactive reasons*, which are related to the need to solve the shortcomings often associated to the traditional FL teaching methods and ii) *Proactive responses*, which are factors concerned with reinforcing multilingualism, as well as with the dominant role of English in Europe as a crucial aspect of the job market (Llinares, 2015; Pérez-Cañado, 2012).

The European Union (EU) is a multilingual area, made up of states with different official languages, as well as numerous regional and minority languages. Consequently, in order to maintain the EU’s aspiration to be ‘united in diversity’, language policies have an essential role. Different initiatives have been undertaken to support CLIL, as it is considered a tool to answer the European need to improve L2 education and multilingualism (Cenoz, Genesee and Gorter, 2014). In what follows, some of the main actions of the European institutions in relation to CLIL will be presented.

Since 1995, the European Commission has introduced a series of policies to improve the teaching and learning of FLs within the EU. The *Council Resolution of 31 March 1995 on Improving and Diversifying Language Learning and Teaching within the Education Systems of the European Union* and the *White Paper on Education and Training (Teaching and Learning – Towards the Learning Society)* (1995) presents the new legislation for multilingualism which was being introduced in Europe at the time. The general aim of this policy is to promote FL learning opportunities in the EU and to prepare multilingual European citizens who will have better opportunities in a globalized Europe (Llinares, 2015). This resolution provides the basis to continue the construction of a multilingual Europe. In order to achieve this, the European Council states the importance of “improving and diversifying the teaching and practice of such [European] languages, thereby enabling every citizen to have access to the cultural wealth rooted in the linguistic diversity of the Union” (p. 1). Thus, the Council emphasized the need to:

- “promote, by appropriate measures, qualitative improvement in knowledge of the languages of the European Union within the Union's education systems [...]
- [...] to encourage diversification in the languages taught in the Member States [...]

(Council Resolution of 31 March 1995, p.1)

Among the measures proposed to improve language learning in schools and universities, “innovative methods” such as “periods of intensive teaching and learning”, “the opportunity for teaching staff on mobility schemes” and “the teaching of classes in a foreign language for disciplines other than languages, providing bilingual teaching” (p.2) were recommended.

In the same year, and along the same lines, the *White Paper on Education and Training* (1995) presented guidelines for support measures and actions to be taken in order to solve identified needs in the fields of education and training. This document set five general objectives: i) encourage the acquisition of new knowledge, ii) bring schools and business closer together, iii) combat exclusion, iv) proficiency in three community languages, and v) treat capital investment and investment in training on an equal basis. The fourth objective, the one our study is concerned with, states that “The European Commission believes that it is necessary to make proficiency in at least two foreign languages at school a priority” (p.13). FLs proficiency is seen as prerequisite for EU citizens to benefit from the opportunities the ‘border-free Single Market’ offers (p.47). Multilingualism is considered an essential support to build up a European feeling of identity. The *White Paper* (1995) also states that language learning at a young age promotes the achievements of good results at school. According to this document, proficiency in FLs “[...] opens the mind, stimulates intellectual agility and, of course, expands people's cultural horizon” (p.47). In order to become proficient in two EU languages, in addition to the mother tongue, the beginning of FL learning at pre-school is also suggested in this document. In primary education, systematic language teaching is seen as crucial, followed by the introduction of a second community FL in secondary school. Moreover, they propose the study of certain content subjects in the first FL. As already stated, the main objective is that “Upon completing initial training everyone should be proficient in two Community foreign languages.” (p.47).

In 2002, the multilingual policy of the EU was further reinforced by means of the *Barcelona European Council Presidency Conclusions* (2002). In the section devoted to education, the objective of enabling European citizens to communicate in two EU languages other than their L1 is restated. In that section, the European Council calls for further action to promote multilingualism and suggests teaching at least two European FLs to young children (p.19).

Since multilingualism is considered to be at the core of the European identity, it comes as no surprise that after the publication of policy documents as the ones presented above, action has been taken to promote language learning and support European language diversity. In this scenario, the implementation of innovative FL teaching methods was given special attention, and CLIL became a valuable tool to achieve these goals. In the Eurydice (2006a) survey *on Content and Language Integrated Learning (CLIL) at School in Europe*, CLIL programmes are seen as a means by which:

“[...] pupils learn school subjects in the curriculum while at the same time exercising and improving their language skills. Subjects and languages are combined to offer them a better preparation for life in Europe, in which mobility is becoming increasingly more widespread and should be within reach of everyone.” (p.3)

In this document, special attention is given to the fact that CLIL content subjects are taught *with* and *through* an FL/SL, not *in* a FL/SL. As regards the status of languages, the language patterns in the EU are varied and several combinations are possible. The TL may be foreign, regional and/or minority languages or other official state languages (although the latter are less frequent). CLIL is defined as:

“[...] a generic term to describe all types of provision in which a second language (a foreign, regional or minority language and/or another official state language) is used to teach certain subjects in the curriculum other than language lessons themselves” (Eurydice, 2006a: 8).

According to this definition, CLIL may be understood as an approach equivalent to immersion or as a methodology that comprises immersion programmes (Lasagabaster and Sierra, 2010). In spite of this observation, some authors have claimed that CLIL belongs to “[...] contexts where the classroom provides the only site for learners’ interaction in

the target language. That is, CLIL is about either foreign languages or lingua francas” (Dalton-Puffer, 2011: 182). Furthermore, although CLIL is an approach meant to be applied to any language, the dominance of English (with the exception of English-speaking countries) as the actual language of instruction in CLIL methodologies is overwhelming (Dalton-Puffer, 2011; Dalton-Puffer and Smit, 2013, Dalton-Puffer et al., 2010; Eurydice, 2006a; Lasagabaster and Sierra, 2010).

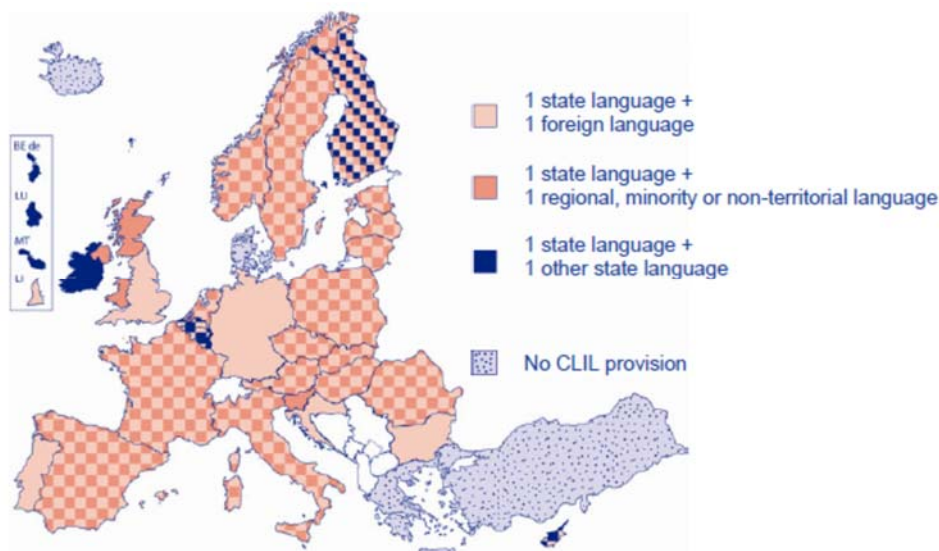
In most European countries, CLIL programmes are offered at primary and secondary education levels, and their duration corresponds to, at least, one period of compulsory education. Nevertheless, and just as with the languages involved, this varies depending on the country. Another characteristic is that, in general, all pupils have access to CLIL programmes, although some countries have established entry conditions that usually depend on an entry test to determine whether the pupils possess appropriate knowledge of the content and the language required for the CLIL class (Eurydice, 2006a).

The 2008 *Communication on Multilingualism: an asset for Europe and a shared commitment* highlights the main aims of the European Commission regarding multilingualism and outlines some more actions to be taken. In this communication, the priority of the EU is reaffirmed: “[...] to raise awareness of the value and opportunities of the EU's linguistic diversity and encourage the removal of barriers to intercultural dialogue” (European Commission Communication, 2008: 5).

More recently, the European Commission report on *Key Data on Teaching Languages at School in Europe* (Eurydice, 2012) devoted a section to CLIL entitled ‘Foreign language provision in the context of CLIL in primary and secondary education’. In this report, the EU states that CLIL is being implemented in most European countries (except for Denmark, Greece, Iceland and Turkey) and describes CLIL as “[...] a form of education provision according to which non-language subjects are taught either through two different languages, or through a single language which is 'foreign' according to the curriculum” (Eurydice, 2012: 39). Moreover, it is stated that when two languages are used in CLIL, the status of the TLs varies from country to country. Regional and minority languages are widely used as the medium of instruction in addition to the FL. This is usually the case in countries and regions with more than one official language, and/or with one or more regional/minority language (Eurydice, 2012). Figure 3 illustrates the status of the TLs used for CLIL programmes across Europe. As can be seen in the map,

in some countries different situations may take place, hence the presence of different statuses of the TL (e.g. Spain, France or Finland).

Figure 3 Status of TLs used for CLIL in primary and/or general secondary education, 2010/11 (taken from Eurydice, 2012: 41).



The absence of official admission criteria for CLIL programmes in most EU countries is restated in this document. However, the existence of some exceptions, i.e. Poland, Romania and Liechtenstein, is acknowledged. Furthermore, in countries with no official admission regulations, schools may implement their own criteria. As can be gathered from this report, as well as from the previous ones, the implementation of CLIL in Europe still presents many contextual differences (Ruiz de Zarobe, 2015).

The introduction of this innovative approach to language teaching has meant a major change in the EU and CLIL programmes have quickly extended across Europe. These programmes are becoming increasingly accepted in the continent, and CLIL is gradually becoming a recognized teaching approach available now to a great number of language learners (Heras and Lasagabaster, 2015).

Because of the important similarities of CLIL programmes and its predecessors in other parts of the world, and before offering a detailed definition of CLIL in the European context, here follows a brief overview of the similarities and differences between these teaching approaches.

4.2 CLIL predecessors: Immersion programmes and content-based instruction

Although CLIL is a relatively new approach in Europe, it has some antecedents, or inspiration sources, in North America such as the immersion programmes in Canada and Content-based Learning (CBL) or CBI in the United States (Euridyce, 2006a; Pinter, 2011). Some authors have argued that teaching a non-language subject *with/through* a TL is a practice distinctive of immersion and CBI only, whereas others have pointed out that it is also present in European bilingual programmes (Cenoz, 2013; Ruiz de Zarobe, 2015). For instance, in the Basque Autonomous Community in Spain a bilingual programme has been implemented for over 30 years in which several subjects are taught in Basque. Basque is not a FL in this region, but a SL to many of the locals. This type of programmes also exists in other bilingual regions in Spain, such as Galicia, Navarre and Catalonia.

As in CLIL, the main goal of immersion, CBI programmes and bilingual education is that learners acquire knowledge of both subject matter content and the TL, combining the teaching of these two aspects in the classroom. Other essential property shared by these teaching approaches is that the TL becomes the medium of instruction of the subject matter (Cammarata and Tedick, 2012; Loewen, 2015; Lyster, 2007).

Nevertheless, even though these three contexts “[...] are similar in their approach to classroom instruction” (Loewen, 2015: 149), they have specific characteristics that differentiate them from each other (Dalton-Puffer, 2011; Dalton-Puffer and Smit, 2013; Dalton-Puffer et al., 2010; Lasagabaster and Sierra, 2010). As will be discussed below, in immersion/bilingual and CBI programmes the teaching of language and content is ‘balanced’ (see Lyster (2007)), whilst some authors claim that in CLIL this is somehow different, and the ‘integration’ of these two aspects rather means the teaching of both at the same time (Llinares and Peña, 2014). There are also additional aspects that distinguish these teaching approaches, but before we move on to present those, an overview of the so-called predecessors of CLIL will be provided. However, it must be noted that other scholars argue that these programmes are, in fact, different labels for the same teaching approach (e.g. see Cenoz (2015)). This point will also be discussed.

4.2.1 CLIL, Immersion and CBI, same or different?

Immersion programmes originated in Quebec in the 1960s with an early total immersion programme in which English monolingual children received instruction entirely in French (until Grade 3, when English was introduced). These programmes successfully spread to other regions of Canada and to other parts of the world as an approach to improve learners' language proficiency level. The initial idea was that by mere immersion the students would just acquire the language. However, research has shown that more than that is necessary to become a competent user of the language. Thus, Lyster (2007) proposed a counter-balanced approach to this type of instruction, in which learners' attention should shift between content and form. Immersion programmes have been defined as

“[...] a form of bilingual education that aims for additive bilingualism by providing students with a sheltered classroom environment in which they receive at least half of their subject-matter instruction through the medium of a language that they are learning as a second, foreign, heritage, or indigenous language. In addition, they receive some instruction through the medium of a shared primary language, which normally has majority status in the community” (Lyster, 2007: 8).

The other approach considered as predecessor of CLIL are CBI programmes. CBI is a teaching approach common in the United States that may include different immersion programmes (Lyster, 2007). It started in the 1980s taking as a reference the success of the Canadian immersion programmes. Met (1998) defines CBI as “[...] an approach to second language instruction that involves the use of a second language to learn or practise content” (Met, 1998: 35). CBI is grounded in both the acknowledged benefits of FonF, i.e. language acquisition is promoted by brief attention to language structures during ‘larger, meaning-focused interaction’ (Long, 1996), as well as by models of ‘incidental and implicit learning’, namely that the L2 is acquired incidentally as learners focus on the content (Dalton-Puffer, 2007). Another important aspect, shared by the other approaches presented above, is that language focus should enable L2 learning, as well as the learning of the content (Loewen, 2015; Valeo, 2013).

Empirical evidence shows that immersion and CBI programmes have been beneficial for the development of the learners' knowledge of the L2 and for academic achievement,

providing “[...] not only the cognitive basis for language learning, however, but also the requisite motivational basis for purposeful communication” (Lyster, 2007: 2).

On the downside, L2 proficiency has not always been reported to reach the expected levels (Cammarata and Tedick, 2012; Loewen, 2015). Research has provided evidence that whilst learners reach native-like proficiency in receptive skills, productive skills remain at a lower level of command (Perez-Cañado, 2012). This has been argued to be due to “[...] a lack of systematic attention to language development during subject matter instruction” (Cammarata and Tedick, 2012: 253).

In order to achieve the maximal potential of immersion and CBI programmes, researchers have suggested the need for a more balanced teaching of language and content for systematically addressing both (Cammarata and Tedick, 2012; Lyster, 2007), “[...] rather than resorting to traditional decontextualized grammar instruction on the one hand, and content instruction with only incidental mention of language on the other” (Lyster, 2007: 138). This final aspect is one of the key features considered to differentiate CBI and immersion from CLIL. Before detailing the specific features of CLIL, this teaching approach will be presented against the backdrop of CBI and immersion programmes in order to compare these three methodologies.

From the definitions above, the aim of immersion, CBI and CLIL is to teach content and language at the same time. Their goal is that the learners become proficient in the L1 and L2, and acquire knowledge of the content subject. Another aspect that these programmes share is that their teaching approach is communicative, as another important goal of these methodologies is to achieve effective communication (Lasagabaster and Sierra, 2010).

Nevertheless, in spite of the acknowledged characteristics CLIL has in common with the other language learning approaches, many authors claim that CLIL programmes have some distinguishing features that immersion and CBI do not share (Dalton-Puffer et al., 2014; Lasagabaster and Sierra, 2010). These reported differences normally focus on the goals of these teaching methodologies, the learners’ and practitioners’ profiles, the TL used as the medium of instruction, and the relation between content and language instruction (balance vs. integration) (Cenoz et al., 2014). The differences between these methodologies, or lack of them, remain a controversial topic as some counter-arguments to the aspects traditionally used to differentiate them have been proposed, concluding that

the differences are unsupported and that these teaching approaches are not so clearly distinguished (Cenoz et al., 2014).

Starting with the most controversial feature, most researchers agree that the language of instruction of CLIL programmes is a FL⁹, not a SL (Dalton-Puffer and Smit, 2013; Lasagabaster and Sierra, 2010; Llinares and Peña, 2014), and consequently, learners have very limited access to the TL outside the classroom. Thus, CLIL learners receive a lower amount of hours of exposure and contact with the TL, as it is not as available as in immersion settings, where the language of instruction is often an official local language (Pérez-Cañado, 2012). Nevertheless, it must be noted that some authors have argued against this distinguishing feature, since according to the Eurydice report (2006), minority languages can also be used in CLIL (Cenoz et al., 2014).

Another important difference, connected to the previous aspect, is the fact that in immersion programmes the teachers are usually NS. In CLIL programmes, however, the teachers are NNS of the TL and normally content experts, rather than language experts (Lasagabaster and Sierra, 2010). The onset age of these programmes also presents a distinction: learners in CLIL programmes normally start at a later age than those in immersion and CBI (Dalton-Puffer, 2011).

Lasagabaster and Sierra (2010) stress that the teaching materials are another distinguishing factor. In immersion programmes, the materials are the same as the ones used by NS, whereas CLIL materials are often adapted to FL learners. Moreover, immersion programmes aim at a native-like command of the language. CLIL programmes however, do not aim at such a far-reaching objective. For instance, the Basque Department of Education proposed (for English) a B1 level of the Common European Framework of Reference for Languages (CEFR) at the end of secondary school (Lasagabaster and Sierra, 2010).

One more essential aspect is the way in which content and language are integrated. In contrast to the ‘balancing’ of the two aspects present in immersion and CBI programmes, which include some focus on form and on meaning, integration in CLIL means

⁹ Even though, as seen in the European language policies discussed above, the EU does not make a direct reference to FLs as the means of instruction in CLIL lessons, and rather offers a quite flexible range of language type, some authors have made of this aspect a unique feature of European CLIL (e.g. Dalton-Puffer, 2011).

“addressing both at the same time” (Llinares and Peña, 2014: 17), without preference for one over the other (Coyle, 2007). This difference however, has been somehow softened. Lyster (2007) modified the original idea of ‘balance’ in immersion and proposed a more counter-balanced approach in these programmes, which encouraged giving equal weight to focus on meaning and focus and form. This has been seen as “movement away from relying solely on the idea of the self-propelled, implicit language learner” (Dalton-Puffer, 2011: 194).

Finally, contextual and historical aspects have also been addressed as reasons to consider CLIL different (Dalton-Puffer, Llinares, Lorenzo and Nikula, 2014). CLIL is different from immersion and CBI in that it is considered European, as it originated as a European concept and from linguistic needs of the EU (Muñoz, 2007b). Moreover, the fact that the EU has supported and promoted the implementation of CLIL programmes has also contributed to its consideration as a European teaching approach (Dalton-Puffer et al., 2014).

In any case, regardless of the differences between these teaching approaches, they are all based on similar ideas and the influence of the Canadian immersion programmes and American CBI on CLIL must be acknowledged. In the literature, it is clear that there is an ongoing discussion about the relationship between CLIL and these other teaching approaches and, to our knowledge, no consensus has been reached in this regard. In some cases, the European origins of CLIL seem to outweigh the rest of the distinguishing aspects argued to be unique of this methodology. As Dalton-Puffer (2011) stated “[...] whether a concrete program is referred to as immersion or CLIL often depends as much on its cultural and political frame of reference as on the actual characteristics of the program” (p. 183). In what follows, the definition of CLIL will be addressed and its main distinctive characteristics will be enumerated.

4.3 CLIL: Definition and main characteristics

The term CLIL emerged in the mid 1990s and it has been defined as “[...] any educational situation in which an additional language and therefore not the most widely used language of the environment is used for the teaching and learning of subjects other than the language itself” (Wolff, 2007: 16). However, as discussed above, CLIL does not consist of just teaching a subject matter in a different language to that of the local population. Therefore, we believe Coyle, Hood and Marsh’s (2010) definition to be more precise:

“[...] a dual-focused educational approach in which an additional language is used for the learning and teaching of both content and language” (Coyle et al., 2010: 1). This definition, although it still presents CLIL as an umbrella term that may include a variety of bilingual education programmes which use an additional language as the medium of instruction for different school subjects (Cenoz et al., 2014; Llinares, 2015; Mehisto, Marsh and Frigols, 2008), highlights the double focus of CLIL on both language and content, with no preference of one over the other. CLIL has also been described as “a foreign language enrichment measure packaged into content teaching” (Dalton-Puffer, 2011: 185). More recently, Llinares (2015) underlines another distinguishing feature of CLIL: “[...] a clear notion of the integratedness of the teaching and learning of content and language.” (p. 69). In her definition, Llinares (2015) distances CLIL from previous characterizations (as a dual approach in which the teaching of content and language is balanced) which have led to controversy since they could be applied to other teaching approaches (Cenoz et al., 2014). In what follows, the most salient features of CLIL will be listed.

First, CLIL is considered to be European, as it emerged as an answer to European needs (Heras and Lasagabaster, 2015; Lorenzo, Casal and Moore, 2010). Moreover, many CLIL programmes have been supported and funded by EU policies and that may have also influenced its consideration as typically European (Dalton-Puffer et al., 2014). Another important feature of CLIL programmes is that they are implemented in courses in which the learners have already acquired literacy skills in their L1 (Dalton-Puffer and Smit, 2013). Also, CLIL programmes are implemented in mainstream education, and therefore, they are available to virtually all citizens (but see below).

The TL in CLIL programmes is usually a FL or a lingua franca, and therefore not used by the local community (Dalton-Puffer and Smit, 2013; Dalton-Puffer et al., 2014). This also implies that the teachers are normally NNS of the TL, and often experts on the content they teach, rather than on the language (Dalton-Puffer et al., 2010). Moreover, CLIL lessons are timetabled as content subject sessions while the TL continues to be a subject of its own, to be taught as an FL by language experts on top of the content lessons that are taught through the TL (Dalton-Puffer et al., 2014). This way, the amount of exposure to the TL in CLIL programmes is larger compared to the hours of FL teaching in MS settings (Dalton-Puffer and Smit, 2007).

The larger amount of hours exposure to the TL benefits language implicit learning in a similar way to immersion programmes (Heras and Lasagabaster, 2015; Muñoz, 2007b). Typically, CLIL programmes offer around 50% of the curriculum in the TL (Dalton-Puffer, 2011), with the added advantage of not taking up time from the other subjects in the curriculum (Lasagabaster and Sierra, 2010). Additionally, CLIL programmes have been claimed to be more efficient, since two scholastic components, content and language, are addressed in the time normally devoted to one. These aspects are very important, especially when working with YLs. As discussed in Chapter 3, the meaningful use of the TL, together with large amounts of quality exposure, are considered as more important than age of onset (see García Mayo and García Lecumberri (2003); Muñoz, (2006)), since they allow YLs to benefit from the ‘implicit learning mechanisms’, characteristic of children (Muñoz, 2006, 2007b; Nikolov, 2009; Nikolov and Mihaljevic-Djigunović, 2006, 2011). Therefore, CLIL becomes a valuable tool to improve the language learning outcomes in primary education.

CLIL is considered to be a more natural way of acquiring a language and, among other benefits, it has been claimed to be very valuable because “[...] it provides plenty of real and meaningful input to the learners and raises their overall proficiency in the target language” (Coyle, 2007: 548). Thus, the input provided in CLIL lessons is qualitatively different to the one in MS classrooms. The TL is used to convey information which renders it communicatively more purposeful than the language in traditional FL contexts, which is frequently manipulated for the sake of language teaching (Dalton-Puffer, 2011; Lázaro Ibarrola and García Mayo, 2012; Muñoz, 2007b; Ruiz de Zarobe, 2008). Therefore, as Ruiz de Zarobe (2008) proposes “[...] adopting a content-based approach within foreign language learning can thus guarantee more and richer opportunities for using the language in meaningful ways” (p. 63).

CLIL lessons promote interaction in the language classroom, not only between the teacher and the students, but also among learners (Nikula, 2007). In fact, it has been reported that learners see the TL in a CLIL lesson as a tool for communication, rather than as an object of study (Agustín Llach, 2009). Thanks to the double focus on meaning and content and to the real communicative situations that take place in the CLIL classroom, learners also acquire more language specific terminology (Heras and Lasagabaster, 2015). This has been claimed to be so because “[...] classroom content is not so much taken from

everyday life or the general content of the target language culture but rather from content subjects, from academic/scientific disciplines or from the professions” (Wolff, 2007, pp. 15-16). CLIL methodologies are also beneficial for the learners’ motivation and attitude to learn the TL, as language anxiety is believed to decrease because the focus is placed not only on the language forms but also on the content (Heras and Lasagabaster, 2015; Lasagabaster, 2011; Muñoz, 2007b).

As CLIL was originally aimed at improving learners’ FL proficiency, most research has focused on the effect of CLIL on linguistic outcomes (Ruiz de Zarobe, 2015). As for the advantages of CLIL methodologies over MS approaches, the literature has documented numerous benefits of this teaching approach regarding L2 acquisition (Ruiz de Zarobe, 2011). As regards the effect of CLIL on content learning, most research has shown that CLIL also exerts a positive impact. To our knowledge, research to date has not reported statistical differences in content learning between learners in CLIL and those receiving the content lessons in the L1 (Admiraal, Westhoff and de Bot, 2006; Jäppinen, 2005; Ruiz de Zarobe and Lasagabaster, 2010; Seikkula-Leino, 2007). Still, these findings need to be taken with caution since in different contexts the results have also shown that learners may experience some difficulties (Airey and Linder, 2006; Hellekjær, 2010).

The aspects presented above portray the essence of CLIL programmes, and provide the key to their consideration as a more efficient way to teach a FL than MS methods (Dalton-Puffer, 2011). Table 12 illustrates the main benefits of CLIL methodology.

Table 12 *Benefits of CLIL methodology.*

Naturalistic setting for language acquisition:

It makes the learning of a language more meaningful.

Content and language integration provides a more ‘authentic’ communicative act.

Greater amount of exposure to the TL:

Language addressed in the time normally devoted to a subject matter.

Focus on meaning reduces learners’ language anxiety and increases their motivation.

Positive context for content learning.

The purported benefits of CLIL are based on psycholinguistic theories: CLIL has taken some elements of Krashen’s Monitor Model (1985), such as the importance of naturalistic settings and the value of comprehensible and meaningful input, as well as the role of the

affective filter, for language learning. CLIL is beneficial for the latter because in CLIL lessons language mistakes are supposedly not penalised, and consequently, language anxiety is believed to decrease (Dalton-Puffer, 2011). CLIL has also been influenced by Long's Interaction Hypothesis (1996): meaningful interaction and meaning negotiation are seen as sources of comprehensible input, essential for language learning. The CLIL classroom provides a context in which interaction is promoted. Swain's Output Hypothesis (1993) also plays an important role, since language production is believed to stimulate attention to language forms, not only to meaning. Finally, from a socio-cultural perspective, CLIL offers a context in which language is learnt "through the participation in social events" (Dalton-Puffer and Smit, 2007: 10).

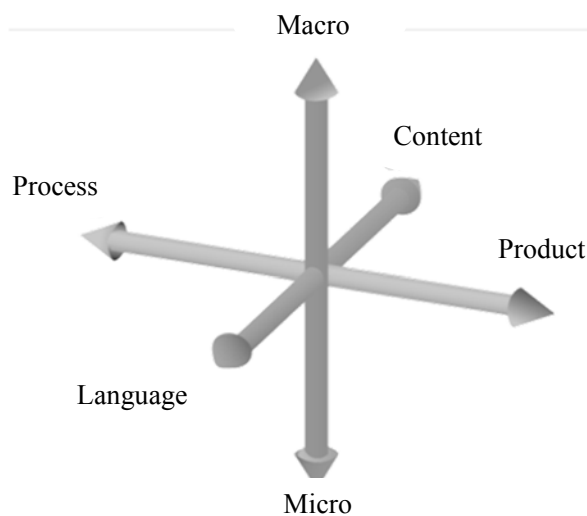
As shown in this section, in spite of sharing some common features with other teaching approaches, CLIL has its own distinguishing characteristics which make it a unique teaching approach, different from the programmes which are believed to have inspired and influenced it. In what follows, a review of the research on CLIL will be provided, as well as a revision of the effect on learning outcomes of this teaching approach.

4.4 CLIL research review: CLIL and learning outcomes

Dalton-Puffer and Smit (2007) distinguished two dimensions from which CLIL has been approached. These authors categorised studies taking into consideration whether they focused on either macro- or micro-level dimensions of CLIL and whether they were process- or product-oriented. Micro-level approaches focus on the immediate participants (i.e. teachers and learners) and the CLIL classroom. On the other hand, the macro-level perspective, "[...] is concerned with taking an outside view of the conditions under which CLIL happens and on courses of action which can be taken to implement CLIL" (Dalton-Puffer and Smit, 2007: 13).

Product-oriented studies are concerned with the outcomes of CLIL whilst process-oriented research addresses the course of the methodology being implemented. Later on, a third dimension was added to categorise studies according to their focus on either language or content (Dalton-Puffer et al., 2010). The new three-dimensional categorisation is illustrated in Figure 4 below.

Figure 4 *Three-dimensional CLIL research space (taken from Dalton-Puffer et al. (2010: 10)).*



Llinares (2015) combined two dimensions of Dalton-Puffer and Smit's (2007) categorisation and offered a further conceptualisation of four types of research studies on CLIL. Product-oriented macro studies are defined as those reporting on “already implemented CLIL programmes” (e.g. Coyle et al., 2010), whilst process-oriented macro studies normally address the process of implementation of CLIL at different levels and contexts (e.g. Eurydice, 2006a; b; Lorenzo et al., 2010), as well as reports concerned with learning arrangements and task types deemed suitable for CLIL programmes (Dalton-Puffer and Smit, 2007). On the other hand, product-oriented micro studies have predominantly compared language attainment of CLIL students with non-CLIL learners (grammar/written production (e.g. Ackerl, 2007; Hüttner and Rieder-Bünemann, 2007; Jexenflicker and Dalton-Puffer, 2010), vocabulary (e.g. Sylvén and Ohlander, 2014; Xanthou, 2011) and the effect of CLIL on learners' motivation (Lasagabaster, 2011). Some researchers have focused on content learning, but the literature is comparatively scarce (e.g. Admiraal et al., 2006; Airey and Linder, 2006; Hellekjær, 2010; Jäppinen, 2005). Finally, Llinares (2015) describes process-oriented micro studies as those addressing the “CLIL classroom as interactional space” (Llinares, 2015: 60) (e.g. Dalton-Puffer, 2007; Nikula, 2007). We will start our review of CLIL research in Europe with studies addressing the effect of CLIL on content learning.

In Finland, Jäppinen (2005) conducted a long-term study (2001–2003) in which the influence of CLIL on the thinking and content-learning processes in mathematics and

science was analysed. The participants were 7 to 15 years old learners. The CLIL learners (N= 335) were taught through English, French, or Swedish, whereas the control group (N= 334) were taught in their L1 (Finnish). In most of the age groups analysed, the cognitional development in CLIL equalled the one achieved in MS classes. No statistical differences were found between CLIL and non-CLIL learners in the youngest groups (7-9 years old) in mathematical thinking and learning processes while in science, these groups appeared to have difficulties with very abstract topics. Nevertheless, as these learners grew up and their thinking processes developed, they showed higher cognitional levels than the control group (p. 158). The most encouraging results were obtained in the second age group (10-14 year-olds), as CLIL significantly benefited mathematics and science learning and thinking processes. Finally, no statistically significant differences were found among the older learners, which was attributed to the smaller amount of CLIL hours received by this age group. In sum, the data presented in this study contribute to indicate that CLIL environments, although initially more demanding, provide favourable conditions for thinking and content learning.

Seikkula-Leino (2007) also carried out her study in Finland and, contrary to the results obtained by Jäppinen (2005) in the 10-14 year-olds group, reported no statistically significant differences between CLIL learners and non-CLIL learners (grades 5 and 6) in mathematics. On the other hand, concurring with previous research, a higher motivation in the CLIL groups was identified. Finally, this author found that in non-CLIL groups “[...] pupils were strongly overachievers, meaning that pupils overachieved in both subjects – Finnish language and mathematics” (p. 335). In the CLIL classes however, learners were overachievers in their L1 or in mathematics, but not in both subjects.

In the Netherlands, Admiraal et al., (2006) conducted a longitudinal study on secondary education learners’ proficiency in English and Dutch and their achievement in content subject when taught through the TL (English). CLIL learners were more proficient in the TL in terms of oral and reading skills. Receptive skills, however, did not seem to be affected. With regard to their content knowledge, no negative results were found.

In spite of the encouraging results reported above, these findings were only partially supported by research carried out in Scandinavian contexts (Airey and Linder, 2006; Hellekjær, 2010). These studies compared reading comprehension in English and in the L1 (Swedish or Norwegian) and, although no significant differences among the two

teaching methods at tertiary education levels were found, some difficulties with understanding concepts and taking notes in the TL were reported (Airey and Linder, 2006; Hellekjær, 2010).

Airey and Linder (2006) state that there are indeed differences between learning through one's L1 and a FL. In their study, although learners felt that the language of instruction did not influence their learning, under stimulated recall they identified a number of difficulties when learning through a FL (for instance, devoting a great effort to taking notes). A decrease in classroom interaction when the medium of instruction was the TL was also reported. Learners seemed to be reluctant to speak during the class, although they did ask questions at the end. Nevertheless, these results differ from the ones obtained by Nikula (2007) in Finland, who investigated language use in the CLIL classroom, and reported more teacher-student interaction in upper-secondary CLIL programmes than in MS methodologies. This might be explained by not only the different levels and demands, but also by the way students are selected to be part of the CLIL groups in these two contexts (Bruton, 2011a; Hellekjær, 2010).

Hellekjær (2010) also reported that, although no significant differences in reading comprehension scores were found between MS and CLIL learners, students found learning through a TL more laborious. Difficulties in following the line of thought of the teacher and with taking notes were identified. These findings were attributed to the possibility that less positive results in CLIL lessons may have been influenced by the mechanics of taking notes in the L2 (Hellekjær, 2010).

In sum, although some contradictory results have been obtained, what seems to be unquestioned by the research to date is that learners taught through a FL/SL are able to attain, at least, similar levels of content knowledge as when taught in their L1 (Ruiz de Zarobe, 2015).

The findings concerning language skills are more distinct and research has shown numerous advantages of CLIL. Benefits not only for reading, but also speaking (Admiraal et al., 2006) and writing (Hüttner and Rieder-Bünemann, 2007) have been reported. Moreover, CLIL classrooms provide a communicative setting to use the TL which boosts learners' motivation and self-confidence (Ruiz de Zarobe, 2015). Following Ruiz de

Zarobe (2015), we have classified the studies according to the linguistic aspect they address.

Regarding written production, overall benefits have been documented, although some linguistic areas have been found to be more affected by CLIL instruction than others (Ackerl, 2007; Hüttner and Rieder-Bünemann, 2007). Hüttner and Rieder-Bünemann (2007) compared the micro- and macro-level production of 44 EFL learners in Vienna (age 12). The overall results show that the CLIL learners had an advantage in the command of the micro-level features of narratives (tense consistency and verb forms) and, although to a lesser extent, of the macro-level features (plot development and story resolution). The authors suggest that some macro-level features may be affected by general cognitive skills whilst micro-level skills seem to be more influenced by the specific characteristics of CLIL instruction.

Jexenflicker and Dalton-Puffer (2010) also analysed the written production of Austrian 11th grade learners (age 16). The results revealed that, although the two groups (mostly) covered the content requirements, the essays by the CLIL learners presented overall advantages over those written by their MS counterparts. The differences were identified mainly in vocabulary, expression and grammatical accuracy, as well as text format and register. Although problems were found in textual organization and discourse competence in the two groups, these were more noticeable in the non-CLIL group. Concurring with Hüttner and Rieder-Bünemann (2007), Jexenflicker and Dalton-Puffer (2010) state that “CLIL instruction affects those areas most which concern purely linguistic skills (i.e. grammar and vocabulary)” (p. 182).

In another study also carried out in Austria, Ackerl (2007) analysed the written production of ten 12th grade learners (age 18), five following a CLIL course and five a non-CLIL methodology. This study provides further support for the positive impact of CLIL on written production. The findings suggest that CLIL learners, although also produced a large number of errors, use a wider range of expressions, vocabulary and verb tenses, as well as more complex sentences, and make fewer errors than non-CLIL learners. Reflecting earlier work, the results here exhibit how some aspects are more positively influenced by CLIL (lexical richness) than others (verb tenses).

As regards the influence of CLIL on vocabulary learning, studies are still scarce. Sylvén and Ohlander (2014) conducted a large-scale longitudinal study in Sweden and compared

upper-secondary education learners (age 15-16) in CLIL programmes to age-matched non-CLIL students. Their study focuses on written academic skills, specifically on the effect of CLIL on vocabulary acquisition and motivation. The results obtained are in line with those reported in previous studies, that is, that CLIL learners clearly outperformed their non-CLIL counterparts. Following Bruton's (2011a) appeal for more critical awareness, the authors acknowledge the fact that CLIL programmes are not mandatory, and that, therefore, it is possible that only the most motivated learners, and those with a higher level of competence in the TL, attend CLIL classes. A closer look into the results shows that male learners outperformed the female. This was attributed to the assumption that only the most proficient males enrolled in CLIL, while females saw it also as an opportunity to improve their TL proficiency.

Further evidence supporting the benefits of CLIL on vocabulary acquisition is provided by Xanthou (2011). As with previous findings, significant differences were identified in the vocabulary acquisition of 11 years old CLIL pupils (L1= Greek). Content subject (i.e. science) learning was also analysed in this study and, concurring with previous research (Ruiz de Zarobe and Lasagabaster, 2010; among others), no significant differences were found between CLIL and non-CLIL learners.

Although still more research is needed and some conflicting aspects have been identified, numerous studies support the benefits of CLIL for language learning (Coyle, 2007; Ruiz de Zarobe, 2015; among many others). It has been demonstrated that different language aspects are influenced in different ways, some developing faster than others, although overall outcomes are encouraging. Moreover, there is ample evidence that CLIL instruction does not negatively affect the learning of subject matter content. Finally, research suggests that CLIL learners tend to be more motivated (Lasagabaster, 2010; Pladevall-Ballester, 2015), which positively affects their language learning process. Nevertheless, other authors claim that instead of increasing motivation, CLIL may produce a loss of self-esteem in learners who are not confident enough to use a language they do not fully control (Cenoz et al., 2014).

As seen in this brief review, numerous differences that make the comparison of findings difficult are present in the studies, such as learners' characteristics (e.g. age, L1), types of CLIL programmes (compulsory vs. voluntary), and measures used to examine different variables. It has also been argued that the rapid spread of CLIL across the EU may be not

be thanks to its actual positive effect on language acquisition but to “claims of the success of CLIL without substantial empirical evidence” (Cenoz et al., 2014: 14), and that further research that controls for some variables that may not have been accounted for before is needed (Bruton, 2011a,b). These and other not so encouraging aspects of CLIL will be discussed in the last section of the current chapter. Before that, the implementation of CLIL in Spain and the principal research projects carried out in this country, together with their main findings, will be reviewed.

4.5 CLIL in Spain

CLIL has rapidly spread all over Europe in the last ten years (Llinares, 2015). In Spain CLIL programmes have followed the same trend (Ruiz de Zarobe and Lasagabaster, 2010) and are being implemented to varying degrees (Heras and Lasagabaster, 2015). Some authors have claimed that Spain is one of the countries at the head of Europe in CLIL practice and research (Coyle, 2010; Heras and Lasagabaster, 2015).

The concept CLIL has been translated into Spanish as *Aprendizaje Integrado de Contenidos y Lenguas Extranjeras* (AICLE). According to the Eurydice report (2006b), the curriculum in CLIL programmes in Spain is to be covered in at least two languages (as in the rest of EU countries), within two different approaches, depending on whether they are in a monolingual or in a bilingual community:

- “it may be taught in the official language of the State (Spanish) but also in a joint official language other than Spanish (Basque, Catalanian, Valencian and Galician);
- it may be taught in the official language of the State (Spanish) but also partly in one or two foreign languages, such as English in accordance with the Ministry of Education and Science/British Council Agreement” (Eurydice, 2006b: 3).

With regard to the CLIL subjects and the number of hours of CLIL per week in pre-primary and primary education, the EU recommends teaching at least 40% of the curriculum in primary education, and 30% in pre-primary. It is also specified that CLIL hours should be taken from other subjects, always provided that these content subjects keep the teaching hours allocated to them in the core curriculum. The provision of Spanish Language and Literature must not be compromised by the teaching of subjects in English

in the first stage of primary education. The minimum amount of hours in English per week is specified in the *Documento Técnico de Mínimos* (Table 13).

Table 13 *Minimum weekly timetable of CLIL lessons (adapted from Eurydice, 2006b: 6).*

Stage	Age	Subjects	Weekly timetable
Pre-primary education	3-6 years old	Education comprising all subject areas.	7-9 hours.
First stage of primary education	6-7 years old	Part of the contents of the natural, social and cultural environment; artistic education (plastic arts); English; mathematics (just basic concepts); sometimes music (within artistic education) and physical education (psychomotor skills).	9-10 hours.
Second stage of primary education	8-9 years old	The same subject areas as in the first stage.	10-11 hours.
Third stage of primary education	10-11 years old	The same subject areas as in the other two stages.	11-12 hours.

In Spain there are no admission requirements to enter CLIL programmes in the early stages of education, nor do pupils have to pay supplementary fees either. In pre-primary and primary education, pupils are not required to sit an entry test or to have a certain level of competence in the TL. In contrast, students in secondary education have to pass a test to assess their level of proficiency in the TL. As regards the language used as medium of instruction, in pre-primary and primary education the curriculum is developed in both the L1 and the TL, whereas in secondary education the subjects taught through the TL are never taught in Spanish. The methodology for all the levels of these programmes is “communicative, participatory, active and motivating” (Eurydice, 2006b: 6). Finally, learners’ TL attainment and content knowledge are evaluated following continuous assessment.

Even though the Organic Law of Education 2/2006 (Ley Orgánica de Educación LOE 2006) establishes the core regulations concerning education in Spain, each autonomous

community regulates its implementation within their territory. Consequently, CLIL programmes may vary from one region to the other (Llinares, 2015; Ruiz de Zarobe and Lasagabaster, 2010). However, this aspect is not unique in Spain and it can be generalised to other countries (Ruiz de Zarobe and Lasagabaster, 2010).

The autonomous communities which have carried out most research on CLIL have been Andalusia, the Basque Country, Catalonia, Galicia and Madrid. As already discussed, the implementation of CLIL may vary from one autonomous community to the other, since the educational policies present differences in each region. Still, all the CLIL projects share the common goal of improving the FL proficiency of the learners, as well as the core features defined above. Figure 5 below portrays a simplified version of the Spanish linguistic map.

Figure 5 *Simplified version of the Spanish linguistic map (Taken from Fernández Fontecha, 2009: 5).*



In **Andalusia**, the regional Andalusian government launched in 2005 the *Plan de Fomento del Plurilingüismo* (the Plurilingualism Promotion Plan). This policy was implemented over the years 2005-2009 and aimed at improving the linguistic competence in this region, this way adhering to the multilingual policies of the EU. The plan proposes 72 actions, and has developed a network of over 400 bilingual primary and secondary schools. Among the different measures implemented, half the schools in the region became bilingual, teaching up to 40% of the curriculum in English. Another action

proposed was the introduction of two FLs as means of instruction in the whole educational network (Lorenzo et al., 2010).

Within this framework, Lorenzo et al. (2010) conducted one of the first large scale multidimensional studies on European CLIL. This study compared the language attainment of CLIL learners (TL English N= 754; TL French N= 423; TL German N= 143) in primary and secondary education with that of learners of the same age in MS courses (N= 448). It must be noted that the control groups were only L2 English. Three different TLs were examined, focusing on four main aspects: “[...] linguistic outcomes and competence levels; acquisitional routes and individual differences; L2 use in CLIL classrooms; and educational effects beyond the L2” (Lorenzo et al., 2010: 436). The results clearly show that CLIL learners outperform their counterparts in non-CLIL classes. The positive outcomes relate not only to the increased exposure to the TL and incidental learning but also to the meaningful language use and meaning-focused methodology, which also contributed to reduce learners’ language anxiety. Another benefit of CLIL found in this study is a greater cohesion between language and content teaching, promoted by more cooperative work between teachers in the schools.

Interestingly, the level of proficiency of the learners of the three FLs was comparable. This finding is surprising mainly because the EFL learners had only had a year and a half of instruction through this language, whilst the German and French had experienced bilingual education since they started primary school. Thus, Lorenzo et al. (2010) provide further evidence that show that late-starters may make up for an early start thanks to their more developed meta-cognitive abilities, and that the amount and quality of exposure may have a more decisive role than age (Muñoz, 2007b).

Concerning the amount of L2 use in the classroom, their results suggest that teaching assistants use the L2 the most, similarly to instructors in full-immersion programmes, whereas language teachers would represent partial-immersion and content teachers seem more inclined to code-switch.

These findings, however, have been discussed by Bruton (2011a), who argues that the absence of a pre-test that showed the learners’ initial level of proficiency in the TL challenges the validity of the reported linguistic benefits of CLIL. This aspect is related to another point: since no pre-tests were conducted, the level of the control group may have already been lower than that of the CLIL learners. Bruton (2011a) claims that this

is quite likely, arguing that the learners who choose CLIL may be more motivated (Ruiz de Zarobe and Lasagabaster, 2010; San Isidro, 2010), have a higher level of command of the TL, or belong to families of 'higher socio-economic-status' (Bruton, 2011a: 237). Therefore, although all learners in a school have the possibility to attend CLIL programmes, the ones with the stated characteristics are the most likely to choose this option, giving rise to what Bruton (2011a) calls 'disguised selection' (p. 240). Another argument for the likelihood of different levels is that CLIL programmes and teachers receive more support than MS approaches. CLIL coordinators have more time to devote to the design of the courses and teaching materials. The presence of teaching assistants is a variable that plays an important role. Finally, the data collection instrument used for measuring L2 use in the class has some shortcomings. Questionnaires may not be sufficiently reliable: in fact, studies on L1 use in the classroom have reported greater use than Lorenzo et al. (2010) (e.g. Carless, 2004). Also, since no observable data were collected, the teachers' actual L2 use, as well as their command of the TL cannot be reported. In sum, Bruton (2011a) claims the need for research that takes the above stated limitations into consideration in order to assess the true effects of CLIL.

The **Basque Country**, unlike Andalusia, is a bilingual region in which Basque and Spanish are official languages and, therefore, both are taught throughout compulsory education. In this region, three models of education are available: Model A, in which Spanish is the language of instruction and Basque is taught as a subject matter three or four hours per week; Model B, in which subjects are taught in Basque and in Spanish, 50% of the teaching time allocated to each language; and Model D, in which the subjects are taught in Basque, and Spanish is taught during four hours per week. In 2003 the *Plurilingual Experience* was launched, which states that a FL must be the medium of instruction of at least 7 hours per week in Compulsory Secondary Education, and from 20% to 25% of the subjects in Post-Compulsory Education (Ruiz de Zarobe and Lasagabaster, 2010).

Ruiz de Zarobe (2008) conducted a longitudinal study in this autonomous community in which she analysed the differences in the oral production of CLIL and MS learners. The participants were 89 secondary school learners, who were divided into three groups according to the school programmes they followed. All the participants were bilingual (Spanish and Basque) and English was their third language (L3). The non-CLIL group

(N=29) received three hours of English per week. The CLIL1 group (N= 24) started the CLIL programme when they were 14, and received 3-4 hours of Social Science through the TL (English), on top of the regular EFL classes. The CLIL2 group shared the same characteristics as CLIL1, but received two curricular subjects through the TL (Social Sciences: 3-4 hours a week, and Modern English Literature: 2 hours a week). Five aspects of speech production were analysed: pronunciation, vocabulary, grammar, fluency and content. The overall results showed that in all the categories the two CLIL groups performed significantly better. What is more, there was a direct relationship between the amount of CLIL instruction and the results obtained: Group CLIL2 performed significantly better than the other two groups in all the aspects analysed. When the different school years were compared, the positive results were confirmed: the CLIL groups outperformed the non-CLIL, although the differences between the CLIL programmes were not so remarkable. In the 4th Year, CLIL2 performed significantly better in all categories, except for the results in vocabulary where no significant differences were found. In the last year before university, only the non-CLIL and the CLIL2 groups were compared and, interestingly, statistical differences were found only in two categories: vocabulary and grammar. The author concludes that the results suggest that the more exposure to the TL, the greater the proficiency achieved by the learners (but see criticism by Bruton (2011b)).

Lázaro Ibarrola and García Mayo (2012) also explored the effects of CLIL on 15 EFL Spanish-Basque high-school students' oral production over a two-year period. Their results show that the use of the learners' L1 in discourse markers and repair sequences decreased significantly, although this decrease did not correlate with an increase in their use of English discourse markers or paraphrasing of words they did not know. Their study was also concerned with morphosyntax, and the results in this regard were more optimistic and showed a significant development, suggesting that the CLIL learners were 'one step ahead' when compared to non-CLIL learners.

One of the research groups within the Language and Speech Laboratory (LASLAB) group (www.laslab.org) from the University of the Basque Country has worked with schools in the Basque Country and Navarre during the years 2013-2016 and has investigated the oral performance of YLs of EFL, both in CLIL and non-CLIL programmes, when carrying out different collaborative tasks. This research project, led by Dr. García Mayo, is entitled *Oral interaction among young learners of English as a foreign language: The impact of*

the use of negotiation and corrective feedback strategies during communicative tasks on language learning and was funded by the Spanish Ministry of Education (FFI2012-32212). The current dissertation is in fact part of this larger research project which addresses the need to further investigate YLs' FL acquisition, particularly from an interactionist perspective. Among other aspects, the researchers in this group have examined the effect of CLIL on the learners' TL (English) acquisition. The project started in 2013 and, as reviewed throughout the literature presented in the previous chapters, several studies have been conducted to date (e.g. García Mayo and Imaz Agirre, 2016; García Mayo et al., in press; Lázaro Ibarrola and Hidalgo, 2017a; Martínez Adrián and Gutiérrez Mangado, 2015).

Regarding content learning, several longitudinal studies have been conducted in this region and the results have shown that CLIL learners achieve a similar level of content knowledge as the learners in MS courses, who receive the content lessons in their L1 (see for instance Ruiz de Zarobe and Lasagabaster (2010)). Some authors have suggested that the positive results obtained may have been influenced by motivational factors, as it has been reported that CLIL boosts learners and teachers' motivation (Lasagabaster, 2011; Lorenzo et al. 2010; among others). As discussed above, these findings have been partially supported by studies in other parts of Europe: while CLIL provides a positive context for content learning, the results are still not significantly better than those obtained by non-CLIL learners (Admiraal et al., 2006; Jäppinen, 2005).

In **Catalonia**, the CLIL-SLA Project is carrying out numerous studies dealing with learners in grades 5 and 6 of primary education. The project consists of a two-year longitudinal study which investigates the effects of CLIL on the learners' competence in the TL (English) and attitude towards it. A special focus is placed on controlling some of the methodological limitations found in previous CLIL research, specifically the amount of TL the learners in the different programmes are exposed to and the initial proficiency level in English of CLIL and non-CLIL groups (Bret Blasco, 2014).

Pladevall-Ballester (2015) focused on the perceptions of CLIL stakeholders, i.e. learners, teachers and parents. The learners (N= 197) were in their 5th year of primary education (age 9-10) and attended five different schools in Catalonia. Five CLIL teachers and 159 parents also participated in the study. By means of interviews and questionnaires, the participants' opinions after a year of CLIL implementation in the schools were analysed.

The results reveal that learners, teachers and parents consider CLIL to be motivating and to have a positive effect on SLA. Learners in general (except for the low achievers) are satisfied with the CLIL teaching approach, above all because they deem the learning and use of the TL more meaningful. Moreover, most learners consider the CLIL lessons ‘quite easy’, except for the oral production in the TL which was reported as problematic by over half of the participants. Parents seem to consider CLIL a better way to improve learners’ proficiency in the TL (i.e. English), although some believe that the exposure to the TL is not enough to attain proficiency at the same time they show concerns about their children’s L1 development and content knowledge acquisition. Finally, teachers believe that CLIL “[...] was a positive experience and a valuable tool which increased the learners’ motivation and oral comprehension” (Pladevall-Ballester, 2015: 50). They also considered that learners in general had no problems following the lessons, although were worried about the low-achievers who had more difficulties. All teachers were aware of the students’ improvement of their language skills, above all their oral comprehension, motivation and willingness to use the TL. Still, the practitioners showed some concern about the lack of materials, support and CLIL training, as well as about their ability to teach the content.

Like the previous two autonomous communities, **Galicia** also has experience in bilingual immersion programmes involving the local language (Galician). Spanish and Galician are the two official languages in this region, Galician being the vehicular language of, at least, 50% in all the educational levels up to university. Like in the rest of Spain, there is no entry test for the CLIL programmes, in accordance with the equal opportunities policy established by the European Community (San Isidro, 2011).

San Isidro (2010) analysed the linguistic competence in the TL (English) of CLIL and non-CLIL learners. The participants (n= 278; 154 CLIL, 133 non-CLIL) attended 10 schools in Galicia and were in the 4th grade of compulsory secondary education. The results revealed a significantly higher overall command of the TL as well as of all the individual language skills (i.e. reading, writing, listening and speaking) of the CLIL learners. The author, however, acknowledges that these findings should be taken with caution as participation in the CLIL programme was optional and, therefore, the learners’ motivation as well as their command of the TL may have been different to that of their non-CLIL counterparts. Thus, San Isidro (2010) calls for further research that includes

longitudinal studies and the examination of not only the linguistic development of the learners, but also of the effect of the CLIL approach on content learning.

In contrast to the three regions presented above (The Basque Country, Catalonia and Galicia), **Madrid** is a non-bilingual autonomous community and, consequently, teaching content through a TL is a relatively new approach (Llinares and Dafouz, 2010). In the 2004-2005 academic year, the CAM Bilingual Project was implemented in a number of primary schools in the region of Madrid. The first year 26 state schools took part in the project and by now it incorporates 336 state primary schools and 98 state secondary schools, as well as 163 state-subsidised schools. The schools that participated in this project teach at least a third of the curriculum in the TL (English). In this region, numerous research groups have begun to study the implementation of CLIL in the educational system at various levels (pre- and primary, secondary and tertiary). For instance, the UAM-CLIL Project, which started in 2005, is concerned with identifying the students' linguistic needs in specific subjects and providing support to secondary school CLIL teachers. This research group "set out to create a corpus of samples of CLIL students' language, both spoken and written, which would show how these pupils deal with the difficulties involved in expressing the content of a discipline in English" (Llinares and Whittaker, 2010: 126).

Llinares and Whittaker (2010) compared the oral and written production of CLIL and MS secondary school students (age 13-14) of history. They presented a detailed systemic functional linguistic (Halliday, 1994) analysis of the learners' performance when carrying out the same task in their respective language of instruction, i.e. TL (English) and in the learners' L1 (Spanish). The results show that both groups expressed content with similar types of verbs and process types, and that there was little difference between the two modes in the two languages. For instance, a limited use of clause connectors was found in the oral and written performance of both groups. Moreover, a differentiating characteristic maintained by the participants is the more frequent use of relational processes to make comparisons in written than in spoken production. Despite these similarities, the findings also showed that CLIL learners used more relational processes than the non-CLIL groups. Modality was scarcely expressed in general (mostly used to express obligation), although the CLIL groups made a more varied use (including also probability and usuality). The CLIL groups focused more on time and place both in

written and oral production, whilst the MS learners used cause and manner. Non-CLIL learners also used a wider variety and a higher proportion of circumstances, “[...] showing more developed awareness of the register of the discipline.” (Llinares and Whittaker, 2010: 134). The authors conclude that in order to take advantage of the full potential of CLIL, learners need to learn the language of the academic discourse of the different disciplines. Therefore, training becomes essential to make teachers aware of the linguistic features needed for a specific content subject, including different genres and both in the written and spoken modes.

As seen in this literature review, studies in Spain have reported that CLIL is a positive teaching approach that brings multiple benefits to the learners. CLIL stakeholders already seem to have this impression, as reported by Pladevall-Ballester (2015). Their intuition is supported by studies which show that CLIL learners outperform non-CLIL in their TL general proficiency (Lázaro Ibarrola and García Mayo, 2012; Ruiz de Zarobe, 2008) and at least obtain similar results in their content learning (Ruiz de Zarobe and Lasagabaster, 2010). Moreover, CLIL learners are more motivated (Lasagabaster, 2011), which may also exert an influence on their results. However, some limitations have been identified, which are also present in CLIL programmes in other parts of Europe, such as the deficiencies in teacher training, the fact that CLIL programmes are optional in many schools (Bruton, 2011b) and that comparisons between CLIL and non-CLIL learners’ language proficiency do not always take into account the amount of hours of exposure to the TL these learners receive (Bret Blasco, 2014; Martínez Adrián and Gutiérrez Mangado, 2015). Moreover, as Llinares and Whittaker (2010) point out, more attention to the functions of language in the CLIL classroom is needed. Another controversial aspect is that, in studies that control for TL exposure, non-CLIL learners are older than CLIL learners and, therefore, they may be at a different maturational stage. As Martínez Adrián and Gutiérrez Mangado (2015) claim, age at testing may be more important than type of exposure. These and other potential flaws in research on CLIL will be further discussed in the last part of the current chapter. Still, in spite of these shortcomings and in light of the results obtained, CLIL in Spain seems to be on the right track to become a very valuable tool which will definitely improve Spanish learners’ FL proficiency. In the next section, our focus will narrow to the research carried out in Navarre, the region where our study has been carried out.

4.5.1 CLIL in Navarre

In the context of Navarre, immersion programmes have a long tradition. Like in the Basque Country, in Navarre there are immersion-like programmes in some schools where all the subjects are taught in Basque (the second official language of the region). According to the Eurydice report (2006b) the Autonomous Community of Navarre has four bilingual educational establishments, two corresponding to pre-primary and primary education and two to secondary education. In 1997, the pre-primary programme was launched and at the beginning of the academic year 2002/2003, the primary level programme started. A year later, in 2003/2004, the bilingual programme started in secondary education.

Multilingual programmes start in the second stage of pre-primary education (3-6 years old) and continue through Compulsory Secondary Education (12-16 years old). The learners' proficiency in the TL (English) is tested before admission in the centre, although it is not used for selection purposes, but rather to keep a record of their academic development. Regarding the content subjects that may be taught in English, except for Spanish and Basque, any area is eligible. In the Eurydice Report (2006b), subjects such as technology, music, social sciences and sciences are listed as the content subjects usually taught through the TL in Secondary Education, although these are selected by each school and might include other subjects. Even though the methodology might change from school to school, this report recommends teaching the subjects in the TL eleven to twelve hours in pre-primary and primary school. In fact, "[...] in the first year or two, one half of the content may be offered in the CLIL target language and the other half in Spanish" (Eurydice, 2006b: 8). According to the *Curriculum for Secondary Education (English)*, secondary students must have 5 sessions of EFL per week, whereas Science and Geography and History will be allocated the same hours as in the current Spanish education system.

As mentioned above, a subgroup within the LASLAB group from the University of the Basque Country, supported by the Spanish Ministry of Education and Competitiveness, has worked with two primary schools in Navarre examining the oral performance of CLIL and non-CLIL YLs of EFL. For instance, the already reviewed papers by Azkarai and Imaz Agirre (2016), García Mayo and Lázaro Ibarrola (2015) and Lázaro Ibarrola and

Hidalgo (2017a,b)¹⁰. These papers focus on how CLIL YLs interact with each other and negotiate for meaning when performing a collaborative task, an aspect which, to our knowledge, had not been investigated before in a CLIL setting. Results show how YLs in CLIL programmes are able to negotiate for meaning and use a variety of strategies to do so. In García Mayo and Lázaro Ibarrola (2015), CLIL learners were reported to use more interaction strategies than their non-CLIL school counterparts. However, Azkarai and Imaz Agirre (2016) found that some negotiation strategies were initiated more frequently by non-CLIL learners as a function of the task examined.

As presented above, the influence of CLIL on YL's L1 use has also been analysed. In Navarre, García Mayo and Lázaro Ibarrola (2015), García Mayo and Imaz Agirre (2017) and García Mayo and Hidalgo (2017) reported a lower reliance of the CLIL learners on the L1 than their MS age-matched peers.

4.6 Conclusion: Research gaps in CLIL and future directions

In the current chapter, we have offered a detailed description of CLIL and analysed its main characteristics, framing it against its predecessors (i.e. Canadian immersion programmes and CBI). A revision of relevant research studies and main findings within this approach has also been provided. In this last section, we will examine some of the critical voices that have recently argued against a number of the above presented aspects of CLIL.

A growing body of research has reported overall beneficial effects of CLIL over MS approaches. Nevertheless, as can be ascertained from the description of CLIL features and the literature review, there is room left for discussion over what can be considered CLIL (Dalton-Puffer and Smit, 2013). In fact, it has been claimed that “[...] what characterizes CLIL more than anything is the remarkable variety of practices that can be found under its umbrella. (For a particularly expansive understanding see Mehisto, et al., 2008: 13; see also Coyle et al., 2010)” (Dalton-Puffer et al., 2010: 2). Thus, some authors ask for a clarification of the term and the methodology to be followed (Cenoz et al., 2014).

The distinctiveness of CLIL is in itself a controversial aspect. Some researchers stress CLIL's uniqueness and, therefore, the need to study it on its own (Dalton-Puffer et al., 2014; Lasagabaster and Sierra, 2010; Perez-Cañado, 2012). Others, however, maintain

¹⁰ See Chapters 1 and 3.

that no distinction can be made between CBI and CLIL (Cenoz, 2015; Cenoz et al., 2014). For instance, Cenoz et al. (2014) argue that the fact that CLIL features are not clearly outlined makes it difficult to distinguish CLIL from other teaching approaches that use an L2 as the medium of instruction. Following Met's (1998) classification of immersion as a type of CBI, Cenoz (2015) states that "[...] there are no differences between CBI and CLIL regarding their essential properties" (p. 21), and proposes sharing the research findings of these teaching approaches.

Nonetheless, given the adaptability of CLIL to different contexts, differences in its implementation may be found (Coyle, 2007). Thus, the flexible nature of this approach has been argued to influence the results when they are transferred from one context to another uncritically (Cenoz et al., 2014). This may lead to interpretations and practices that might be far from the initial idea of language and content integration. Therefore, for some authors, a clear account of the benefits of CLIL in specific contexts is essential. Other authors, however, declare that this is an aspect that might be also present in other educational research areas (Dalton-Puffer et al., 2014; Cammarata and Tedick, 2012).

Furthermore, we cannot forget that CLIL programmes are optional in many schools, which might mean that the most motivated learners (Lasagabaster and Sierra, 2010; Ruiz de Zarobe and Lasagabaster, 2010; San Isidro, 2010), and probably those with higher content subject scores and TL proficiency levels, will be the ones attending CLIL classes (Bruton, 2011a, b; Ruiz de Zarobe, 2015). This is another factor which, when not taken into consideration, might influence the results of CLIL studies (Bruton, 2011a, b). Thus, some authors claim that, in order to develop the true potential CLIL has, an in-depth analysis is needed to account for the commonalities of this approach in diverse contexts. The training of CLIL teachers is an additional predicament that has to be overcome. CLIL instructors must take on the responsibility of teaching through a FL, in which they need to be competent. Moreover, they must know how to exploit content materials for language teaching (Ruiz de Zarobe, 2008).

Regarding research methodology, numerous studies have compared CLIL to MS lessons without considering the fact that CLIL offers learners a greater number of hours of exposure to the TL (Ruiz de Zarobe, 2015). Consequently, it has been argued that more detailed research is needed to examine to what extent the positive results obtained in CLIL studies are influenced by its unique features rather than by CLIL itself and the integration

of language and content. Some authors point out that most studies do not control for these external variables, which may exert an influence on the results, and claim that findings, therefore, need to be interpreted with caution (see Bruton (2011a,b); Pérez-Cañado (2012)).

One more aspect that has been criticized is that most research on CLIL has focused on learners' linguistic outcomes and compared the language proficiency of CLIL students and non-CLIL learners whilst comparatively few studies have focused on content outcomes (Cenoz et al., 2014; Lasagabaster and Ruiz de Zarobe, 2010; Llinares, 2015). Many authors have claimed that more input from content specialists is necessary in order to shed more light on the effects of CLIL on the acquisition of content knowledge (Cenoz et al., 2014; Dalton-Puffer and Smit, 2007; Dalton-Puffer et al., 2010; Llinares, 2015; Llinares and Peña, 2014).

Finally, there is limited research on how language and content can be best integrated in CLIL (Cenoz et al., 2014). Llinares (2015) offers a review of recent research for understanding language and content in integration. Some authors argue that the double focus on content and language intended in CLIL programmes is very difficult to achieve in the classroom (see Dalton-Puffer (2007)).

Although these and other factors need to be taken into consideration when analysing CLIL, research to date has provided evidence for many advantages of CLIL over MS methodologies. Numerous beneficial aspects for the learning of both language and content have been reported. In summary, CLIL has the potential of offering (provided there is a specific and successful teacher training) a naturalistic setting for language acquisition that makes language learning more meaningful. Moreover, the integration of content and language provides learners with more 'authentic' communicative acts. Meaningful communication is essential for language learning, and it is believed to also improve the learners' communicative competence by allowing them to take part in real communication. Another valuable feature of CLIL is that it is considered to be more efficient, as it addresses two aspects in the time normally devoted to one. This way, the amount of exposure to the TL in CLIL programmes is larger compared to the hours of language teaching in MS contexts (Dalton-Puffer and Smit, 2007). Finally, CLIL programmes' focus on meaning has positive effects on the affective level, reducing the

learners' language anxiety and increasing their motivation (Heras and Lasagabaster, 2015; Lasagabaster, 2011).

It has been reported that CLIL learners outperform learners in MS classes in their command of the FL. The reported benefits are in terms of general proficiency (Coyle, 2007; Dalton-Puffer, 2007; Lorenzo et al., 2010) and also in some specific areas of the language, such as pronunciation and vocabulary (Dalton-Puffer, 2007; Lázaro Ibarrola and García Mayo, 2012; Ruiz de Zarobe, 2008). Additionally, some authors have underscored that CLIL methodologies promote specific linguistic abilities or behaviours typically associated with effective language learning. These include risk-taking and problem solving skills, linguistic confidence, student independence and linguistic spontaneity (Coyle, 2007; Dalton-Puffer, 2007; among others). Consequently, CLIL is being implemented in European schools and it is considered a valuable tool to promote bilingual education, with no detriment of content learning. Still, further research is needed in order to develop this methodology to its full potential.

PART II
THE PRESENT STUDY

CHAPTER 5 THE STUDY

5.1 Motivation of our study

As can be inferred from the literature review outlined in the previous part of the current dissertation, there are areas related to child FL learning that remain under-researched.

Thanks to teaching approaches such as TBLT and CLIL, there is an increasing amount of peer interaction taking place in second and foreign language lessons. Even so, the opportunities learners normally have to use the TL are limited, partly because of the ratio of learners to teacher in a classroom and sometimes also because of the teacher-centred approach used for teaching. As Van den Branden (1997) states, in these groups “the less assertive and less proficient learners receive minimal output opportunities” (p. 598). Oral collaborative tasks are considered a valuable tool to push learners to use the TL meaningfully and generate large amounts of output (García Mayo, 2007; Pica, 2013; Van den Branden et al., 2009). Nevertheless, some aspects of the effect of peer-peer interaction on FL learning remain unclear. In order to help to illustrate how YLs interact in an EFL setting, in the present dissertation we have chosen an oral collaborative task in which learners had to interact without any help from the teacher.

The question of potential differences related to the learners’ age is another aspect of interest. As discussed in Chapter 3, childhood is a period of many changes and, consequently, tasks that may be appropriate for a specific age group might not be so beneficial for another (Muñoz, 2007b; Nicholas and Lightbown, 2008; Pinter, 2006, 2011). As far as we know, only García Mayo and Lázaro Ibarrola (2015), García Mayo and Imaz Agirre (2016) and Azkarai and Imaz Agirre (2016) have addressed age as an independent variable that affects the use of interactional strategies by young EFL learners when performing an oral collaborative task. Whereas García Mayo and Lázaro Ibarrola (2015) and Azkarai and Imaz Agirre (2016) found that the group of younger learners used more CAs than the group of older YLs, García Mayo and Imaz Agirre (2016) reported no statistically significant differences between their participants’ use of negotiation strategies. These results obtained in FL contexts however, do not concur with those reported by Oliver (2009) in an ESL setting since, as already seen, the younger learners in her study made limited use of some negotiation strategies.

Moreover, we would also like to examine the effect of TR on YLs' general performance. When working with collaborative tasks, learners have to focus both on form and meaning, which becomes especially hard due to the spontaneous nature of oral communication. TR has become a valuable way of diverting learners' attention from meaning to form. Repetition is believed to lead to improvements in aspects of the FL production, such as fluency, accuracy, complexity, as well as less NoM, and generally a more efficient organisation of the learners' language resources (Bygate and Samuda, 2005; Gass, et al., 1999; Mackey et al., 2007; Pinter, 2007). Furthermore, TR is common practice in language lessons because practitioners do not need to design and explain a new task every time. So far, the relation between this variable and YLs' use of negotiation strategies remains unclear, as very few publications are available in the literature that discuss this task condition (however see Mackey et al., 2007; García Mayo and Imaz Agirre, 2016; Lázaro Ibarrola and Hidalgo, 2017b). Similarly, it is unclear how YLs' general competence measured in CAF is affected by TR. Research on these aspects has so far provided rather inconclusive results.

Finally, since our study is set in an FL context, and in line with previous studies on interaction in FL settings, the learners' L1 will play an important role, as students will likely resort to it at some point when engaged in communicative tasks (Storch and Aldosari, 2010; Tognini and Oliver, 2012). Research has reported a limited use of the L1 in the FL classroom and, when used, it serves functions that facilitate task completion (Azkarai, 2013; Azkarai and García Mayo, 2015; García Mayo and Hidalgo, 2017; García Mayo and Lázaro Ibarrola, 2015; Lázaro Ibarrola and Hidalgo, 2017a; Storch and Wigglesworth, 2003; Swain and Lapkin, 2000). This dissertation considers not only the relationship between proficiency (related to age) and L1 use, but also whether and how the use of the learners' L1 is affected by the repetition of a task. The functions the L1 serves in learner-learner interaction will also be addressed.

In the present study, we hope to shed some light on those research gaps described above. Within the interactionist framework, we will analyse the oral production of two age groups (ages 8-9 and 10-11) of Spanish EFL learners at a basic level of proficiency when performing a collaborative task. Specifically, the aim of this study is to investigate:

- i) the nature of the NoM as well as the different negotiation strategies used by young EFL learners.

- ii) the differences regarding the two groups (their age and school year) when performing a two-way collaborative task, that is, whether younger and older YLs carry out the task similarly. This will be operationalised by measuring the NoM strategies they use, CAF and L1 use of each age group.
- iii) the effect of TR on the participants' output. We will focus on how NoM strategies, learners' CAF and the use of the L1 vary upon TR.

In the current chapter, the research questions and hypotheses of this dissertation are introduced. The participants and the setting of our investigation will also be described, as well as the materials used to conduct the experiment and the procedure to carry out the study. Finally, we will present the guidelines followed to codify and analyse the data.

5.2 Research questions and hypotheses

In this study, we analyse the oral production of two age groups of learners (ages 8-9 and 10-11) while performing an oral collaborative task in pairs in a Spanish school following a CLIL methodology. All the participants were learning EFL in the same school and the pairs were matched for age and proficiency level. The task was a PPT and it was repeated three times following the same procedure but changing the content (procedural repetition). The objective is, on the one hand, to describe the nature of students' interactions by examining their NoM strategies and, on the other hand, to observe if procedural TR has any effects on these negotiations (amount or type) and on the learners' general performance (CAF and L1 use). In order to achieve our research aims the following research questions guided the current study:

Negotiation of meaning

RQ1 To what extent do young EFL learners of different age groups negotiate for meaning when performing an oral collaborative task?

RQ2 What NoM strategies do YLs use in their oral interactions with age- and level-matched peers? What is the purpose of their negotiations?

RQ3 Are these strategies and their functions the same in the two age groups?

Task repetition

RQ4 Does procedural TR have an effect on YLs’

- use of NoM strategies?
- general performance (CAF)?
- L1 use and the functions it serves?

Age

RQ5 Does age have an effect on YLs’

- use of NoM strategies?
- general competence (CAF)?
- L1 use and the functions it serves?

On the basis of the literature reviewed above, we expect to find NoM among the learners, along with differences in the overall task performance and the NoM strategies younger and older YLs use. Moreover, we believe that procedural TR will exert an influence on the output of the two age groups from task to task. Hence, the following hypotheses will be tested:

Hypothesis regarding young EFL learners’ ability to negotiate when performing an oral collaborative task

Peer interaction is believed to trigger large amounts of NoM and, consequently, of modified output both in young and adult learners (García Mayo and Lázaro Ibarrola, 2015; Oliver, 1998, 2009). Additionally, jigsaw tasks, like the one used in our study, are considered the type that promotes interaction the most.

1. Taking this evidence into account, we expect that the pairs of EFL YLs participating in our study will negotiate for meaning in order to understand each other and complete the task.

Hypotheses regarding YLs’ use of NoM strategies

The literature has documented how young L2 learners use a variety of NoM strategies when performing collaborative tasks. However, strategies commonly used in adult negotiation, such as comprehension checks, are rarely used (Gagné and Parks, 2013; Oliver, 1998, et passim). Additionally, differences in the NoM by YLs in FL and SL contexts have been reported.

2. Thus, in the present study we expect our EFL young participants to use a variety of NoM strategies, similar to those reported in FL studies addressing YLs' oral interaction.
3. Accordingly, we believe that repetitions will be the most frequent strategy whereas CAs in general, and comprehension checks in particular, will be scarcely used.

Hypothesis regarding possible age differences in the participants' NoM

To our knowledge, only three studies have considered age-related differences in the NoM strategies YLs use: Oliver (2009) in a SL context; and García Mayo and Lázaro Ibarrola (2015) and Azkarai and Imaz Agirre (2016) in a FL setting. Whereas the findings reported by the FL studies concur in the older YLs using fewer NoM strategies than the younger learners, Oliver (2009) did not report significant differences between the two age group learners. As for the potential differences in the use of strategies and the functions they serve, Oliver (2009) reported a greater use of strategies concerned with 'other' by the older YLs, whilst the strategies concerned with 'self' were similar in the production of both younger and older learners.

4. Thus, we expect to document a lower use of NoM strategies by FL YLs than those reported for SL YLs, and a range of NoM strategies similar to the one observed in previous studies. Additionally, we expect the younger participants (age 8-9) to make a more extensive use of these strategies during their interactions than their older counterparts (age 10-11). Regarding the functions NoM strategies may serve in the output of each group, we believe that older YLs' will probably show a greater concern about their interlocutors' needs.

Hypotheses regarding the effects of procedural TR on the participants' oral performance

As indicated above, the impact of procedural TR on NoM has been rarely investigated. According to Mackey et al. (2007), procedurally familiar tasks generate more opportunities to use feedback and more actual use of feedback, although this task condition was not the one that triggered the most strategies. García Mayo and Imaz Agirre (2016), however, reported no significant effects of different types of TR on NoM strategies. On the other hand, Lázaro Ibarrola and Hidalgo (2017b) concluded that a decrease in the use of NoM strategies was found after the learners' third task performance.

García Mayo and Imaz Agirre (2016) repeated the task only twice, which may be the reason for the different results obtained.

5. Thus, either no significant changes or a decrease in the NoM strategies used by the participants in the current study is expected.

Task familiarity, achieved through TR, exerts a positive influence on language learning and on learners' general performance (Ahmadian and Tavakoli, 2010; García Mayo et al., in press; Kim, 2013) as it leads to a more efficient organisation of the language resources (Kim and Tracey-Ventura, 2013; Mackey et al., 2007; Pinter, 2007; Samuda and Bygate, 2008). However, inconclusive findings have been reported as regards the different aspects of CAF (e.g. Kim and Tracey-Ventura (2013), Patanasorn (2010)). Still, even though the literature dealing with the effect of TR on YLs' general performance is scarce, most of the studies have identified beneficial effects on fluency (Bret Blasco, 2014; García Mayo et al., in press; Pinter, 2007; Sample and Michel, 2014)

6. Therefore, we expect our learners to improve in terms of overall general competence, and anticipate that the aspect that will be affected the most by procedural TR is fluency.

We have also examined how procedural TR influences YLs' L1 use. As already seen, only Azkarai and García Mayo (2016) examined this aspect of YLs' oral performance, and a significant drop of the overall L1 use was reported. The most frequent functions at the two testing times are borrowings and appeals for help. The frequency of two functions (expressions of lack of knowledge and phatics) changed upon procedural TR: whereas phatics decreased significantly at T2, expressions of lack of knowledge increased significantly.

7. Thus, fewer L1 terms in the last task performance are expected.
8. We anticipate the vocabulary function to remain stable across task performances, and to be the most common L1 function. Discourse markers, which are equivalent to phatics in Azkarai and García Mayo (2016), are expected to decrease.

Hypotheses regarding possible age differences in the participants' oral performance

When the patterns of NoM in the oral performance of young FL learners of different ages have been compared, older YLs have been found to use fewer NoM strategies (García Mayo and Imaz Agirre, 2016; García Mayo and Lázaro Ibarrola, 2015). To our knowledge, only García Mayo and Imaz Agirre (2016) considered age-related differences in tandem with TR in the use of NoM strategies by YLs (albeit only two repetitions and with a 2-month interval between each task) and reported no significant impact of TR on the NoM by the two age groups.

9. We expect the younger participants to make a more extensive use of the NoM strategies than the older YLs at the three data collection points. We do not anticipate significantly different behaviours in the NoM of the two age groups upon TR.

As for the impact of TR on general competence, we believe that the only empirical study that has addressed the differences in general competence, operationalised as CAF, between different age YLs is García Mayo et al. (in press). Differences between two age groups (8-9 and 9-10 years old) were observed mainly in the increased fluency of the younger learners at the last TR and more accurate production of the older YLs.

10. Since the older YLs had received more hours of instruction in the TL, we anticipate oral production of this group to be more target-like in terms of CAF.
11. We expect the older learners' production to be more target-like and the younger participants to be more fluent by the last performance.

As regards L1 use, most research has demonstrated that learners make a limited L1 use and that even when the L1 is used, it benefits task completion (Azkarai and García Mayo, 2016; García Mayo and Hidalgo, 2017; García Mayo and Lázaro Ibarrola, 2015; Lázaro Ibarrola and Azpilicueta Martínez, 2015; Lázaro Ibarrola and Hidalgo, 2017a; Storch and Aldosari, 2010). Most studies have established a relationship between proficiency and L1 use. Nevertheless, some recent studies have pointed to other variables, such as motivation, which might also exert some influence (García Mayo and Lázaro Ibarrola, 2015). Still, no clear conclusion has been reached yet concerning this aspect. When L1 functions have been analysed, three main uses have been identified, namely vocabulary, metacognitive talk and discourse markers. Vocabulary and task management (metacognitive talk) have

been claimed to be the most common L1 functions (García Mayo and Hidalgo, 2017; Lázaro Ibarrola and Hidalgo, 2017a; Storch and Aldosari, 2010).

12. We anticipate a limited use of the L1 by our participants. In consonance with the generally acknowledged fact that a greater proficiency level leads to less L1 use, we expect that the younger learners (age 8-9) will use their L1 more frequently than their older counterparts (age 10-11). As for the functions the L1 serves, we expect vocabulary and metacognitive talk to be the most frequent L1 uses in the performance of the two groups.

5.3 Participants and setting

In order to address the objectives as stated, the oral production of two groups of young EFL learners while performing a two-way collaborative task was analysed. We have worked with learners of two different ages, specifically 8-9 and 10-11 years old. This age range is considered to belong to the middle-childhood stage which, according to Berk (2006), comprises children between 7 and 11 years old. Within the middle-childhood stage, children gradually incorporate adult-like features (Nicholas and Lightbown, 2008), are more capable of logical thinking and able to categorise and organise objects (Philp et al., 2008). We have chosen this age range because it represents a period in which children are able to take turns, talk about a topic for longer than younger children, and are aware of the pragmatics of speech acts (Philp et al., 2008). On the other hand, although children of these ages share this developmental stage, Pinter (2007) suggests that children aged 10-11 are more able to provide support to their interlocutor when engaged in conversation. Therefore, we would like to analyse whether younger children (age 8-9) are also able to fulfil their role as interlocutor or not.

Eighty (N= 80) young EFL learners participated in this study: the younger learners group consisted of 40 pupils in their 3rd year of primary education (age 8-9) and the older group was made up of 40 learners in their 5th year of primary school (10-11 years old). Out of the 80 participants in the study, 41 were girls and 39 were boys. The students attended a state primary school in Navarre that follows a CLIL programme, which is mandatory for all learners. This eliminates the potential limitation of what Bruton (2011a: 240) calls a 'disguised selection', that is, that only the most motivated learners and those with higher-than-average proficiency participated. According to school internal tests and external assessment (Programme for International Student Assessment (PISA)), the learners'

proficiency level in the TL can be equalled to an A1 level of the CEFR, although the group in Year 5 had a slightly better command of the TL. Although there was a variety of L1s among the pupils in this school, Spanish is the language of the community they live in, as well as that of the school, and all the participants speak Spanish fluently. Table 14 shows the distribution of the participants in this study.

Table 14 *Participants' profile.*

	3rd Primary Education	5th Primary Education
Participants	40	40
Age	8-9	10-11
Gender	19 males/21 females	20 males/20 females
Years studying English	5	8
English proficiency	A1	A1 ⁽⁺⁾

In this school, all subjects are taught half in the learners' L1 (Spanish) and half in the TL (English). Therefore, our participants receive around 50% of the teaching time in English (which corresponds to approximately 13 hours per week). Specifically, these pupils receive two sessions of math, natural science and physical education in English, and one of art per week. English language is studied in literacy class and is allocated six sessions. Thus, the approximate total amount of exposure to the TL these learners receive is 396 hours per school year. The teaching method is a typical communicative classroom where all skills are integrated by making an extensive use of games, songs and all sorts of communicative tasks. Teaching materials designed for native English children, such as magazines and newspapers, story books, folk songs and rhymes, as well as radio interviews, are also used in the lessons. Thanks to the communicative approach of CLIL, pupils are accustomed to and feel comfortable interacting in the TL with the teacher and, to a lesser extent, with their classmates.

Each grade in the bilingual programme followed in this school is divided into two groups of students, A and B, which receive the same subject in each respective language (i.e. Spanish and English). The Spanish and the English teacher work with the two groups, each of them through one of the languages. They work together and devote time each week to lesson planning and coordinating the activities to be carried out in class. The

school holds that tasks are never duplicated, but that the lessons in the two languages complement each other.

5.4 Materials and procedure

5.4.1 The task in our study: The picture placement task

In order to maintain ecological validity, the PPT used in the present dissertation was designed in collaboration with the teachers from the two classes that participated in this study. In conjunction to this teacher-researcher collaboration criterion, the task was designed based on others used in previous studies: a jigsaw type task was selected because, as stated in the *Task Classifications* section in Chapter 2, this task type is considered one of the most appropriate to foster interaction.

Jigsaw tasks, and therefore our PPT, have been presented as one of the most productive task types for L2 acquisition (Butler and Zeng, 2014; Ellis, 2003; Oliver, 2009; Pica, 2013; Pica et al., 1993; Pica et al., 2006). They have a closed outcome, that is, a single possible solution, and require interaction among the participants in order to achieve a common goal (Pica et al., 1993). The PPT we have used is a two-way task in which both learners in the dyad hold essential information to fulfil the task, interaction becoming essential for task completion. Thanks to the opportunities for interaction this type of task offers, conditions for learners working in pairs to negotiate for meaning are created. Since mutual understanding is needed, speakers sometimes need to clarify or explain the meaning they intend to convey (thus producing modified output and comprehensible input) and offering feedback to their interlocutors in response to their output. This way, these tasks provide learners with plenty of opportunities to attend to meaning, function and form (Pica et al., 2006).

Therefore, this task type is considered one of the most efficient at promoting language learning. While interacting, learners need to carry out a linguistic exchange in which they need to make a meaningful use of the TL and will very likely have to modify their production and engage in NoM (Varonis and Gass, 1985). We believe this to be an essential feature since there is abundant evidence suggesting that NoM facilitates language acquisition (Ellis, 2003, 2005; Gass and Mackey, 2007; Loewen, 2005; Long, 1981, 1983; McDonough 2005; McDonough and Mackey, 2006, 2008).

Our task was mainly designed to shed light on YLs' interactional patterns and, to a lesser extent, it intended to establish a context for description at a basic level. With regard to the characteristic real-world target of pedagogic task, the posters used portray real-life scenes, showing places well-known to the participants in this study (i.e. a classroom, a playground or a living room), thus allowing the learners to experience and use authentic and meaningful language.

As stated at the beginning of this section, we have tried to ensure the ecological validity of our study. In order to achieve this goal we have worked with teachers and have designed the PPT with the aim of implementing it within real classroom conditions. In fact, before the researcher started to collect the data, the teachers used a version of the task during their actual class time. However, in the present study, as in many other research studies (Pica, 2005, Sample and Michel, 2014), the tasks were carried out in a controlled environment with the researcher and the pair of learners outside the classroom. This has been done to guarantee a uniform implementation of the task across the different participants, even though we are aware of the concerns about the validity of this type of data. Table 15 summarizes the characteristics of the PPT.

Table 15 *Picture placement task characteristics.*

Procedure	Without seeing each other, participants interact to place the photos on the posters in order to make them identical
Type	Jigsaw
Information flow	2-way
Exchange of information	Required
Outcome	Closed

To perform this task, the participants worked in age- and level-matched pairs. Two identical posters (one for each learner) and two identical sets of 6 photos of children (these pictures will be referred to as *a, b, c, d, e, f*) were designed. Student A in the pair had her poster with pictures *a* and *b* already placed on it and pictures *c, d, e* and *f* outside the poster, while student B had the other poster with pictures *c* and *d* on it and pictures *a, b, e* and *f* outside the poster (See Appendix C: Posters used in the study). Each pair of learners sat at two tables separated by a folding screen so that they could not see each other and were forced to rely exclusively on oral English. The goal of the task was that

the participants, without seeing each other or each other's materials, interacted in English in order to complete their posters so that both ended up having the same pictures in the same places on their respective posters. In order to do so, the participants had to use the TL to ask questions as well as give information to their partners.

5.4.2 Procedure

Before collecting the data, the researcher obtained written permission for the learners' participation in the study from the school and the children's parents (see Appendix A). The actual data collection procedure took a total of three school sessions (one per week). Since, to our knowledge, there is not an established interval between repetitions (nor the number of repetitions has been set), we have followed previous research with YLs in EFL settings, specifically Mackey and Oliver (2002), Sample and Michel (2014) and Pinter (2007), and had the students perform the different tasks three times with a one-week interval.

The same task type was repeated three times on a weekly basis by every dyad. From now on, we will be referring to each TR session as Time 1, Time 2 and Time 3 (T1, T2, T3). Among the different types of TR, we have worked with procedural repetition. Thus, although the posters and pictures changed between data collection points, the task procedure remained the same. It must be noted that the differences in content were minimal, as the three tasks contained pictures of children performing simple actions (e.g. reading, eating, playing ball) and illustrated places with which the learners were familiar (e.g. school, rooms in a house, parks) (Posters and pictures appear in Appendix C).

The tasks were carried out in a separate room where the participants were video-recorded by the researcher. After the participants were given the instructions, the learners carried out the task working autonomously. The researcher only intervened to move the task along if the participants got stuck, to check if they had finished and to answer some questions about vocabulary.

There was no time limit so the participants could use as much time as they needed to perform the tasks. Table 16 shows the mean time the participants devoted to complete the task and the time range of the three task performances. The time learners spent at the beginning of each task discussing who would start is included. Year 3 learners needed an average of 6 minutes 13 seconds to complete the PPT. The session the participants

devoted more time to was T1 (7 minutes 26 seconds in average), whereas T3 was the session in which the learners invested less time (5 minutes 23 seconds). On the other hand, the mean time for Year 5 was 4 minutes and 49 seconds. Just like for Year 3 learners, the session in which the learners invested more time was T1 (5 minutes 51 seconds) and T3 was the one in which the learners needed less time (3 minutes 47 seconds). The mean times the learners in both groups needed to complete the tasks decrease upon TR. Learners in Year 3 needed more time than their older fellow learners to complete each task. Statistical analyses revealed that, in fact, the influence of the group on the time devoted to complete the tasks was significant ($F(1,78) = 556.83, p < .001$).

Table 16 Time invested in each task by the participants in each group.

	Task 1		Task 2		Task 3	
	Y3	Y5	Y3	Y5	Y3	Y5
Mean	0:07:26	0:05:51	0:06:03	0:04:48	0:05:23	0:03:47
Range	0:12:37 - 0:02:50	0:12:20 - 0:01:55	0:13:00 - 0:03:00	0:11:55 - 0:01:35	0:11:16 - 0:02:18	0:07:02 - 0:01:54

After running the data through Wilcoxon signed-rank tests, it became evident that the repetition of the PPT led to a statistically significant decrease in the time invested in each task in the two groups (Year 3: T1 – T3: $Z = -4.02, p < .001$; Year 5: T1 – T3: $Z = -4.23, p < .001$) (see Table 17).

Table 17 Time differences across tasks.

	<i>T1-T2</i>	<i>T2-T3</i>	<i>T1-T3</i>
Year 3	$Z = -2.50, p = 0.01$	$Z = -2.20, p = 0.41$	$Z = -4.02, p < 0.001$
Year 5	$Z = -3.25, p = 0.01$	$Z = -2.2, p = 0.02$	$Z = -4.23, p < 0.001$

As a whole, the Year 3 group needed 2 hours, 28 minutes and 46 seconds to complete T1, whereas the Year 5 learners invested 1 hour, 57 minutes and 9 seconds. T2 was completed by Year 3 in 2 hours, 1 minute and 3 seconds and T3 in 1 hour, 47 minutes and 30 seconds. The Year 5 group, on the other hand, invested 1 hour, 33 minutes and 57 seconds to carry

out T2 and 1 hour, 15 minutes and 39 seconds for T3. Table 18 illustrates the overall times the two age groups devoted to complete the three tasks.

Table 18 Overall time invested in each task by the two groups.

	Task 1		Task 2		Task 3	
	Y3	Y5	Y3	Y5	Y3	Y5
Time	02:28:46	01:57:09	02:01:03	01:33:57	01:47:30	01:15:39

5.5 Codification and data analysis

In the present section, the aspects of our participants' performance selected for analysis will be described, together with the measures used to examine their production and the codification tools. Additionally, the statistical tests used to analyse our data will also be presented.

The learners' output was transcribed verbatim by the researcher. A total of 11 hours, 6 minutes and 4 seconds were transcribed into the CHAT (Codes for the Human Analysis of Transcripts) format (MacWhinney, 2000). To determine the role of procedural TR in the oral performance of our participants as well as the differences between age groups, the transcripts were coded for the constructs of interest for our study, namely NoM strategies, L1 use and CAF. The transcripts were analysed using the CLAN (Computerised Language Analysis) tools (MacWhinney, 2000). This format is used in the Child Language Data Exchange System (CHILDES) Project, a widely used system for analysing child oral production. The total number of transcripts examined in this research study is 120, 60 from each group. An independent rater coded 30 of the participants' interactions, which constituted 25% of the whole dataset. Inter-rater reliability was calculated using simple percent agreement, which resulted in 95%, and any remaining discrepancies were solved individually on a case-by-case basis. Statistical analyses were then conducted using SAS (Statistical Analysis System) software system. In what follows, a description of all the codified features is provided.

i) Negotiation strategies:

Following the studies summarised in the literature review, all the NoM strategies found in our data were coded (namely comprehension checks, acknowledgements, utterance

completions, clarification requests, confirmation checks, self-repetitions, recasts, and explicit corrections). We have considered the categorisation provided by Lázaro Ibarrola and Hidalgo (2017a) and the NoM strategies were further classified into: a) *Strategies to prevent communication breakdowns*, b) *Strategies to confirm successful communication*, c) *Strategies to repair communication breakdowns* and d) *Strategies to focus on form*. As indicated in Chapter 1, in this new categorisation, the NoM strategies are further classified according to the function they serve, so that the limitation of double-categorising some strategies is prevented. This way, the possibility of an overestimation of the total number of strategies used by the speakers is avoided as utterances that could fall within two different types (e.g. repetitions and comprehension checks) are not classified twice according to their type, but once, according to the function they perform (a limitation acknowledged by Oliver herself (1998: 381)). Example 34 below, taken from our own data base, illustrates this issue:

Example 34

1. *CHI2: |I have it in front of the door|
2. *CHI1: |**of the door?**| [Confirmation check]
3. *CHI2: |yes|

Based on Lázaro Ibarrola and Hidalgo's (2017a) categorisation, the second turn in the previous exchange has been tallied only once as a strategy meant to repair a communication breakdown, specifically a confirmation check. Following Oliver's (1998, 2009) classification would have led us to consider turn 2 as, in Oliver's words, 'a multifunctional utterance'. However, we believe that even though this utterance may be categorised under two types of strategies, it performs a single function, in this case to make sure that CHI1 has understood what her partner said correctly. Thus, this type of utterance may be considered 'multi-type' rather than 'multi-function' (Lázaro Ibarrola and Hidalgo, 2017a).

That said, we now proceed to depict the language features codified as NoM strategies for the purpose of the present dissertation, which have been classified as follows, and are illustrated with examples from our data:

a) *Strategies to prevent communication breakdowns:*i) *Comprehension checks:**Example 35*

1. *CHI1: |{I have} in the class I have a childrens| is a boy {a boy with the hair with a t-shirt brown ay} with a t-shirt black {near is}| have the hair {the xxx} like brown {like} | in her t-shirt puts six {six} :: and is in the blackboard :: **you know Gloria?**

[Comprehension check]

2. *CHI2: |yes yes yes|

Example 35 illustrates a comprehension check, a strategy used by the speaker to check that the previous utterance has been correctly understood by the listener. In this particular example, speaker CHI1 produces a relatively long utterance, in which one of the items from the poster is described in detail. After all the information is given to the listener, the speaker employs this strategy in order to see whether the interlocutor has been able to follow the speech by asking him “you know?”, this way avoiding a potential communication breakdown. Speaker CHI2 confirms he has understood and they proceed with the task.

ii) *Mere self-repetitions:**Example 36*

1. *CHI2: |is in the door?|
2. *CHI1: |yes|
3. *CHI2: |**I put in the door?**|

[Mere self-repetition]

Example 36 is an instance of mere self-repetition. CHI2 repeats her previous utterance despite the affirmative response given by her partner, to make sure CHI1 has understood her. Therefore, this self-repetition aims at preventing a communicative breakdown in communication.

b) *Strategies to confirm successful communication:*

i) *Acknowledgements:*

Example 37

1. *CHI1: |eh down in the room of the Christmas I have a girl {in} eh eh
with a {book green} (.) green book :: and is in the sofa with a
green t-shirt|
2. *CHI2: |ok| [Acknowledgement]
3. *CHI 2: |I have a girl eating a sandwich with a purple jacket {in the} (..) in
his bedroom in the bed|

In the second turn of Example 37, CHI2 acknowledges he has understood the previous utterance by saying ‘ok’ and immediately moves on to describe another item in the task. This way, both interlocutors understand that step in the task is over and another one can be taken.

ii) *Utterance completions:*

Example 38

1. *CHI2: |a little boy?|
2. *CHI1: |yes it have a t-shirt with many colours like (...)|
3. *CHI2: |**blue** (..)| [Utterance completion]
4. *CHI1: |blue red and black (.) no?|
5. *CHI2: |yes|

Example 38, on the other hand, is an instance of utterance completion and it illustrates how CHI2 assists his interlocutor by providing, in turn 3, the term CHI1 was looking for in turn 2. Thus, by means of this utterance, CHI2 is showing explicitly that he understands what the interlocutor intends to communicate.

c) *Strategies to repair communication breakdowns:*

i) *Clarification requests:*

Example 39

1. *CHI1: |where do you have the girl?|
2. *CHI2: |**what?**| [Clarification request]
3. *CHI2: |where do you have that girl|

Example 39 above features a clarification request. There is a communication breakdown and to repair it, CHI2 needs CHI1 to clarify what he or she has just said. To do this, CHI1 makes use of a very frequent form of clarification request (i.e. what) in turn 2.

ii) Confirmation checks:

Example 40

1. *CHI1: |where the girl :: who is eating the sandwich :: can be?|
2. *CHI2: |I have it in front of the door|
3. *CHI1: |**of the door?**| [Confirmation check]
4. *CHI2: |yes|

An example of a confirmation check is observed in the third turn in (40), when CHI2 partially repeats the interlocutor's previous utterance in order to confirm whether he has understood or heard it correctly. As can be seen in this example, this strategy takes the form of a partial other-repetition, which, according to Oliver (2009), happens quite frequently.

iii) Repair self-repetitions:

Example 41

1. *CHI2: |the girl is in the slide?|
2. *CHI1: |the in the?|
3. *CHI2: |**the girl is in the slide?**| [Repair self-repetition]

Example 41 illustrates this strategy by showing how CHI2, in order to repair the communication breakdown pointed out by CHI1's clarification request in turn 2, repeats his previous utterance (turn 1) in the last turn.

d) Strategies to focus on form:

i) Explicit corrections:

Example 42

1. *CHI2: |es que no sé cómo se dice una palabra (*I don't remember how to say one word*)||como se decía (..) (*How do you say (..)*)|
2. *CHI1: |**how can I say!**| [Explicit correction]
3. *CHI2: |how can I say colgante? (*pendant*)|

In example 42 above, CHI2 is struggling to find a term in the TL and wants her partner to help her. However, before she finishes uttering her question, CHI1 directly corrects CHI2's previous utterance and gives her the TL language equivalent. This explicit correction is accepted by CHI2, who incorporates it in her next utterance.

ii) Recasts:

Example 43

1. *CHI1: |I have a girl :: that is in the sofa (.)| is with a jacket blue|
2. *CHI2: |**a blue jacket**| [Recast]
3. *CHI1: |with a blue jacket|

In the second turn in Example 43, CHI2 recasts the part of her partner's previous utterance which she noticed was not targetlike. In the next turn, CHI1 repeats the correction offered by CHI2.

We now proceed to explain the measures used to analyse the strategies illustrated above. As in previous studies, rates of use have been obtained by dividing the number of strategies over the total number of Analysis of speech units (AS-units).

According to Foster, Tonkyn and Wigglesworth (2000) an AS-unit is "a single speaker's utterance consisting of an independent clause, or sub-clausal unit, together with any subordinate clause associated with either" (Foster et al., 2000: 365). Based on these authors' definition, an independent clause may be made up of one or more phrases, which can be elaborated to a full clause from the context, or of a minor utterance of the type "oh poor woman", "thank you very much" or "yes" (Foster et al., 2000: 366). AS-units have been defined for the purpose of analysis of oral data and have been commonly used in previous studies analyzing oral production and learners' interaction (e.g. Kim and Tracy-Ventura, 2013; Lázaro Ibarrola and Hidalgo, (2017b); Sample and Michel, 2014).

In order to provide a clearer account of the guidelines we have followed when analysing AS-units, consider the following examples, taken from our participants' actual oral production. Based on Foster et al.'s (2000) study, AS-units have been marked by an upright slash (/), whereas boundaries between clauses are indicated by a double colon (::) and 'false starts, functionless repetition and self-corrections are put inside brackets {...}' (p. 365). Furthermore, coordinated phrases are considered to belong to the same AS-unit. Nevertheless, contrary to Foster et al.'s (2000: 367) proposal, and taking into

consideration the nature of our participants' speech (characterised by numerous pauses) coordinated sentences with a pause longer than 0.5 seconds in between, or the first phrase marked by rising or falling intonation, have not been considered as two different AS-units in the present dissertation, but have been tallied as only one.

Example 44

1. *CHI1: |and another boy?|
2. *CHI2: |a boy that is with a blue t-shirt|
3. *CHI1: |yes|

Example 45

1. *CHI2: |I don't know| I don't have it in the picture|

Example 46

1. *CHI2: |where is the girl that :: have a glasses and :: is reading a book?|

In example 44, one AS-unit consisting of a single clause can be observed in each turn, whereas (45) illustrates two AS-units featured in the same turn. Example 46, on the other hand, shows one AS-unit made up of three clauses: a main clause with one subordinate and one coordinated clause.

As already stated, false starts, functionless repetitions and self-corrections are not counted as AS-units; only the final version of the speaker's output has been considered in our analysis. See examples 47 and 48.

Example 47

1. *CHI1: |ehm {I have near the chair of the (..) and in front the (..)} is
behind the chair :: and in front of the door in the class|

Example 48

1. *CHI2: |{in the right side (..) no no! (..)} up the sofa where the Olentxero
is|

Examples 47 and 48 above have been counted as one AS-unit each. Example 47 consists of two clauses and (48) of only one. The first words in each example have been taken as a false start, followed by the speaker's self-correction while trying to express a single thought.

An overview of the codified NoM strategies is offered in Table 19.

Table 19 *Codified NoM strategies.*

Function	Strategy	Definition	Example	Measure	Previous studies
Preventing communication breakdowns	Comprehension checks	Speakers check whether the listeners understand.	*CHI1: is in the door in the middle of the door (.) no? *CHI2: uhum. *CHI1: you see? *CHI2: yes (.) now me.	Comprehension Checks/AS-units	Oliver (1998, 2009); García Mayo & Lázaro Ibarrola (2015)
	Mere self-repetitions	Speakers repeat their utterance to make sure they are understood.	*CHI2: you have a girl in the classroom? *CHI1: a girl in the classroom yes. *CHI1: yes I have it.	Mere Self-repetitions/AS-units	Lázaro Ibarrola & Hidalgo (2017a)
Confirming successful communication	Acknowledgements	Listeners confirm the previous utterance has been understood.	*CHI1: the boy with the jacket orange is in bedroom? *CHI2: no no is (..) no is a boy is a girl. *CHI1: ok (.) the girl with the book is in the classroom?	Acknowledgements /AS-units	Lázaro Ibarrola & Hidalgo (2017a)
	Utterance completions	Speakers complete the interlocutors' utterance.	*CHI1: in what door? *CHI2: in the picture of the (..) in the (...) in the (...). *CHI1: in the class?	Utterance Completions / AS-units	Lázaro Ibarrola & Hidalgo (2017a)

Repairing communication breakdowns	Clarification requests	Listeners ask the interlocutor to clarify what he or she had just said.	*CHI2: the boy with the biscuit are in your xxx? *CHI1: what? *CHI2: the boy with the biscuit are in your picture?	Clarification Requests/AS-units	Oliver (1998, 2009); García Mayo & Lázaro Ibarrola (2015)
	Confirmation checks	Listeners make sure they have understood the previous utterance correctly.	*CHI1: the girl with the book is in the classroom? *CHI2: classroom? *CHI1: Uy in the classroom (.) in the bedroom?	Confirmation Checks/AS-units	Oliver (1998, 2009); García Mayo & Lázaro Ibarrola (2015)
	Repair Self-repetition	Speakers, in order to repair a communication breakdown, repeat a previous utterance.	*CHI2: you have a girl in the playground? *CHI1: a? *CHI2: girl in the playground?	Repair Self-repetitions / AS-units	Lázaro Ibarrola & Hidalgo (2017a)
Focusing on form	Explicit corrections	Listeners make the speakers aware of a mismatch between their utterance and the TL.	*CHI2: it is on the door of the second on the picture of de abajo. *CHI1: down. *CHI2: yes.	Explicit corrections /AS-units	Lázaro Ibarrola & Hidalgo (2017a)
	Recasts	Reformulation of a previous utterance.	*CHI1: and with the (...) glasses (.) the left (.) *CHI2: at the left of the door. *CHI1: yes.	Recasts/ AS-units	Oliver (2009)

ii) L1 use

All the utterances containing L1 (Spanish) terms were coded. From the different classifications of L1 functions (Alegría de la Colina and García Mayo, 2009; Muñoz, 2007; Storch and Aldosari, 2010), we have followed Azkarai and García Mayo (2015) as the uses they describe provide a very adequate categorisation for the functions we have identified in our data: metacognitive talk, phatics, off-task talk, and vocabulary¹¹. The phatic function however has been substituted for discourse markers in our study, which include different word classes such as conjunctions, interjections, adverbs and lexicalised phrases (Schiffrin, 2006: 321). Following Muñoz's (2007) classification, within the vocabulary function we have included appeals for help, borrowings and foreignisings. In what follows, each of the functions the L1 served in our study, specifically a) metacognitive talk, b) discourse markers, c) off-task talk, and d) vocabulary, are illustrated with corresponding examples.

- a) **Metacognitive talk**: The speakers' use their L1 to talk about the task itself. This category comprises functions such as task management, planning or checking for understanding of the task. To give an illustration of this use, see examples 49 and 50, in which the speakers organise the conversation turns.

Example 49

1. *CHI1: |**empiezo yo**| (*I start*)

Example 50

1. *CHI2: |**me toca a mí :: (.) verdad**| (*now me (.) right?*)

- b) **Discourse markers**: This use of the language is intended to serve social and interactive purposes rather than to convey specific information. Typical instances of this function are expressions such as 'well', 'ok', or 'so'. In example 51, the speaker uses a discourse marker to start her contribution.

¹¹ Nevertheless, in our data we have not found one of the categories they list, namely grammar talk, probably due to the different age of the participants (i.e. university students in Azkarai and García Mayo (2015) whereas we have worked with YLs) and the type of task employed (Azkarai and García Mayo's (2015) used two speaking tasks (PPT and picture differences) and two speaking+writing tasks (dictogloss and text editing)).

Example 51

1. *CHI1: |**bueno** (*well*) yes a little bit|

c) *Off-task talk*: The L1 is used to talk about a topic that is not related to the task itself. See example 52, in which the participants talk about the temperature in the classroom.

Example 52

1. *CHI2: |**tío hace tanto calor** (*man it's so hot*)|

2. *CHI1: |**yo tengo frío** (*I'm cold*)|

3. *CHI2: |**pues yo calor** (*well I'm hot*)|

d) *Vocabulary*: The L1 is used to deliberate over vocabulary, to ask for unknown words or simply to replace an unknown L2 term. Based on Muñoz's (2007) categorisation, within this function three subtypes are included:

d.1 Appeals for help: When speakers ask for an L2 word in their L1. See example 53, in which the speaker uses a whole sentence in the L1 to ask about an unknown vocabulary item.

Example 53

1. *CHI1: |¿**cómo se dice galleta?**| (*how do you say cookie?*)

d.2 Borrowings: The use of an L1 term (or from any other previous language) in the TL discourse (Muñoz, 2007). In Example 54 below, the learner replaces an L2 term (*purple*) with the corresponding one from the L1 (*morado*).

Example 54

1. *CHI1: |the girl of the jacket eh **morado** (*purple*) are in the (..)?|

d.3 Foreignisings: Phonological or morphological adaptations of an L1 term to the rules of the TL (Muñoz, 2007). See Example 55.

Example 55

1. *CHI2: |you have a boy with the t-shirt :: put SIX next to the teacher table?|
2. *CHI1: |no|
3. *CHI2: |{to the table of the (.) to the} next to the table of the **alums** (*pupils*)?|
4. *CHI1: |no|

In this exchange, the participants are trying to locate one of the children in their posters, and in turn 3, CHI1 uses ‘*alums*’, from Spanish ‘*alumnos*’, when meaning *pupils*.

In order to obtain rates of use, the number of L1 uses has been divided by the total number of AS-units. L1 uses range from individual L1 terms, as in examples 54 and 55, to whole AS-units, as in (53). The basic guideline followed has been to tally as one item the use of the L1 that intended to convey a single idea. Each L1 function has also been analysed individually in order to see which one was most frequently used by our participants.

Table 20 displays the L1 functions detailed above.

Table 20 Codified L1 functions.

Function		Definition	Example	Measure	Previous studies	
L1 use	Metacognitive	L1 to talk about the task itself.	*CHI2: is a (..) girl with a jacket with flowers? *CHI1: yes. *CHI2: ah ya sé quien es! (<i>oh now I know who it is!</i>)	Metacognitive Talk/ AS-units	Alegría de la Colina & García Mayo (2009); Azkarai & García Mayo (2015); Muñoz (2007); Storch & Aldosari (2010)	
	Discourse markers	L1 for interactive purposes.	*CHI1: the girl with the sandwich is in the next to the flowers red? *CHI2: uhmmm? *CHI1: bueno (<i>well</i>) next to the flowers.	Discourse Markers/ AS-units	García Mayo & Hidalgo (2017)	
	Off-task	L1 to talk about an unrelated topic to the task.	*CHI2: tío para de tocarme (<i>man stop touching me!</i>)	Off-task / AS-units	Alegría de la Colina & García Mayo (2009); Azkarai & García Mayo (2015)	
	Vocabulary	Appeals for help	L1 to ask for unknown vocabulary.	*CHI2: where is the boy that have a (.) ¿cómo se decía cortina? (<i>how do you say curtain?</i>)		Muñoz (2007)
		Borrowings	L1 terms.	*CHI2: the girl are reading? *CHI1: yes. *CHI2: what colour is the (.) camiseta (<i>t-shirt</i>)?	Vocabulary/ AS-units	Muñoz (2007)
Foreignisings		L1 terms adapted to the TL.	*CHI1: a boy with a jacket blue and red xxx with a (..) in the head eh a (.) som (<i>hat</i>) colour white.		Muñoz (2007)	

iii) General competence: Complexity, accuracy and fluency (CAF)

To analyse our participants' general competence (CAF), we have based our analysis on the measures used by Sample and Michel (2014) and Bret Blasco (2014), as these studies were carried out in EFL settings and with YLs of a similar age to that of our participants. Thus, we codified for measures of structural and lexical complexity, specific and global measures of accuracy and speed fluency. Based on the two studies mentioned above, we have tallied the total number of words, clauses and AS-units. Specifically, the following measures have been used:

- **Complexity:**

a. Structural complexity:

- i) Ratio of words per AS-unit (Sample and Michel, 2014).
- ii) Complexity by coordination (ratio of clauses per AS-unit) (Bret Blasco, 2014). It must be noted here that Bret Blasco employed a *unit* the author designed for her research study. In the current dissertation however, we have used AS-units instead.

b. Lexical complexity: Vocabulary diversity measured with D value (which calculates type-token ratio against increasing token size) (Sample and Michel, 2014). A higher level of D indicates greater lexical diversity and, therefore, a richer vocabulary. This measure is implemented with the CLAN program command *VocD* (MacWhinney, 2000).

- **Accuracy:**

a. Global accuracy:

- i) Proportion of error-free AS-unit.
- ii) Proportion of errors per AS-unit.

AS-units which did not contain any type of morphological, syntactic or lexical error have been considered as error-free. Nevertheless, pronunciation errors have not been taken into account as they lie beyond the scope of the current dissertation. Consider examples 56 and 57 as illustrations of our participants' error-free AS-units (EFAS) and an exchange containing several non-error-free AS-units (NEFAS) respectively.

Example 56

1. *CHI1: |{the boy} where is the boy?| [EFAS]
2. *CHI2: |{the boy is in the (..)} how do you say *portería*?| [EFAS]
3. *CHI1: |{in the} (..) a goal place? | I don't know| [EFAS]
4. *CHI2: |the boy is in the goal| [EFAS]

In the example above, we can see an exchange in which all the AS-units have been counted as error-free.

Example 57

1. *CHI1: |it's a girl?| [NEFAS]
2. *CHI2: |yes|
3. *CHI1: |ok|
4. *CHI2: |eh it's a boy?| [NEFAS]
5. *CHI1: |yes :: what clothes (..) (..) have?| [NEFAS]
6. *CHI2: |the girl have a t-shirt (..) purple| [NEFAS][NonCV]

During the exchange illustrated in example 57 above, different error types can be observed. The first two questions by CHI1 (turn 1) and CHI2 (turn 4) do not follow the question structure of subject-verb inversion. The speakers employ the order of a declarative sentence with rising intonation instead, that is, they seem to be at the second stage in question formation (Pienemann and Johnston, 1986). The next question by CHI1 (turn 5) portrays again difficulties with question formation rules in the TL (English). In this instance, not only is the auxiliary verb (*does*) omitted but also the subject of the sentence (*the girl/she*), representing also a clear example of L1 (Spanish) structural transfer. The last utterance in this exchange (turn 6) also contains an error in the verb form, as it does not comply with subject-verb agreement rules. This utterance also illustrates the type of error chosen for the specific accuracy measure in the present dissertation (correct verb use), which is described below.

b. Specific accuracy: Since our task mainly required the description of places, objects and children, and this type of discourse requires the use of verbs, the proportion of correct verb use (in terms of tense and subject-verb agreement) has been analysed (correct verb forms/total number of verbs used) (Bret Blasco, 2014). See examples 57 (above), 59 and 60 for instances of what has been tallied

as incorrect verb forms (NonCV), and example 58 below to illustrate a correct verb form (CV).

Example 58

1. *CHI2: |where is the boy?| [CV]
2. *CHI1: |the boy is {in the (..) is eh (..)} in front of the table| [CV]

Example 59

1. *CHI1: |was a girl in the playground?| [NonCV]
2. *CHI2: |a girl?|

In example 59, although the subject-verb agreement requirement is fulfilled, the tense used is not appropriate for the communicative situation. Additionally, other typical cases which have been taken as NonCV are those in which the verb, even when adopting adequate morphological forms (e.g. tense and number), did not match the communicative needs of the situation, that is, it did not convey the meaning required. See example 60:

Example 60

1. *CHI1: |you have a girl in a blackboard?| [NEFAS]
2. *CHI2: |yes|
3. *CHI1: |you are a girl in the park?| [NEFAS][NonCV]
4. *CHI2: |no|

The AS-unit in the third turn in Example 60, apart from being tallied as a NEFAS because of the incorrect question structure, the verb choice in it has also been considered as NonCV, since CHI1 asks the interlocutor whether he *is a girl in the park*, instead of asking, for instance, if *there is*.

- **Fluency:**

Following Bret Blasco (2014), we have analysed speed fluency and breakdown (dis)fluency. We have also taken into account our participants' level of proficiency (i.e. A1), as well as the characteristics of their oral production, which normally contains numerous pauses as well as false starts and repetitions.

- a. *Speed fluency:* Number of words produced by the participants divided by the total number of minutes needed to complete the task (Bret Blasco, 2014; Sample and Michel, 2014). Thus, an increase in the number of words per minute will be interpreted as an increase in fluency.

- b. Breakdown (dis)fluency:* The occurrence of L1 use over the total number of AS-units. A great value in this measures implies little fluency.

In order to offer a more visual illustration of the analysis of these aspects of speech, Table 21 summarizes the CAF measures adopted in the current study.

Table 21 CAF measures.

Complexity	Structural	Words/AS-unit Clauses/AS-unit
	Lexical	D value
Accuracy	Specific	Correct verb forms/total number of verbs
	Global	Errors/AS-unit Error-free AS-units/total number of AS-units
Fluency		Words/minute
		L1/AS-units

In what follows, the statistical analyses conducted in order to examine our data will be introduced. The data, coded in terms of the above described range of measures of NoM strategies, CAF and L1 use, were run through the corresponding statistical tests. Two different tests have been carried out in order to investigate the effects of procedural TR and age group, the two independent variables in our study. Generalised linear mixed-effects model (GLMEM) tests have been run because of the count nature of two of our dependent variables (i.e. NoM strategies and L1 use). For the analyses of the continuous data in CAF, a general linear mixed model (GLMM) has been used. One test has been run for each dependent variable. These tests allow for the consideration of the potential correlation of the participants working in pairs in the two groups as well as testing time effect. GLMEM allows for Poisson mixed-effect model and random effects, while the GLMM takes mixed-effect model and random effects. The significance level was fixed at $\alpha = 0.05$.

Additionally, with the aim of assessing the relationship between the different aspects of the participants' general performance (CAF), Spearman's rank-order correlation tests have also been carried out. This test has been selected instead of Pearson's because the variables analysed did not present a normal distribution.

In this section we have tried to provide a clear account of the operationalisation of the areas of L2 knowledge considered in the current dissertation in order to analyse young EFL learners' L2 oral production. The data coding system, namely, the codification of NoM strategies, L1 use and CAF, has been detailed and illustrated with examples from our own dataset. In the next chapter, the results obtained in the quantitative analysis are presented.

CHAPTER 6 RESULTS

This chapter discusses the results obtained from the analyses of our data in order to answer the research questions posited in Chapter 5. The different aspects each question addresses (i.e. the impact of age and procedural TR on NoM strategies, CAF and L1 use) have been investigated separately. Thus, the chapter is divided into three sections. First, we provide the results regarding the amount and type of NoM strategies the participants in Year 3 and Year 5 used at T1. In the second section, the results regarding the effect of procedural TR on the performance of both groups are presented. Additionally, we analyse whether there are differences between the two age groups in order to find out any age effects in the way YLs complete the task, specifically how they negotiate for meaning, the development of their general performance (CAF) and L1 use¹². Finally, a summary of the results will be offered in the last part of the chapter.

6.1 Negotiation strategies in Year 3 and Year 5

As noted above, only the results obtained at T1 will be presented in this section. Since TR may have an effect on the number of strategies learners use, the first time the participants in this study faced the task has been taken as the one that would best represent the way they negotiate for meaning. Consequently, the data collected at this task time have been analysed to examine this aspect.

As can be seen in Table 22, the examination of the participants' performance reveals that the learners in the two age groups (8-9 and 10-11 year-olds) negotiate for meaning when engaged in conversational interaction with age- and level-matched peers. The participants in this study employed all the NoM strategies identified in the literature, namely comprehension checks and mere self-repetitions, acknowledgements and utterance completions, clarification requests and confirmation checks, explicit corrections and recasts. However, the analysis of the oral production of the two groups showed a different use of some of these strategies. These differences become most evident in the type of NoM strategies each group employed most frequently: whereas Year 3 learners mainly used clarification requests (26.95%) and repair self-repetitions (23.4%), the most common strategy by far in the production of Year 5 learners are acknowledgements (which represent 41.99% of the total amount of strategies employed by these learners).

¹² Supplemental tables displaying the results that did not reach statistical significance and/or have not been detailed in the current chapter are presented in Appendix D.

As for the least used NoM strategies, the two groups coincide in a minimal use of three strategies: comprehension checks (Year 3: 1.42%; Year 5: 2.76%), explicit corrections (Year 3: 0.71%; Year 5: 0%), and recasts (Year 3: 2.13%; Year 5: 1.1%). As illustrated in Table 22, the rest of the NoM strategies identified in these learners' oral production follow different patterns in the output of each group.

Table 22 *NoM strategies used by Year 3 and Year 5 learners during oral interaction at T1.*

Functions	Strategies	Year 3		Year 5	
		Counts	Percentage	Counts	Percentage
Preventing communication breakdowns	Comprehension checks	2	1.42%	5	2.76%
	Mere self-repetitions	14	9.93%	12	6.63%
Confirming successful communication	Acknowledgements	21	14.9%	76	41.99%
	Utterance Completions	5	3.55%	19	10.5%
Repairing communication breakdowns	Clarification request	38	26.95%	21	11.6%
	Confirmation check	24	17.02%	26	14.36%
	Repair self-repetitions	33	23.4%	20	11.05%
Focusing on Form	Explicit correction	1	0.71%	0	0%
	Recasts	3	2.13%	2	1.1%

We will not go into more detail regarding the individual NoM strategies since, as mentioned in the previous chapter, we have followed Lázaro Ibarrola and Hidalgo's (2017a) categorisation of NoM strategies based on the functions these strategies serve, which, from now onwards, will be the basis of our analysis.

The rates of use of some of the general functions served by the NoM strategies follow the same trend as the use of the individual strategies described above and, consequently, some differences between the two groups are observed. Table 23 details the counts and rates of use of the functions of the strategies to negotiate for meaning employed by the participants in our study.

Table 23 Percentages of the functions served by the *NoM strategies used by Year 3 and Year 5 learners during oral interaction at T1.*

	Year 3		Year 5	
	Counts	Percentage	Counts	Percentage
Preventing communication breakdowns	16	11.35%	17	9.39%
Confirming successful communication	26	18.44%	95	52.49%
Repairing communication breakdowns	95	67.37%	67	37.02%
Focusing on Form	4	2.84%	2	1.10%

Beginning with the analysis of the functions served by the NoM strategies used by the participants in Year 3, our results show that, clearly, the most frequent function in these learners' production is to repair communication breakdowns. This type has been identified 95 times in their first performance, which represents 67.37% of the total amount of strategies employed at T1. This function is followed by strategies used to confirm successful communication, employed 26 times which correspond to 18.44% of the total NoM strategies. Strategies to confirm successful communication are less frequent (to a statistically significant degree) than those used to repair communication breakdowns in the performance of the Year 3 group ($t = -5.36, p < .001$). Preventing communication breakdowns is the third most frequent function, identified on 16 occasions which represent 11.35% of the total strategies. This function is statistically significantly less frequent than strategies used to repair communication breakdowns, which, as already stated, is the most common function in the output of this group ($t = -6.03, p < .001$). However, the differences between NoM strategies to confirm successful communication and those to prevent communication breakdowns do not reach statistical significance ($t = 1.40, p = .86$). Finally, only 4 strategies to focus on form were identified in the production of this group, representing 2.84% of the total identified in these learners' first performance. Not surprisingly, the rate of use of this strategy is statistically significantly lower than that of the two most common ones (focus on form vs. confirming successful communication ($t = -3.19, p = .034$); focus on form vs. repairing communication

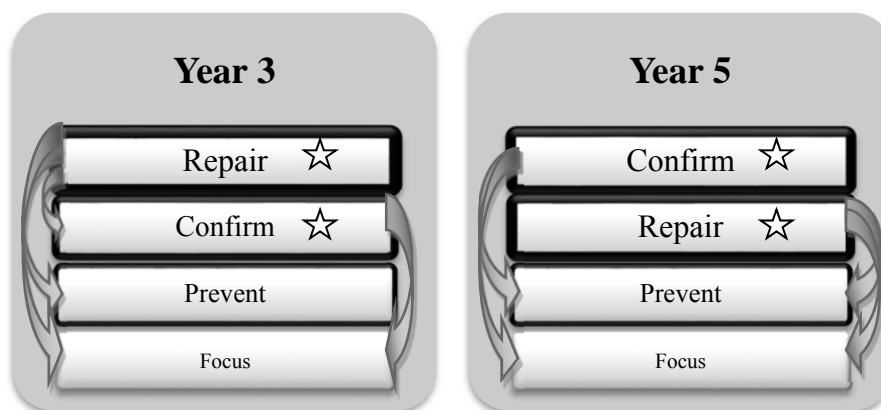
breakdowns ($t = -5.68, p < .001$). In contrast, no statistically significant differences were found between strategies used to focus on form and those to prevent communication breakdowns, that is, between the two less frequent functions ($t = -2.27, p = .314$). All in all, it is evident that in these learners' interactions there is a clearly prevalent function served by the NoM strategies: to repair communication breakdowns. There are two other functions that are recurrent, but not as frequently employed, namely strategies used to prevent communication breakdowns and those to confirm successful communication. Lastly, there is one function that is merely nonexistent: focus on form.

On the other hand, the most frequent function of the NoM strategies employed by the participants in Year 5 is to confirm successful communication during interaction, employed 95 times in the first task, which represents 52.49% of the total number of strategies. This function, however, is not statistically significantly more frequent than that to repair communication breakdowns ($t = -2.00, p = .48$), which is the second most commonly employed by the learners in this age group, with 67 instances identified constituting 37.02% of the total sample. As in the Year 3 group, the third most frequent function is to prevent communication breakdowns, which appears 17 times in our data (9.39%). The differences between the second and the third most frequent functions are statistically significant ($t = -4.53, p = .0002$). This function is also significantly less frequent than the most common one in the oral production of the Year 5 group, to confirm successful communication ($t = -5.92, p < .001$). Finally, corroborating the results obtained in the analysis of the output of the learners in Year 3, focus on form is the least common by far, used only twice, adding up to 1.10% of the total amount of NoM strategies identified in the production of the older YLs. The strategies serving this function, as was the case in the Year 3 group, are statistically less frequently used than the two most common functions (focus on form vs. confirming successful communication ($t = -4.94, p < .0001$); focus on form vs. repairing communication breakdowns ($t = -4.48, p = .0003$)). However, no statistically significant differences were found between the number of strategies used to focus on form and those to prevent communication breakdowns ($t = -2.70, p = .13$). Thus, in the older YLs' production two common functions were identified and both were equally frequent when compared to the rates of use of the other functions: NoM strategies to confirm successful communication and those to repair communication breakdowns. This group of older YLs does not only negotiate to repair communication breakdowns, as their younger counterparts do, but also to confirm successful

communication. On the other hand, their ability to negotiate in order to prevent communication breakdowns seems to be still low, and their focus on form, lacking.

Figure 6 displays the differences between the rates of use of the NoM strategies in the production of each group. The stars and the arrows mark the functions which were statistically significantly different from one another within the same age group.

Figure 6 Differences between the functions served by NoM strategies in each group.



Post-hoc tests were also run in order to investigate if there were significant differences in the use of NoM strategies by these two age groups. The results reveal statistically significant differences only in the use of strategies to confirm successful communication, which were much more frequent in the production of the older learners: Year 5 learners employed 95 NoM strategies with this function whereas Year 3 learners only 26 ($t = -4.47, p = .0003$). With respect to the rest of the functions, the production of the two groups of participants did not differ significantly. Table 24 details these findings (statistical significance has been marked with an asterisk).

Table 24 *Within- and between-group differences in the functions served by NoM strategies at T1.*

Strategy	Group	Strategy	Estimate	SD	DF	t Value	Pr > t	Adj P
Confirm	3	Focus	1.8718	0.587	274	3.19	0.0016	0.0337*
Confirm	3	Prevent	0.4855	0.347	274	1.40	0.1632	0.8575
Confirm	3	Repair	-1.2958	0.242	274	-5.36	<.0001	<.0001*
Prevent	3	Repair	-1.7813	0.295	274	-6.03	<.0001	<.0001*
Focus	3	Prevent	-1.3863	0.611	274	-2.27	0.0240	0.3144
Focus	3	Repair	-3.1676	0.558	274	-5.68	<.0001	<.0001*
Confirm	5	focus	3.8607	0.781	274	4.94	<.0001	<.0001*
Confirm	5	Prevent	1.6635	0.281	274	5.92	<.0001	<.0001*
Confirm	5	Repair	0.3492	0.174	274	2.00	0.0462	0.4817
Focus	5	Prevent	-2.1972	0.815	274	-2.70	0.0074	0.1277
Focus	5	Repair	-3.5115	0.784	274	-4.48	<.0001	0.0003*
Prevent	5	Repair	-1.3143	0.290	274	-4.53	<.0001	0.0002*
Focus	3 – 5	Focus	0.6929	0.960	274	0.72	0.4710	0.9963
Confirm	3 – 5	Confirm	-1.2960	0.290	274	-4.47	<.0001	0.0003*
Prevent	3 – 5	Prevent	-0.1180	0.408	274	-0.29	0.7728	1.0000
Repair	3 – 5	Repair	0.3490	0.240	274	1.47	0.1420	0.8213

All in all, it can be seen that the two groups of participants negotiated for meaning when completing the PPT. The two most frequent strategies in the production of the two groups coincided: repairing communication breakdowns and confirming successful communication were the functions performed by most of the strategies identified in our dataset. Whereas the counts and statistical analysis indicate that Year 3 learners negotiate mainly to repair communication breakdowns, Year 5 children seem to negotiate not only to repair breakdowns, but also to confirm successful communication, as no statistically significant differences were identified between these two functions in the output of the older YLs. On the other hand, both groups were similar in their use of NoM strategies to prevent communication breakdowns doing so on fewer occasions than the two most frequent, and focus on form was certainly rare in the production of the two age groups. Regarding individual NoM strategies that serve these functions, a parallel pattern

emerged in the output of the two groups: the learners in both groups used the three strategy types to repair communication breakdowns (namely clarification requests, confirmation checks and self-repetitions) similarly. Within the confirming successful communication function, acknowledgment of understanding was the most frequent strategy in the production of both groups. Even though strategies to prevent communication breakdowns were not as frequent as the two previous ones, and comprehension checks were scarce, the learners mainly resorted to another strategy that serves this function (i.e mere self-repetitions). Explicit corrections and recasts to focus on form were equally uncommon. Finally, with regard to the comparison of the two groups, only the strategies employed to confirm successful communication were found to be significantly more frequent in the output of the older learners, whereas no differences were identified among the rest of the functions at T1.

6.2 Effect of procedural task repetition and age on learners' oral performance

This section describes the results obtained from the GLMEM and GLMM tests performed to identify the main effects and interaction effects of the two independent variables, namely testing time (T1, T2 and T3) and age group (Year 3 and Year 5) on the learners' oral production (NoM, CAF and L1 use) in an attempt to identify the differences between the performance of the two age groups across tasks. Post-hoc tests were performed to investigate where exactly significant differences occurred, both across tasks and between groups. By means of these tests, we will identify the main effect¹³ for time and the main effect for group on each variable and will be able to determine whether the mean score of a specific dependent variable is significantly different across the three testing times in the output of the two age groups. Additionally, we will report the effect of the interaction between task time and group. If the interaction of the independent variables was found to be significant, it would not be possible to determine which factor exerts a larger influence. As all the participants performed the task three times, the testing time (T1, T2 and T3) served as the within-subjects factor. The different groups (Year 3 and Year 5), on the other hand, served as the between-subjects factors. The set of results provided first for each variable when time effects were identified corresponds to the analyses of the learners' production as a whole, not differentiating between groups, whereas group

¹³ A "main effect" is the effect of one of our independent variables on one dependent variable, ignoring the effects of all other independent variables.

effects refer to the production of each group not considering testing times. Finally, the results of the analyses of the oral performance of each group at each testing time, as well as the differences in the changes each group experienced across tasks, will be reported.

In subsection 6.2.1 we present the results from the analysis of the development of the learners' NoM. In 6.2.2 we examine the CAF measures considered in this study. The results for each of the measures used to investigate CAF at the three data collection points (T1, T2 and T3) will be provided separately. Subsection 6.3 focuses on the last aspect analysed, the participants' use of the L1 and the functions it served during oral interaction at each testing time.

6.2.1 Impact on negotiation strategies

In this section we will report the results obtained from the comparison of the use of negotiation strategies by Year 3 and Year 5 learners in the three tasks in order to find any potential effect of procedural TR on YLs' NoM, as well as any differences between the two age groups. Token counts, percentages, means, standard deviations and minimum and maximum values of the four functions served by NoM strategies at each testing time of the two groups are reported in Table 25.

Table 25 Descriptive statistics for the functions served by the NoM strategies in the performance of the two age groups across testing times (T1, T2, T3).

Group	Time	Variable	Number	Percentage	Mean	SD	Min.	Max.
Year 3	1	Prevent	16	11,35%	0.40	0.84	0	4
		Confirm	26	18,44%	0.65	0.86	0	3
		Repair	95	67,37%	2.37	2.14	0	11
		Focus	4	2,84%	0.10	0.30	0	1
	2	Prevent	13	10,74%	0.32	0.65	0	3
		Confirm	30	24,79%	0.75	0.87	0	3
		Repair	76	62,81%	1.90	1.92	0	9
		Focus	2	1,65%	0.05	0.22	0	1
	3	Prevent	7	6,31%	0.17	0.45	0	2
		Confirm	39	35,14%	0.97	1.86	0	10
		Repair	62	55,86%	1.55	1.58	0	6
		Focus	3	2,70%	0.07	0.27	0	1
Year 5	1	Prevent	18	9,89%	0.45	0.87	0	4
		Confirm	95	52,2%	2.37	1.96	0	7
		Repair	67	36,81%	1.67	2.22	0	12
		Focus	2	1,1%	0.05	0.22	0	1
	2	Prevent	24	13,11%	0.6	1.33	0	8
		Confirm	97	53%	2.42	2.19	0	9
		Repair	61	33,33%	1.52	1.68	0	8
		Focus	1	0,55%	0.02	0.16	0	1
	3	Prevent	4	2,86%	0.10	0.38	0	2
		Confirm	93	66,43%	2.32	2.36	0	8
		Repair	40	28,57%	1	0.96	0	3
		Focus	3	2,14%	0.07	0.27	0	1

The employment of NoM strategies to prevent communication breakdowns (namely comprehension checks and mere self-repetitions) and to repair communication breakdowns (confirmation checks, clarification requests and self-repetitions) by Year 3 and Year 5 groups follows a similar pattern, and no statistically significant group effects were found on these two NoM functions (preventing: $F(1, 38) = 0.04, p = .8337$; repairing: $F(1, 38) = 3.07, p = .08800$). As displayed in Table 25 above, in the performance of the two groups the use of these strategies slightly diminishes from T1 to T2 (except for Year 5 strategies to prevent breakdowns which increase at this testing time), followed by a sharper fall at T3 compared to the initial score obtained at T1. Not surprisingly, significant effects of time were identified for these two functions (preventing: $F(2, 196) = 9.00, p =$

.0002; repairing: $F(2, 196) = 5.45, p = .0050$). Table 26 shows the results obtained in the analyses of the influence of testing time and age group on our participants' use of strategies to negotiate for meaning.

Table 26 Main effects and interaction effects of task time and age group on NoM strategies.

	Effect	Num. DF	Den. DF	F Value	Pr > F
Prevent communication breakdowns	Group	1	38	0.04	0.8337
	Time	2	196	9.00	0.0002*
	Group*time	2	196	2.50	0.0847
Confirm successful communication	Group	1	38	19.45	<.0001*
	Time	2	196	0.80	0.4515
	Group*time	2	196	1.08	0.3400
Repair communication breakdowns	Group	1	38	3.07	0.0880
	Time	2	196	5.45	0.0050*
	Group*time	2	196	0.33	0.7177
Focus on form	Group	1	38	0.74	0.3963
	Time	2	196	0.56	0.5719
	Group*time	2	196	0.21	0.8080

The comparisons of the output of all the participants, not considering them as two individual groups but as only one, show how the decrease that took place between T1 and T3 in the learners' use of these two functions is statistically significant (preventing: $t = 3.74, p = .0007$; repairing: $t = 3.29, p = .0034$). Moreover, statistically significant differences were found between T2 and T3 in the use of NoM strategies to prevent communication breakdowns ($t = 4.19, p < .0001$). The differences between T2 and T3 in the use of strategies to repair communication breakdowns, however, did not reach significance ($t = 2.24, p = .0664$).

As regards learners' use of strategies to confirm successful communication (acknowledgements and utterance completions), no significant effects of task time were found ($F(2, 196) = 0.80, p = .4515$). The number of strategies serving this function in the output of the two groups remains stable across task performances. In contrast, group effects were identified, indicating that the two groups used strategies serving this function in different ways ($F(1, 38) = 19.45, p < .001$). As indicated by the results of the post-hoc tests, Year 5 learners produced a significantly larger number of instances of strategies to

confirm successful communication at each testing time than Year 3 learners did (See Table 1 in Appendix D).

Finally, the last function served by NoM strategies, focusing on form (which comprises explicit corrections and recasts), was found to be the least common by far in the production of the two groups of YLs. The results obtained from the comparisons of the use of this function at the three testing times show effects neither for task time nor for age group on the number of strategies performing this function (Age: $F(1, 38) = .74, p = .40$; Time: $F(2, 196) = 0.56, p = .57$). More detailed information about the differences found between the NoM strategies identified at each testing time is presented in Table 27.

Table 27 Differences in the use of NoM strategies across task times.

	Time	Estimate	SD	DF	t Value	Pr > t	Adj P
Prevent communication breakdowns	1-2	-0.04002	0.1903	196	-0.21	0.8336	0.9759
	1-3	1.1654	0.3116	196	3.74	0.0002	0.0007*
	2-3	1.2054	0.2874	196	4.19	<.0001	0.0001*
Confirm successful communication	1-2	-0.08197	0.1422	196	-0.58	0.5651	0.8329
	1-3	-0.1921	0.1544	196	-1.24	0.2149	0.4287
	2-3	-0.1101	0.1323	196	-0.83	0.4063	0.6835
Repair communication breakdowns	1-2	0.1585	0.1228	196	1.29	0.1985	0.4024
	1-3	0.4713	0.1432	196	3.29	0.0012	0.0034*
	2-3	0.3128	0.1394	196	2.24	0.0259	0.0664
Focus on form	1-2	0.6931	0.7541	196	0.92	0.3591	0.6288
	1-3	-0.05889	0.5819	196	-0.10	0.9195	0.9944
	2-3	-0.7520	0.7393	196	-1.02	0.3103	0.5669

In what follows, we introduce the differences between the two groups. The final step taken to analyse these learners' NoM are the post-hoc tests to determine the within-and between-subjects differences across tasks. Concerning the use of NoM strategies to prevent communication breakdowns, only the drop in the number of instances of this function in the production of the Year 5 Group from T1 to T3 and from T2 to T3 was statistically significant (T1-T3: $t = 3.12, p = .0254$; T2-T3: $t = 4.06, p = .0010$). The differences between groups and tasks were more pronounced as regards strategies to confirm successful communication: statistically significant differences were found between the three performances of the two groups (T1: $t = -4.12, p = .0008$; T2: $t = -3.86, p = .0021$; T3: $t = -2.99, p = .0365$). Additionally, the differences between the changes across tasks that the production of each group underwent also reached statistical

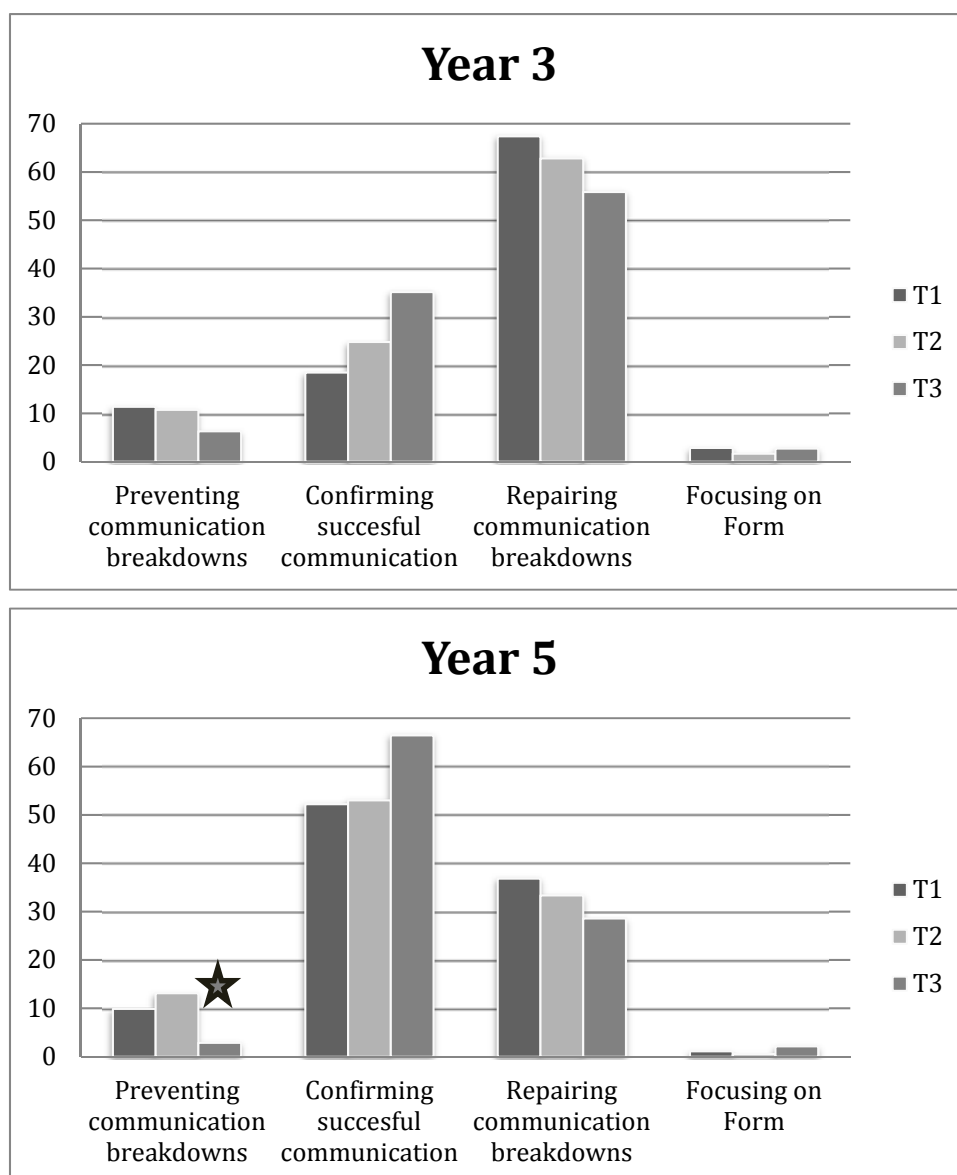
significance (Year 3 vs. Year 5: T1-T2: $t = -4.19, p = .0006$; T1-T3: $t = -4.05, p = .0010$; T2-T3: $t = -3.71, p = .0037$). The differences in the changes between the use of strategies to repair communication breakdowns from T1 to T3 in the production of the two groups also reached statistical significance (T1-T3: $t = 3.38, p = .0113$). These results are detailed in Tables 1, 2, 3 and 4¹⁴ in Appendix D.

Thus, we can conclude that in terms of NoM strategies, age had a modest impact on the learners' interactional patterns, as only one function was found to be significantly influenced by this variable, namely confirming successful communication. The strategies serving this function (acknowledgements and utterance completions) were the only ones that were used statistically significantly more frequently by one of the two groups at the three data collection points, specifically, Year 5. This finding reflects the results reported in the examination of the performance of the two groups at T1, in which this function was identified as the only one that was significantly more recurrent in the older YLs' performance. According to these results, it becomes even more evident that YLs mainly negotiate when they need to repair a communication breakdown, and that their ability to negotiate increases with age, as learners only two years older are able to employ different strategies more frequently (i.e. to confirm successful communication). Age in combination with procedural TR influenced the use of NoM strategies, as each group experienced significantly distinct changes from task to task in their use of strategies to repair communication difficulties and to confirm successful communication. On the other hand, procedural TR appears to have a higher influence than age, as significant differences were found in two of the NoM strategies, specifically those used to prevent and to repair communication breakdowns. However, these results need to be viewed with caution as they may be due to the low overall incidence of NoM strategies, which may also be related to the fact that, through TR, learners need less time and fewer utterances to complete the task, as we will see in the analysis of CAF. Task familiarity, achieved through procedural TR, facilitates task completion in a way that the use of some negotiation strategies (particularly those to prevent and to repair communication problems) seemed less necessary than when learners faced an unfamiliar task. Apart from

¹⁴ Even though neither effects of testing time nor age group were found, post-hoc tests were run on the focus on form data. The results of the within- and between-subjects are presented in Table 4 in Appendix D.

having more attentional resources to devote to language, as we will see in our analyses of CAF, YLs' reliance on other strategies that prevent breakdowns became greater.

The differences in the percentages of use of NoM strategies from task to task by the participants in the present study are featured in Graph 1. As can be seen in the graphs, the patterns of use of the different functions of the NoM strategies follow a similar trend in the two groups. The significant decrease that took place at the third performance of the Year 5 group, within the strategies used to prevent communication breakdowns, is marked with a star. In addition to this, it is worth mentioning that NoM strategies to confirm successful communication follow an increasing trend in the output of the two groups, although the differences across tasks did not reach statistical significance. In contrast, the use of strategies to repair communication breakdowns decreases. This tendency comes as no surprise since more confirmations of successful communication means that fewer communication breakdowns are taking place, hence a smaller need to repair.

Graph 1 Changes in the percentages of functions of the NoM strategies across tasks.

6.2.2 Impact on CAF measures

This section is divided into three subsections, each dealing with a different dimension of language learners' general performance, namely complexity, accuracy and fluency. We will start by reporting the results obtained in the analyses of the measures used to examine the development of complexity upon TR. The second dimension to be presented is accuracy, which will also be analysed considering the three measures this aspect comprises. Finally, the results of the examination of fluency as regards the two measures studied in the current dissertation will be provided.

6.2.2.1 Impact on oral complexity

This subsection explores the potential changes within the complexity measures in the output of the two groups at the three testing times. As stated in the previous section on Data coding and analysis, two measures were selected for the study of structural complexity: i) the ratio of words per AS-unit and the ratio of clauses per AS-unit. Lexical complexity, on the other hand, was calculated using the vocabulary diversity measure D (type-token ratio against increasing token size).

Before explicitly addressing our research question about complexity, we provide an overview of the general performance of the participants in the current study. As shown in Table 28, there is a falling trend in the raw numbers of words, clauses and AS-units tallied in our database. Nevertheless, the differences are not large. It can also be observed how individual differences come into play, as the differences between the maximum and the minimum values in these three variables (namely number of words, clauses and AS-units) in the three task performances of the two groups are large. On the other hand, the means obtained at the three data collection points are very similar.

Table 28 *Descriptive statistics of the words, clauses and AS-units of all the performances for all participants (n= 40).*

			Total	Mean	Median	SD	Min.	Max.
YEAR 3	Words	T1	4944	123.6	95	67.26	46	290
		T2	4972	124.3	111.5	60.87	55	329
		T3	4916	122.9	116.5	56.55	46	240
	Clauses	T1	1346	33.65	28.5	16.29	12	76
		T2	1257	31.43	29.5	11.13	14	58
		T3	1309	32.73	31	14.67	10	73
	AS-units	T1	1245	31.13	26	15.25	10	72
		T2	1162	29.05	26.5	10.85	11	59
		T3	1124	28.10	25.5	12.88	7	62
YEAR 5	Words	T1	5290	132.25	115	62.71	53	276
		T2	5154	128.85	108	67.74	52	312
		T3	4641	116.03	99.5	57.48	35	252
	Clauses	T1	1218	30.45	26.5	13.79	10	59
		T2	1186	29.65	25.5	14.43	13	73
		T3	1085	27.13	26.5	10.48	9	47
	AS-units	T1	1040	26	22.5	12.74	7	50
		T2	1020	25.50	21	12.14	9	62
		T3	914	22.85	22	9.45	5	42

Although there is a decreasing trend in the means of the number of words, clauses and AS-units in each task, neither time nor group effects reached significance for the two first variables, and only the effect of time was significant for AS-units ($F(1, 196) = 3.18, p = .0438$). Table 29 illustrates the results obtained in the statistical analysis of these three dimensions of our learners' production across tasks.

Table 29 *Main effects and interaction effects of task time and age group on words, clauses and AS-units.*

	Effect	Num. DF	Den. DF	F Value	Pr > F
Words	Group	1	38	0.02	0.9030
	Time	2	196	0.92	0.4022
	Group*time	2	196	0.71	0.4926
Clauses	Group	1	38	1.08	0.3063
	Time	2	196	1.24	0.2904
	Group*time	2	196	1.14	0.3205
AS-units	Group	1	38	2.22	0.1442
	Time	2	196	3.18	0.0438*
	Group*time	2	196	0.35	0.7052

Post-hoc analyses were run in order to examine the differences between tasks in more detail. The results indicate that the analysis of the production of AS-units by the two groups of YLs significantly decreases upon TR, as statistically significant differences were identified between T1 and T3 ($t = 2.51, p = .0344$). The results are presented in Table 5 in Appendix D, which also includes the results of the analyses of words and clauses.

In what follows, we present the results obtained from the analysis of the development of our complexity measures. Table 30 details the descriptive statistics for the three complexity measures addressed in our study: the ratio of words per AS-unit, the ratio of clauses per AS-unit and lexical D produced by the learners in the two groups in each task. The results of the analyses performed on the production of the two groups at the three data collection times are presented in Table 31.

Table 30 Descriptive statistics for the three complexity measures in the three tasks.

Group	Time	Variable	Mean	SD	Min.	Max.
Year 3	1	Words/AS-unit	4.03	1.20	1.81	7.09
		Clauses/AS-unit	1.10	0.14	0.78	1.55
		Lexical D	17.45	10.08	4.52	51.48
	2	Words/AS-unit	4.32	1.22	2.24	7.41
		Clauses/AS-unit	1.10	0.12	0.91	1.44
		Lexical D	14.99	8.96	2.87	37.04
	3	Words/AS-unit	4.59	1.34	1.96	7.86
		Clauses/AS-unit	1.19	0.21	0.95	2.00
		Lexical D	15.36	9.51	2.57	41.59
Year 5	1	Words/AS-unit	5.39	1.43	2.81	8.78
		Clauses/AS-unit	1.21	0.16	1.00	1.75
		Lexical D	20.32	7.21	7.48	34.71
	2	Words/AS-unit	5.12	1.11	2.82	7.61
		Clauses/AS-unit	1.17	0.12	0.97	1.44
		Lexical D	18.82	6.40	8.23	32.61
	3	Words/AS-unit	5.26	1.83	2.06	12.33
		Clauses/AS-unit	1.22	0.19	0.94	1.83
		Lexical D	18.98	7.24	6.49	39.28

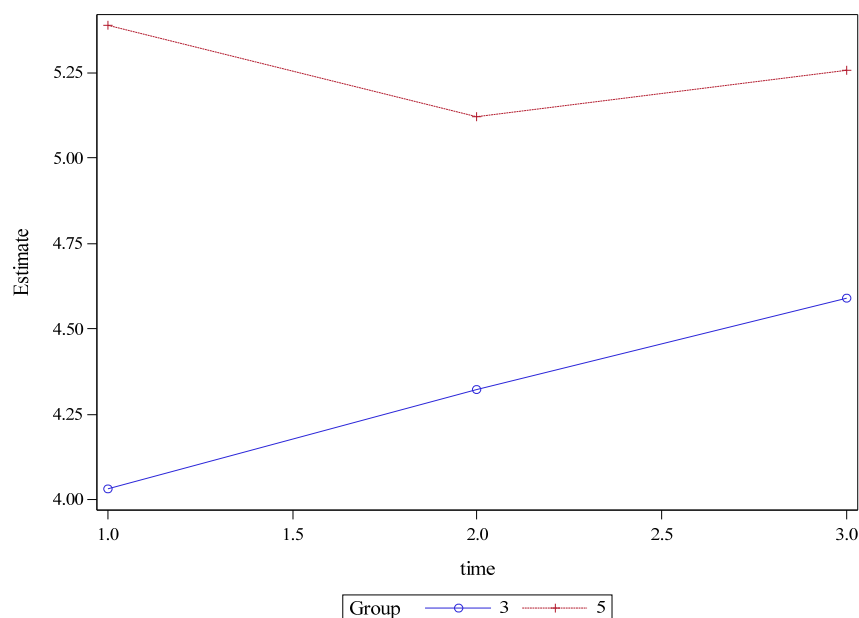
Table 31 Main effects and interaction effects of task time and age group on complexity.

	Effect	Num. DF	Den. DF	F Value	Pr > F
Words per AS-unit	Group	1	38	9.49	0.0038*
	Time	2	196	0.95	0.3894
	Group*time	2	196	2.62	0.0757
Clauses per AS-unit	Group	1	38	4.90	0.0330*
	Time	2	196	6.87	0.0013*
	Group*time	2	196	1.36	0.2594
Lexical D	Group	1	38	3.54	0.0676
	Time	2	191	2.66	0.0724
	Group*time	2	191	0.09	0.9162

First, we will introduce the results obtained from the examination of the ratio of words per AS-unit across tasks in the two groups. The analysis performed revealed significant effects of group for this variable ($F(1, 38) = 9.49, p = .0038$). Both groups in the current study show significantly different behaviours as regards the development of the ratio of words per AS-units, as Graph 2 illustrates. Results reveal a statistically significantly higher production of words per AS-unit by Year 5 learners than that of the younger

learners ($t = -3.08, p = .0038$) (Table 6 in Appendix D). The differences between the changes in the performances of these two groups of learners from T1 to T2 and from T1 to T3 were also statistically significant (T1-T2: $t = -3.10, p = .0268$, T1-T3: $t = -3.27, p = .0158$). On the other hand, procedural TR does not seem to exert an influence on the average number of words that are included in an AS-unit, as no time effects were identified (see Table 31). That is, our results reveal that, even though the two groups behaved differently across tasks, procedural TR itself did not affect the performance of either group as regards the number of words per AS-unit (but see tables 6 and 7 in Appendix D for details).

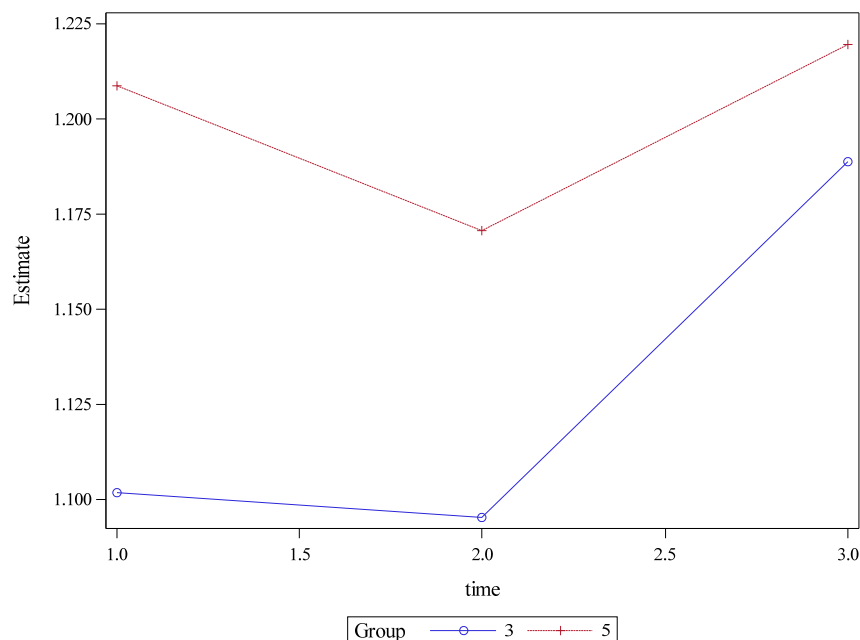
Graph 2 Ratio of words per AS-unit at the three testing times.



In the ratio of clauses per AS-unit significant age group effects ($F(1,38) = 4.90, p = .0330$) and time effects ($F(2,196) = 6.87, p = .0013$) were identified, as shown in Table 31 above. However the differences between tasks take place only from T2 to T3, with a significant increase ($t = -3.56, p = .0013$) (Table 6 in Appendix D). Regarding the differences found between the two groups, significantly more clauses per AS-unit were observed in the production of Year 5 students ($t = -2.21, p = .0330$) than in that of their younger counterparts, which mirrors the results obtained in the analyses of the number of words per AS-unit. As Graph 3 illustrates, there is a steeper decrease in the ratio of clauses per AS-unit in the production of Year 5 learners at T2, followed by a manifest increase at

T3. On the other hand, whereas the decrease observed in the Year 3 group at T2 is less pronounced, the increase in the ratio of clauses per AS-unit these learners' production undergoes is much more pronounced than the one identified in the third performance of the Year 5 group. In fact, within-subjects comparisons show that the differences between the second and the third performance of the Year 3 group, are statistically significant ($t = -3.31, p = .0139$), whereas the differences between these two tasks in the performance of the Year 5 learners do not reach statistical significance. Finally, the differences in the development of the ratio of clauses per AS/unit of these two groups from their second performance to the last time they carried out the task was statistically significant ($t = -3.13, p = .0246$). These results are detailed in Table 8 in Appendix D.

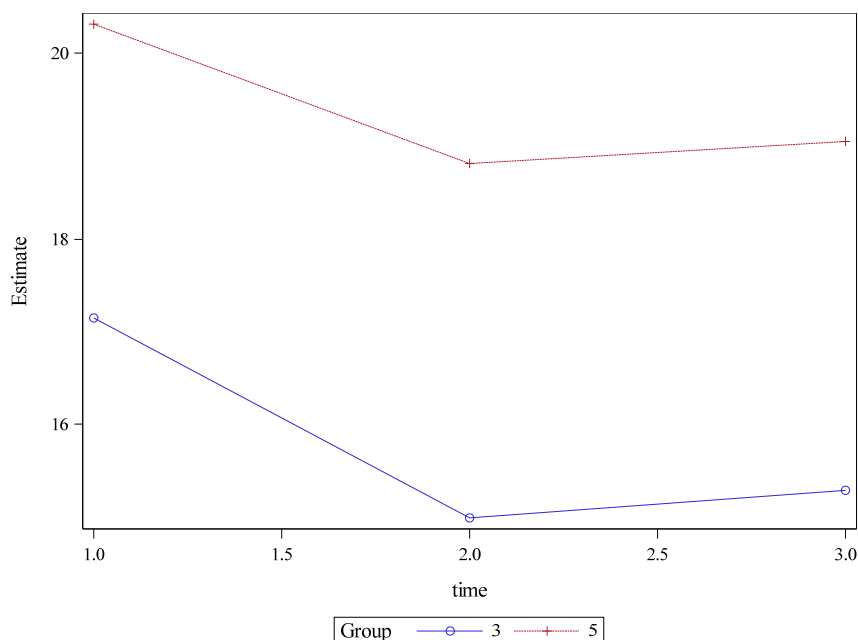
Graph 3 Ratio of clauses per AS-unit at the three testing times.



Finally, as displayed in Table 31, neither time nor group effects are statistically significant as regards lexical complexity. However, the change between T1 and T2 approaches significance in the performance of the participants in the two groups ($t = 2.30, p = .0581$) (Table 6 in Appendix D). Table 9 in Appendix D details the scores obtained in the between-subjects and within-subjects analyses run in order to account for the potential differences between these age groups and the effect of procedural TR on their oral performance. Graph 4 depicts lexical D in the learners' production. Even though lexical complexity was greater in the production of the older YLs, no significant differences can

be reported. The same applies to the differences found between tasks. Lexical D falls significantly from T1 to T2, to then recover slightly at T3, although not achieving the initial levels.

Graph 4 *Lexical D at the three testing times.*



In summary, age group effects were identified in two of the three complexity measures examined, namely the ratio of words per AS-unit and the ratio of clauses per AS-unit. Significant time differences were found only in the learners' production of clauses per AS-unit, which increased at T3. Year 5 learners' output contained significantly more words and clauses per AS-unit than that of their younger counterparts. As a final remark, it is interesting to see how the two measures of structural complexity are affected by the learners' age group, while their lexical complexity does not seem to be affected by this variable.

6.2.2.2 Impact on accuracy

This section reports the results obtained for accuracy. As explained in Chapter 5, the analysis of accuracy in the current dissertation considers the study of both global and specific accuracy. In order to investigate global accuracy, the proportion of errors per AS-unit and the percentage of EFAS over the total number of AS-units at the three data collection times have been computed. On the other hand, the measure used to examine

specific accuracy is the ratio of CV over the total number of verbs produced by the learners. Table 32 displays the preliminary analysis of the scores obtained by the learners in the two groups at each data collection time.

Table 32 *Descriptive statistics for the three accuracy measures in the three tasks.*

Group	Time	Variable	Mean	SD	Min.	Max.
Year 3	1	Errors/AS-unit	0.57	0.25	0.04	1.11
		EFAS/AS-unit	51.39	15.78	5.17	90.00
		Correct Verbs	84.71	19.48	22.22	100.00
	2	Errors/AS-unit	0.69	0.32	0.13	1.38
		EFAS/AS-unit	48.32	15.66	7.14	75.00
		Correct Verbs	87.13	23.64	0.00	100.00
	3	Errors/AS-unit	0.59	0.28	0.14	1.44
		EFAS/AS-unit	51.16	16.89	16.13	83.33
		Correct Verbs	88.28	17.19	28.57	100.00
Year 5	1	Errors/AS-unit	0.67	0.28	0.24	1.56
		EFAS/AS-unit	56.82	15.34	25.00	84.62
		Correct Verbs	81.82	18.03	26.92	100.00
	2	Errors/AS-unit	0.65	0.30	0.08	1.32
		EFAS/AS-unit	58.35	16.88	17.65	93.75
		Correct Verbs	85.99	17.02	15.00	100.00
	3	Errors/AS-unit	0.59	0.29	0.10	1.50
		EFAS/AS-unit	59.02	15.13	29.41	90.91
		Correct Verbs	83.77	15.95	28.57	100.00

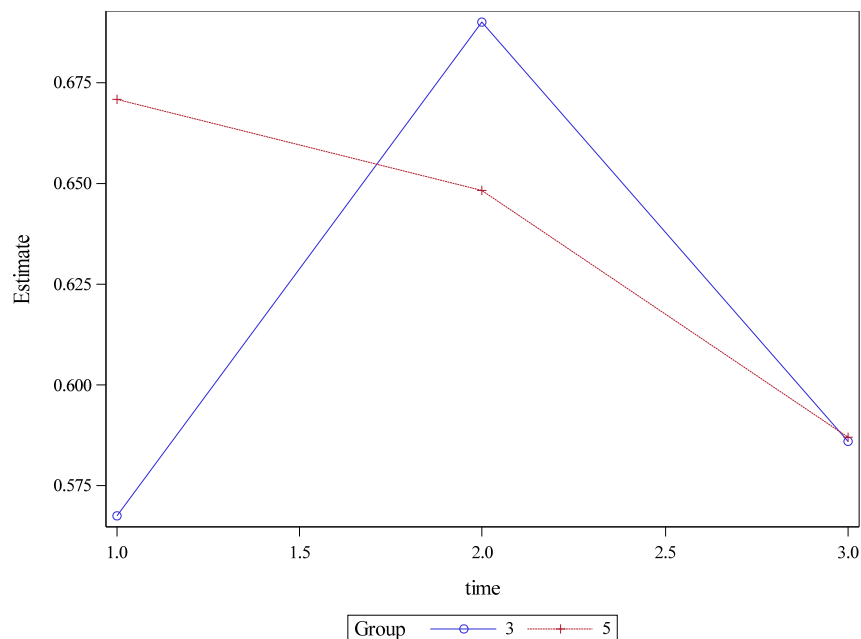
Concerning the first global accuracy measure, the mean scores of errors per AS-unit remain quite similar in the output of the two groups in the three tasks. When comparing the scores these learners obtained in each of the tasks and the differences between the two groups, the results indicate no effects of task time or age group for this accuracy measure (see significance values in Table 33 below). Thus, we can conclude that the number of errors per AS-unit remained stable across tasks and that learners in the two age groups considered produce a similar amount of non-targetlike oral output when interacting with age- and level-matched peers in the TL. The results from the comparisons between each testing time from the two groups are detailed in Appendix D, Table 10.

Table 33 *Main effects and interaction effects of task time and age group on accuracy.*

	Effect	Num DF	Den DF	F Value	Pr > F
Errors/AS-unit	Group	1	38	0.15	0.6998
	Time	2	196	2.61	0.0762
	Group*time	2	196	2.27	0.1062
EFAS/AS-unit	Group	1	38	5.18	0.0286*
	Time	2	196	0.46	0.6350
	Group*time	2	196	0.81	0.4485
CV/Total number of verbs	Group	1	38	0.42	0.5226
	Time	2	196	1.23	0.2950
	Group*time	2	196	0.44	0.6478

Graph 5 illustrates the development of the ratio of errors per AS-unit of the two groups. Although no statistically significant differences between the two groups at any of the testing times were found, in the graph it can be observed how, in spite of an initial higher ratio of errors per AS-unit in the production of the Year 5 learners, this value follows a steady decreasing trend, ending with a lower score at T3, similar to that of the Year 3 participants. In the output of the latter, however, there is a peak at T2, and then the ratio of errors per AS-unit decreases at T3. The details of the between- and within- group comparisons of the scores of this variable are presented in Table 11 in Appendix D.

Graph 5 Ratio of errors per AS-unit at the three testing times by the two groups.



Turning to our second global accuracy measure, the proportion of EFAS over the total number of AS-units, the reported mean scores at each data collection point are also similar in the production of the two groups upon TR: in each task the mean value of this aspect was over 50% (except for T2 in the Year 3 group which was only 1.68 points below 50%). Not surprisingly, the statistical comparison of the differences between the scores obtained in the three tasks shows that, although the highest percentage of EFAS per AS-unit is reached at the last task performance by the two groups, this improvement is not statistically significant for any of the groups. Even though no time effects were identified, significant age group effects were found in the learners' production of EFAS ($F(1,38) = 5.18$, $p = .0286$). These results are detailed in Table 33 above. Although the overall production of EFAS per AS-unit in the output of these two groups was found to be statistically significantly different ($t = -2.28$, $p = .0286$), revealing a higher number of EFAS per AS-unit in the production of Year 5 learners, the post-hoc tests performed indicate no significant within-subjects differences from task to task nor between the differences in the changes the oral production of each group underwent (See Table 12 in Appendix D). Graph 6 shows the differences between these two groups, exhibiting a higher percentage of EFAS per AS-unit in the output of the older YLs at the three testing times.

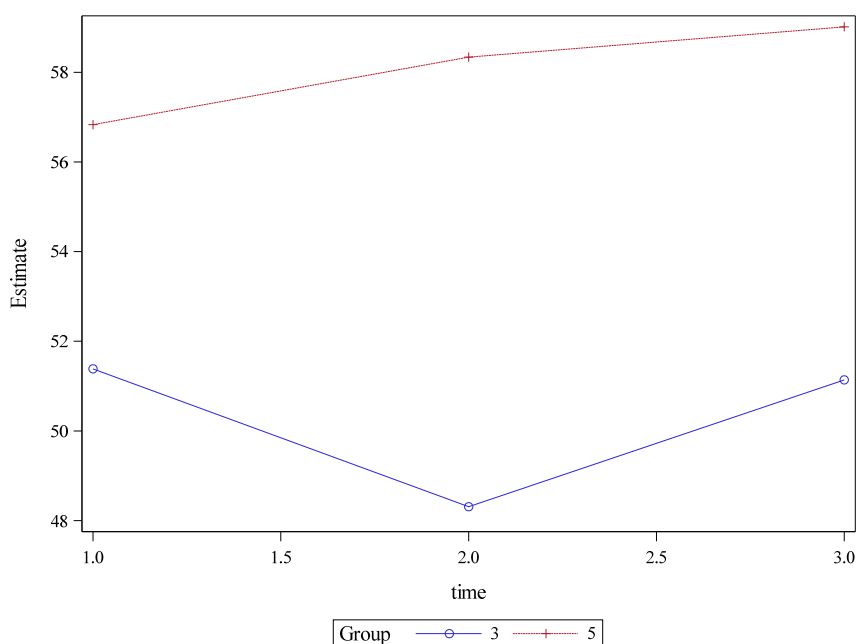
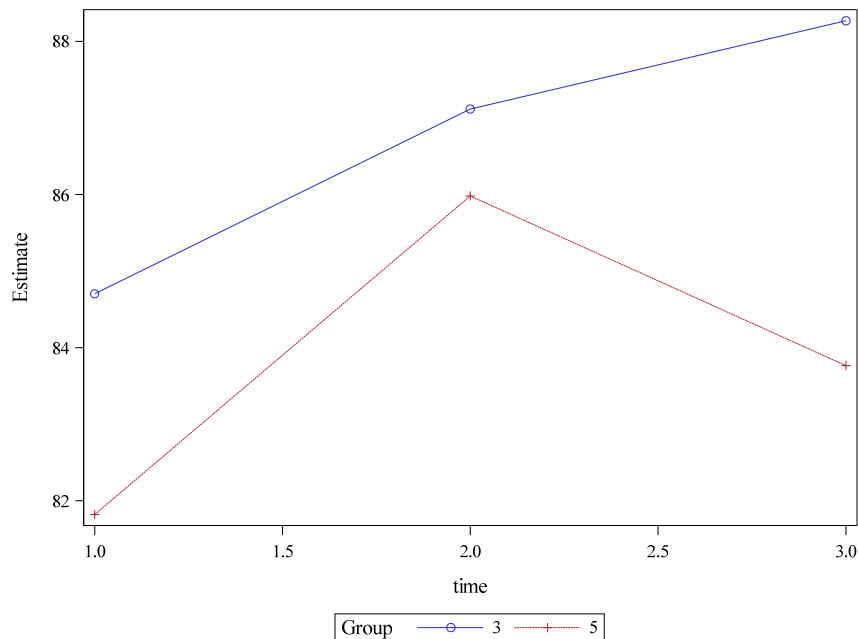
Graph 6 *Percentage of EFAS per AS-unit at the three testing times.*

Table 33 above reports the descriptive statistics of the specific accuracy measure, namely the proportion of CV. As shown in the table, in both groups the percentage of CV over the total number of verbs increases at T2, to then drop at T3. Still, the changes that the ratios of CV experience, analysed as time effects, are not statistically significant for any of the groups. Thus, we can conclude that these learners' specific accuracy does not seem to evolve significantly with the repetition of the PPT. In addition, as shown in Table 33, no group effects were identified for this variable either. The details of the analysis of this variable are presented in Table 13 in Appendix D. See below Graph 7, which represents the development of the production of CV by the participants at the three testing times.

Graph 7 Ratio of correct verbs over the total number of verbs.



The results obtained from the analysis of accuracy in the three measures examined in the current study are presented in Table 33 above. The statistical analyses reveal no significant interactions between time and group for any of the measures, nor were there significant differences between the tasks performances for any of the accuracy measures. Concerning the potential differences between groups, these were statistically significant only as regards the ratio of EFAS per AS-unit, indicating a higher percentage of EFAS in the production of the older YLs.

6.2.2.3 Impact on fluency

The approach adopted to examine the degree of speech fluency the participants unfolded in the development of the three oral tasks was based on two specific measures: speed fluency and breakdown (dis)fluency. Table 34 presents a preliminary analysis of the two fluency measures.

Table 34 Descriptive statistics for the two fluency measures in the three tasks.

Group	Time	Variable	Mean	SD	Min.	Max.
Year 3	1	Words per minute	18.96	8.46	4.05	34.16
		L1 terms /AS-unit	0.21	0.19	0.00	0.88
	2	Words per minute	22.25	7.63	9.05	41.39
		L1 terms/AS-unit	0.21	0.20	0.00	0.95
	3	Words per minute	24.74	8.02	8.14	40.69
		L1 terms/AS-unit	0.18	0.23	0.00	0.92
Year 5	1	Words/ minute	25.28	8.63	6.16	42.32
		L1 terms/AS-unit	0.15	0.12	0.00	0.44
	2	Words/ minute	28.96	8.91	10.68	57.39
		L1 terms/AS-unit	0.10	0.10	0.00	0.43
	3	Words/ minute	30.76	11.08	14.47	70.85
		L1 terms/AS-unit	0.09	0.09	0.00	0.30

Table 35 details the results obtained in the GLMM tests of these two measures. As shown in the table, group-time interaction effects were not statistically significant for the analysis of the ratio of words per minute. On the other hand, significant effects of task time and age group for this fluency measure were identified (Time: $F(2,196) = 20.18, p < .0001$; Group: $F(1,38) = 7.83, p = .0080$).

Table 35 Main effects and interaction effects of task time and age group on fluency.

	Effect	Num DF	Den DF	F Value	Pr > F
Words per minute	Group	1	38	7.83	0.0080*
	Time	2	196	20.18	<.0001*
	Group*time	2	196	0.13	0.8788
L1 terms per AS-unit	Group	1	38	5.52	0.0241*
	Time	2	196	3.32	0.0383*
	Group*time	2	196	0.67	0.5127

Post-hoc tests revealed that the effect of time was significant in the three tasks the learners performed (see Table 36). The scores obtained indicate that procedural TR led to an improvement in this measure, evident in the rise of the average number of words per minute from task to task in the production of the two groups, as Table 34 illustrates. The examination of this aspect of the participants' oral performance also revealed large individual differences within groups, as the minimum and maximum values show. These values follow the same increasing tendency as the mean scores. Specifically, the minimum value at T3 was more than double the initial measures. In addition, the analysis

of the differences between Group Year 3 and Group Year 5 reveals statistically significant values ($t = -2.80, p = .0080$), displaying a more fluent performance of the older YLs.

Table 36 *Task time effects on fluency.*

	Effect	Estimate	SD	DF	t Value	Pr > t 	Adjust- ment	Adj P
Words / minute	Time 1-2	-3.4883	0.6847	196	-5.09	<.0001	Tukey- Kramer	<.0001*
	Time 1-3	-5.6375	0.9261	196	-6.09	<.0001	Tukey- Kramer	<.0001*
	Time 2-3	-2.1492	0.7538	196	-2.85	0.0048	Tukey- Kramer	0.0133*
	Groups 3-5	-6.3514	2.2697	38	-2.80	0.0080	Tukey	0.0080*
L1 terms/ AS- unit	Time 1-2	0.02499	0.0148	196	1.69	0.0927	Tukey- Kramer	0.2118
	Time 1-3	0.04539	0.0189	196	2.41	0.0171	Tukey- Kramer	0.0448*
	Time 2-3	0.02039	0.0196	196	1.04	0.2988	Tukey- Kramer	0.5514
	Groups 3-5	0.08632	0.0367	38	2.35	0.0241	Tukey	0.0241*

The comparison of the changes in the ratio of words per minute in the two groups also revealed statistically significant increases in fluency across tasks (T1-T2: $t = -5.09, p < .0001$, T2-T3: $t = -2.85, p = .0133$; T1-T3: $t = -6.09, p < .0001$). When addressing the within-subjects differences, we found that the increase in the percentage of words per minute that took place between T1-T2 and T1-T3 in each group was also statistically significant (Year 3: T1-T2: $t = -3.40, p = .0105$, T1-T3: $t = -4.42, p = .0002$; Year 5: T1-T2: $t = -3.81, p = .0026$, T1-T3: $t = -4.19, p = .0006$). Surprisingly, when the production of the two groups at each task time was examined, the differences did not reach statistical significance at any of the three data collection times. The rest of the results are detailed in Table 14 in Appendix D. Graph 8 below illustrates the differences between the performances of the two groups, indicating how the ratio of words per minute increased upon TR and that it was higher in the output of the Year 5 learners the three times the PPT was carried out.

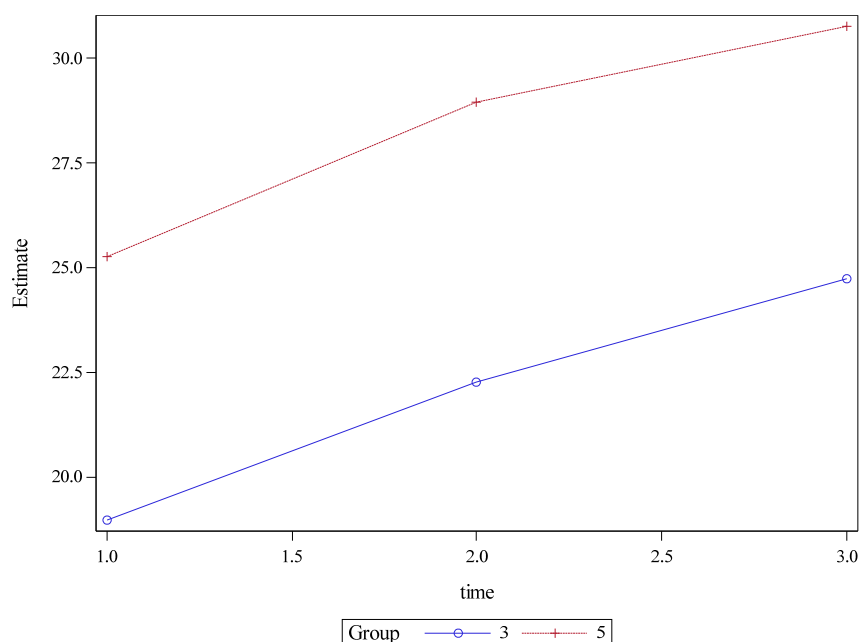
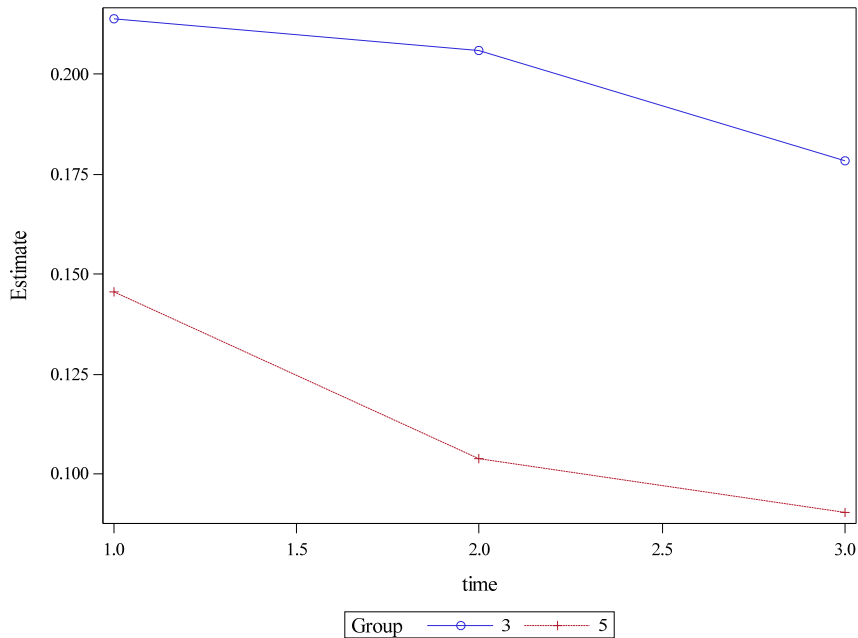
Graph 8 *Development of the ratio of words per minute.*

Table 34 above also includes the results obtained from the examination of breakdown (dis)fluency, operationalised as the amount of L1 use over the number of AS-units. As the table shows, the average number of L1 terms per AS-unit diminishes from task to task. However, in the output of the Year 3 group, the maximum values slightly increase, whereas the same values fall in the performance of Year 5 students. Even so, the differences between the minimum and maximum values are not very large as it seems these learners' reliance on the L1 is quite limited. Table 35 shows the effect of group and time for this variable. Age group effects were found to be statistically significant ($F(1,38) = 5.52, p = .0241$), and so was the effect of task time ($F(2,196) = 3.32, p = .0383$). However, as indicated in Table 36, only the differences found between T1 and T3 reached statistical significance ($t = 2.41, p = .0448$). The differences between the output of the two groups were also statistically significant ($t = 2.35, p = .0241$). Although post-hoc tests did not show statistically significant within-subjects differences, the comparison of the changes in the ratio of L1 terms per AS-unit between T1 and T3 each group experienced was found to be statistically significant ($t = 2.91, p = .0458$) (see Appendix D, Table 15 for the rest of the results). Graph 9 shows a greater L1 use by the younger learners at the three tasks performances. L1 use in both groups decreases at T2, although the fall is steeper in the output of Year 5 learners. However, it is in the comparison of the changes

in the production of the two groups from T1 to T3 that the differences between the groups reach statistical significance.

Graph 9 *Development of the ratio of L1 terms per AS-unit.*



In summary, whereas there were effects for time and age group on fluency and complexity, only age affected accuracy, and in just one of the measures analysed. Complexity, measured as clauses per AS-unit, increases upon TR. Procedural TR also had an effect on the fluency of the two groups. Speed fluency and breakdown (dis)fluency develop in opposite directions: while speed fluency rose, breakdown (dis)fluency fell at each data collection time. Concerning the effect of age group on the learners' task performance, the results obtained show that this variable affected the three dimensions of general performance analysed in this study, in fact, our results suggest that age exerts a larger influence than task time, as it affected more aspects of the learners' CAF. The only accuracy measure found to be different in the two age groups is the ratio of EFAS per AS-unit, which was significantly higher in the production of Year 5 learners. The production of the learners in the Year 5 group is more structurally complex (containing a greater amount of clauses and words per AS-unit), and more accurate (with a higher percentage of EFAS per AS-unit) than that of the Year 3 learners. As regards these learners' fluency, Year 5 learners produce more words per minute and fewer L1 terms per AS-unit than their younger counterparts. Even though Year 5 learners' oral production

was overall more fluent, when the individual performances of the two groups were compared, unexpectedly, no statistically significant differences as regards this fluency measures were reported.

6.2.2.4 CAF correlations

Following Bret Blasco (2014) and Sample and Michel (2014), correlation tests were run to determine the relationship between the different dimensions of general performance considered in the present dissertation, i.e. complexity (structural: words/AS-units and clauses/AS-units; lexical: D value), accuracy (specific: correct verb forms/total number of verbs; global: errors/AS-unit and EFAS/AS-units), and fluency (words/minute and L1/AS-unit). In order to interpret effect size we have followed Cohen's (1988) conventions, which state that a correlation coefficient between 0.10 and 0.29 reflects a small or weak relationship, one between 0.30 and 0.49 represents a moderate relationship, and a correlation coefficient of 0.50 or larger is considered to represent a large or strong correlation between variables. Tables 16 and 17 in Appendix D display the relationships identified between the CAF measures within the two groups at T1, T2 and T3. Only the correlations found between measures corresponding to different aspects of general performance (i.e. CAF) have been reported since, as it is to be expected, strong correlations have been found between the different measures of the same general performance aspect (e.g. words and clauses per AS-unit within complexity, EFAS and errors per AS-unit within accuracy).

First we will present the results obtained in the analysis of the performance of the Year 3 learners at the three testing times. As seen from Table 16 (Appendix D), at T1 there is a moderate positive linear relationship between complexity, in the three measures examined (number of words per AS-unit, number of clauses per AS-unit, and lexical complexity D), and fluency, measured as words per minute. That is, the results show that the number of words per minute and the number of words per AS-unit increase simultaneously ($r_s = 0.46$, $p = .003$). This positive relationship between complexity and speed fluency becomes stronger in relation to the number of clauses per AS-unit ($r_s = 0.52$, $p < .001$). Additionally, lexical complexity D correlates with the two fluency measures, the results showing a moderate positive relationship between the D value and the number of words per minute ($r_s = 0.36$, $p = .027$), and the D value and breakdown (dis)fluency ($r_s = 0.37$, $p = .023$). These relationships suggest that the production of a higher number

of words per minute entails the use of more complex AS-units, as well as larger and more varied vocabulary. On the other hand, trade-off effects seem to take place with the use of more varied vocabulary and breakdown (dis)fluency, measured as the number of L1 use per AS-unit. When YLs employ greater lexical variety, more L1 terms are also used. The last dimensions found to maintain a relationship are accuracy, measured in errors per AS-unit, and complexity, words per AS-unit ($r_s = 0.35, p = .027$). This result also points at the possibility of trade-off between these two aspects of CAF, as longer AS-units seem to correlate to more errors.

At T2, the moderate positive correlation between complexity and fluency continues in the relationship between words per AS-unit and words per minute ($r_s = 0.43, p = .006$), and lexical D and the ratio of words per minute ($r_s = 0.41, p = .004$). In contrast, the correlation between clauses per AS-unit and words per minute disappears, as well as the relationship between lexical D and breakdown (dis)fluency. On the other hand, more measures were found to correlate between accuracy and complexity, making this interaction more noticeable at this second testing time. A positive moderate correlation is observed between the complexity measure words per AS-unit and the accuracy measure errors per AS-unit ($r_s = 0.47, p = .002$), which was already present at T1. Clauses per AS-unit were negatively related to the number of correct verbs per AS-unit ($r_s = -0.34, p = .034$). Lexical complexity D also presented a negative correlation with the number of EFAS per AS-unit ($r_s = -0.43, p = .005$). These findings suggest some trade-off effects between complexity and accuracy since a greater number of words related to more errors, more complex AS-units contained fewer correct verbs and richer vocabulary entailed the production of fewer error-free AS-units. Finally, at T2 a new relationship appeared, namely a negative strong correlation between accuracy measured as the number EFAS per AS-unit and breakdown (dis)fluency ($r_s = -0.50, p = .001$), that is, fewer L1 terms were found in AS-units that did not contain errors. This can be interpreted as a potential negative effect of L1 use.

The results of the correlation tests between the CAF variables in the last performance of the Year 3 learners show that some of the relationships that existed in the other two tasks continue to take place, while some new ones appear. The relationship between complexity and fluency was reasonably constant at each time (T1, T2, T3). The moderate positive correlation between the number of words per AS-unit and words per minute is also present at T3 ($r_s = 0.35, p = .025$), as well as the interaction between lexical D and breakdown

(dis)fluency, which initially took place at T1 but disappeared at T2. This interaction becomes stronger at this third testing time ($r_s = 0.59, p < .001$). The strong negative correlation that appeared at T2 between EFAS per AS-unit (accuracy) and breakdown (dis)fluency continues to be present at T3 ($r_s = -0.62, p < .001$). As indicated previously, a more accurate production seems to entail a lower use of the L1. Another positive correlation was identified between fluency and accuracy: a moderate relationship between words per minute (fluency) and number of correct verbs per total number of verbs (accuracy) ($r_s = 0.35, p = .029$). This means that at the last repetition, speed fluency does negatively relate to accuracy, as both the number of words per minute and the percentage of correct verbs increase simultaneously. However, lexical complexity D displayed a moderate negative relationship with the accuracy measure errors per AS-unit ($r_s = -0.44, p = .005$). This result is significant and reveals that, at this last performance, not only do the trade-off effects identified at T1 and T2 between some of the measures of complexity and accuracy disappear, but also a new positive interaction takes place. At T3, more varied vocabulary production is related to fewer errors per AS-unit.

As shown in Table 17 in Appendix D, the comparisons between the CAF variables of the Year 5 group present a slightly different scenario. At T1, the results show moderate positive correlations between fluency, measured as words per minute, and the specific accuracy measure (correct verbs over the total number of verbs) ($r_s = 0.34, p = .034$), showing that more fluent learners produce more correct verbs. This relationship mirrors the improvement achieved at T3 by their Year 3 counterparts. Speed fluency also correlates with structural complexity in both of our measures, words per AS-unit ($r_s = 0.42, p = .007$) and clauses per AS-unit ($r_s = 0.31, p = .049$). These findings suggest that, as observed in the results obtained in the analyses of the performance of the Year 3 group, more fluent learners produce more complex AS units. Our second fluency measure, L1 use per AS-unit, displays a moderate negative correlation with global accuracy, EFAS per AS-unit ($r_s = -0.46, p = .003$). This means that a larger amount of L1 use entails fewer EFAS. This relationship was also statistically significant in the production of the Year 3 learners in their two last performances. Finally, a moderate positive relationship was found between global accuracy (errors per AS-unit) and structural complexity (words per AS-units) ($r_s = 0.39, p = .012$), that is, as seen in the analysis of the production of the Year 3 learners at T1 and T2, AS-units including more words also contain a greater number of errors, suggesting potential trade-off effects between these variables.

At T2, however, the only interaction among CAF dimensions that appears in our analyses occurs between speed fluency (words per minute) and the two measures of structural complexity, words per AS-unit ($r_s = 0.45, p = .004$) and clauses per AS-unit ($r_s = 0.45, p = .004$). A positive correlation between fluency and complexity also reached statistical significance in the second performance of the younger learners, although it appeared between the ratio words per minute and lexical D, and words per minute and the ratio words per AS-unit. No trade-off effects have been identified in the oral production of the older YLs the second time they performed the PPT.

At the last testing time, the relationship between the ratio of words per minute (fluency) and one of the measures of structural complexity, words per AS-unit, continued to be present and was stronger at this testing time ($r_s = 0.57, p < .001$). This interaction between fluency and complexity also remained constant across all the performances of the Year 3 group. Moreover, the initial negative relationship between breakdown (dis)fluency and accuracy, which disappeared at T2, reappears at T3. One of the aspects of global accuracy, EFAS per AS-unit, shows a strong negative correlation with breakdown (dis)fluency ($r_s = -0.50, p = .001$). Interestingly, a strong negative relationship was also found between these aspects at T2 and T3 in the production of the Year 3 group. As previously stated, this suggests that as the amount of L1 use increases, the number of EFAS decreases and the other way round: more targetlike production contains fewer L1 terms. Additionally, and in line with this finding, a moderate negative relationship between this fluency measure and our specific accuracy measure (correct verbs over the total number of verbs) ($r_s = -0.37, p = .019$) was also identified, which indicates that learners who make a more frequent use of the L1 also produce a smaller percentage of correct verbs.

As can be seen in our results, the only dimensions of general performance that remain stable in their relationship throughout the three testing times in the two groups are fluency (words per minute) and complexity (words per AS-unit). Moreover, in the production of the younger learners, a negative relationship between the number of errors per AS-unit and lexical variety emerged at T3. These findings suggest beneficial effects of procedural TR for these two aspects of general performance. In contrast, some trade-off effects have also been detected, mostly in the performance of Year 3 learners. When the complexity of these learners' production seemed to increase, accuracy would get worse, or the other way round. This became most evident at T2, when all our measures for these dimensions

showed unfavourable relations with each other. One instance of trade-off effect between different dimensions of complexity and accuracy was also found at T1 (words per AS-unit and errors per AS-unit). Within this group, trade-off effects between complexity and fluency also took place at T1 and T3, specifically between lexical D and breakdown (dis)fluency. Finally, the L1 was found to exert a negative influence in the younger learners' accuracy in the last two performances, which was also identified at T1 and T3 in the older learners' group. Within the production of the learners in Year 5, trade-off effects were found between accuracy and complexity at T1. These effects seem to disappear thanks to procedural TR. These findings are displayed in Table 37.

Table 37 Trade-off effects between CAF measures.

	T1	T2	T3
Year 3	Complexity – Accuracy: ➤ Words/AS-unit – Errors/AS-unit.	Complexity – Accuracy: ➤ Words/AS-unit – Errors/AS-unit. ➤ Clauses/AS-unit – Correct verbs. ➤ D – EFAS/AS-units.	
	Complexity – (Dis) Fluency: ➤ D – L1/AS-unit.		Complexity – (Dis) Fluency: ➤ D – L1/AS-unit.
		Accuracy – (Dis) Fluency: ➤ EFAS/AS-units – L1/AS-unit.	Accuracy – (Dis) Fluency: ➤ EFAS/AS-units – L1/AS-unit.
Year 5	Complexity – Accuracy: ➤ Words/AS-unit – Errors/AS-unit.		
	Accuracy – (Dis) Fluency: ➤ EFAS/AS-units – L1/AS-unit.		Accuracy – (Dis) Fluency: ➤ EFAS/AS-units – L1/AS-unit. ➤ Correct verbs/AS-units – L1/AS-unit.

6.2.3 Impact on L1 use

Regarding the general effect of age and procedural TR on the overall L1 use of these two groups of YLs, the results indicate that there is a decreasing trend in the raw number of the learners' use of their L1 as they become familiar with the PPT, as Table 38 details.

Table 38 *Descriptive statistics of L1 use across tasks.*

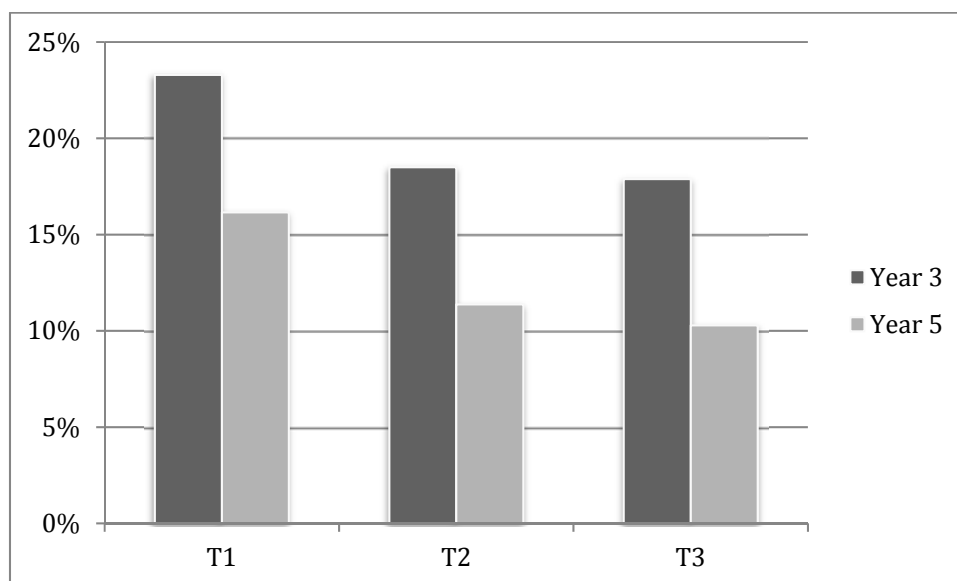
		T1	T2	T3
Year 3	L1 use	289	215	200
	%L1 use/AS-unit	23.29%	18.5%	17.88%
	Mean	20.72%	20.20%	16.97%
	Median	17.62%	16.63%	8.96%
	SD	18.55	19.43	21.98
	Min.	0.00	0.00	0.00
	Max.	87.93%	95.24%	92.31%
Year 5	L1 use	168	116	94
	%L1 use/AS-unit	16.16%	11.37%	10.28%
	Mean	14.54%	10.36%	9.02%
	Median	12.77%	8.45%	5.88%
	SD	11.88	10.21	9.37
	Min.	0.00	0.00	0.00
	Max.	43.75%	43.18%	30.43%

The part devoted to breakdown (dis)fluency in the previous section also measured the effect of procedural TR on the learners' L1 use. These children made a limited use of the L1, always below 23.29%. Specifically, the findings show that, at T1, 23.29% of the AS-units produced by Year 3 learners contained L1 terms. At T2 this percentage fell to 18.5% and at T3 it corresponded to 17.88%. In the performance of the older learners, L1 use represented a lower percentage at the three testing times: at T1, 16.16% of the AS-units had L1 terms, at T2, 11.37% and at T3, 10.28%.

As seen above, group and time effects were significant for this variable (Table 35). As shown in Table 36, the analysis of the L1 use per AS-unit revealed that the decrease observed from T1 to T3 in the learners' performance is statistically significant ($t = 2.41$, $p = .0448$), whereas the differences between T1 and T2, and between T2 and T3, did not

reach statistical significance. In addition, group effects were identified, showing a greater overall L1 use per AS-unit by the younger learners ($t = 2.35, p = .0241$). Nevertheless, even though age exerts a significant effect on the L1 use in the production of YLs, and the results show that Year 3 learners resort to the L1 more frequently in the three tasks, when the performance of the two groups at each testing time was compared, the differences were not statistically significant. Moreover, the overall L1 use by the two groups follows a similar pattern: a constant decrease is observed through the repetition of the PPT. Nevertheless, the drop from task to task did not reach statistical significance in the output of any of the groups at any data collection time. On the other hand, the comparisons of the changes that took place in the ratio of L1 use per AS-unit in the performance of each age group from T1 to T3 was statistically significant ($t = 2.91, p = .0458$), that is, procedural TR had a different effect on the use of the L1 by each of the groups. These results are detailed in Table 15 in Appendix D. Graph 10 illustrates the development of the overall L1 use by the two groups at the three testing times.

Graph 10 Overall L1 across tasks by Year 3 and Year 5.



As Graph 10 shows, even though Year 3 learners resorted to the L1 more frequently than their older counterparts, the differences between the two groups in each task was not statistically significant, nor was the drop in the use of the L1 experienced by either group.

In what follows, we will present the changes of the individual L1 functions upon TR. Table 39 displays the number of instances the participants in the two groups employed

each L1 function in each task as well as the percentages of the total L1 use each function represents.

Table 39 *Descriptive statistics of L1 functions across tasks.*

Group	Time	Function	Total	Percent.	Mean	SD	Min.	Max.
Year 3	1	Metacognitive talk	125	43.25%	3.13	4.84	0.00	21.00
		Discourse markers	26	8.99%	0.65	1.49	0.00	8.00
		Vocabulary	138	46.71%	3.45	4.70	0.00	22.00
	2	Metacognitive talk	64	29.77%	1.60	2.43	0.00	9.00
		Discourse markers	21	9.77%	0.53	1.11	0.00	5.00
		Vocabulary	130	60.46%	3.25	2.87	0.00	9.00
	3	Metacognitive talk	96	48%	2.40	4.57	0.00	20.00
		Discourse markers	19	9.5%	0.48	1.26	0.00	7.00
		Vocabulary	85	42.5%	2.13	3.05	0.00	14.00
Year 5	1	Metacognitive talk	41	24.40%	1.03	1.99	0.00	10.00
		Discourse markers	17	10.11%	0.43	0.90	0.00	3.00
		Vocabulary	110	65.47%	2.75	2.58	0.00	10.00
	2	Metacognitive talk	30	25.86%	0.75	1.85	0.00	9.00
		Discourse markers	15	12.93%	0.38	1.05	0.00	5.00
		Vocabulary	71	61.20%	1.78	2.11	0.00	10.00
	3	Metacognitive talk	27	28.72%	0.68	1.62	0.00	8.00
		Discourse markers	23	24.46%	0.58	1.41	0.00	8.00
		Vocabulary	44	46.8%	1.10	1.24	0.00	5.00

Consistent with the patterns found in most of the other aspects analysed, group and time effects were observed on two of the individual functions the L1 serves. As Table 40 shows, there were statistically significant effects of age group on one of the most frequent L1 functions, metacognitive talk ($F(1,38) = 4.91, p = .0327$), and it approached significance in vocabulary (which is the function most commonly served by the L1 in the

production of the two groups) ($F(1,38) = 4.04, p = .0516$). When looking into the differences between the groups, it becomes evident that the L1 with metacognitive function, as well as when used for vocabulary purposes, is more frequent in the production of the Year 3 group (metacognitive talk: $t = 2.22, p = .0327$; vocabulary: $t = 2.01, p = .0516^{15}$). Additionally, time was also found to have a significant effect on these two functions (metacognitive talk: $F(2, 196) = 3.59, p = .0294$; vocabulary: $F(2,196) = 6.68, p = .0014$). No statistically significant effects were identified with regard to less frequent L1 function in our data, i.e. discourse markers. Finally, no significant interactions between group and testing times were revealed for any of the L1 functions (see Table 40).

Table 40 Main effects and interaction effects of task time and age group on L1 functions.

	Effect	Num DF	Den DF	F Value	Pr > F
Metacognitive	Group	1	38	4.91	0.0327*
	Time	2	196	3.59	0.0294*
	Group*time	2	196	0.55	0.5750
Discourse markers	Group	1	38	0.19	0.6669
	Time	2	196	0.30	0.7399
	Group*time	2	196	0.48	0.6225
Vocabulary	Group	1	38	4.04	0.0516*
	Time	2	196	6.81	0.0014*
	Group*time	2	196	0.91	0.4031

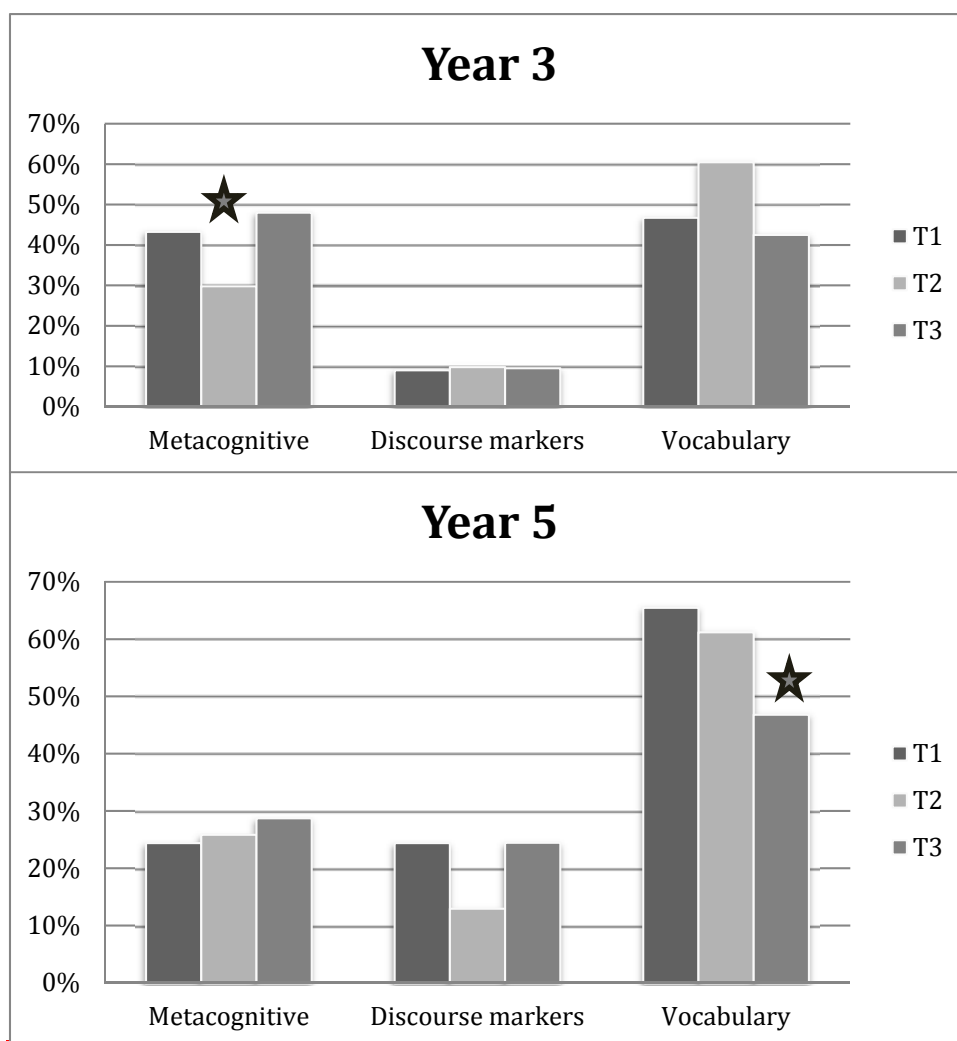
Table 41 shows the results of the analyses of the differences between L1 functions across tasks in the two groups. As illustrated in the table, the differences in the use of the metacognitive function are statistically significant only in the changes that take place between T1 and T2 ($t = 2.54, p = .0318$). Differences across task times in the vocabulary function reach statistical significance in the comparisons between T2 and T3 ($t = 2.92, p = .0109$) and T1 and T3 ($t = 3.63, p = .0011$). As expected, the differences in the use of L1 discourse markers upon TR were not statistically significant.

¹⁵ The results approached significance in the case of this variable.

Table 41 *Task time effects on L1 functions.*

	Time	Estimate	SD	DF	t Value	Pr > t 	Adj P
Metacognitive	1-2	0.4909	0.1934	196	2.54	0.0119	0.0318*
	1-3	0.3409	0.2501	196	1.36	0.1746	0.3626
	2-3	-0.1501	0.2830	196	-0.53	0.5966	0.8566
Discourse markers	1-2	0.1694	0.2566	196	0.66	0.5099	0.7868
	1-3	0.0057	0.3429	196	0.02	0.9868	0.9998
	2-3	-0.1637	0.2957	196	-0.55	0.5806	0.8448
Vocabulary	1-2	0.2488	0.1469	196	1.69	0.0920	0.2104
	1-3	0.7004	0.1932	196	3.63	0.0004	0.0011*
	2-3	0.4517	0.1547	196	2.92	0.0039	0.0109*

The post-hoc analyses of the within- and between-subjects changes indicate that the differences in the L1 vocabulary use of the two groups between T1 and T3, and T2 and T3 were statistically significant (T1-T3: $t = 3.41$, $p = .00101$, T2-T3: $t = 3.28$, $p = .0153$). Additionally, the decrease in the production of vocabulary terms by the learners in Year 5 was statistically significant between T1 and T3 ($t = 3.01$, $p = .0342$). Concerning metacognitive talk, a statistically significant drop was identified in the production of the Year 3 group from T1 to T2 ($t = 3.15$, $p = .0231$). The rest of the comparisons did not reveal statistically significant changes. As for discourse markers, neither task time nor age group were found to exert any significant effect on this L1 function. These results are detailed in Tables 18, 19 and 20 in Appendix D, and presented visually in Graph 11.

Graph 11 *Percentage of use of the L1 functions across tasks.*

Apart from the three main L1 uses presented above, L1 off-task talk was identified in the production of the Year 3 learners. These instances have not been considered in the general count because the examples come from only four children, from 3 different dyads. Table 42 below details the distribution of the use of the L1 for off-task talk in our dataset.

Table 42 *Off-task talk by Year 3 learners.*

Participant	T1	T2	T3
3A	1	1	11
3B	8	1	7
16A	1	1	0
18B	0	1	0

It is apparent from Table 42 that off-task talk came mainly from the participants in dyad 3, who produced most of the instances of this L1 function (92.62% of the total L1 off-task talk) and resorted to this L1 use at the three data collections times. There was a drop in the use of L1 for off-task talk in the second performance by this pair, followed by a rise at T3. This dyad used L1 off-task talk mainly for short expressions of how they felt during the task, which were normally one-clause long. Student 3B also asked once about research procedure, specifically whether the researcher was recording them (*¿nos estás grabando?*). On three occasions this pair engaged in a longer exchange that was not task-related, once when they first completed the task (Example 61), and twice at T3 (Example 62 and Example 63).

Example 61

1. *3B: |**tío hace tanto calor**| (*Man it is so hot!*)
2. *3A: |**yo tengo frío**| (*I'm cold*)
3. *3B: |**pues yo calor**| (*Well, I'm hot*)
4. *INV: |can we continue with the game, please?|
5. *3A: |**tengo mucho calor!**| (*I'm very hot!*)

Example 62

1. *3B: |**no aguantas así**| (*You can't be like that for long!*)
2. *3A: |**a mí me encanta ponerme rojo**| (*I love to turn red*) [holding his breath]

Example 63

1. *3B: |**ah tío me has dado un golpetazo!**| (*Man you hit me!*)
2. *3A: |**porque no tienes que coger eso**| (*Because you shouldn't pick that up*)
3. *3B: |**no estoy cogiendo nada**| (*I'm not picking anything up*)

This pair used their L1 to refer to how hot it was in the room 6 times at T1, and 3 times to talk about the time (specifically about minutes being longer than seconds), or to ask their partner to stop ‘touching them’ (*tío para de tocarme*). At T2 only two instances of L1 off-task talk were identified in this dyad’s output, and again they were playful isolated expressions such as ‘*cállate*’ (shut up!) and ‘*estoy ocupado*’ (I’m busy). T3 is the task with more off-task talk, and this time, two exchanges took place (Examples 62 and 63). The rest, as most of the off-task talk found, were again short expression that did not receive any follow-up on the part of the interlocutor. The increase in the use of this function by 3A is related to his annoyance with his partner’s behaviour, expressed in utterances like ‘*eres un plasta*’ (you are a pain in the neck), o ‘*tío déjame*’ (man leave alone!), and also in the exchanges illustrated in the examples above (62 and 63). In addition, 8 instances were just the participants asking each other, mostly 3A to 3B, to be quiet.

Learner 16B made use of this function once in the two first task performances: once because she felt attacked by her partner and wanted to defend herself, and at T2 just to express how she would like to feel as she thought the girl in the picture felt (see Example 64). Participant 18A used the L1 with this function only once during the entire experiment (at T2), and he did it to express boredom (*¡qué rollo!*).

Examples 64

Time 1

1. *16B: |in the left?|
2. *16A: | yes| sabes cuál es el left no?| (*You know where the left is, don't you?*)
3. *16B: |**sí sí que lo sé** | **no soy como xxx**| (*Yes, yes I know, I'm not like xxx*)

Time 2

1. *16A: |it is in the left|
2. *16B: |**ay qué a gusto con el cojín!**| (*Oh! How comfortable there with a cushion!*)

The examples presented above illustrate how the L1 used for off-task talk played a very limited role, reduced to unusual utterances embedded in task-related discourse, and only employed by very few learners during the development of the task. Furthermore, all the instances came from learners in the younger group. None of the YLs in the Year 5 group resorted to this function at any of the three data collection times.

6.4 Summary

In sum, our results have demonstrated that YLs in these two age groups negotiate for meaning when performing an oral communicative task with age- and level-matched peers. The participants in this study have been able to make use of all the previously identified NoM strategies: comprehension checks, self-repetitions, acknowledgements, utterance completions, clarification requests, confirmation checks, explicit corrections and recasts. Both groups displayed a very low focus on form, as shown by the scant use of strategies serving this function (explicit corrections were the least used strategy in the oral production of the two groups). Additionally, the most common functions of NoM strategies are the same in the two groups: to repair communication breakdowns and to confirm successful communication. Nevertheless, an important difference has been identified: Year 3 learners seem to mainly negotiate when a problem in communication has taken place, therefore resorting to strategies to repair communication breakdowns (most frequently to clarification requests and repair self-repetitions). On the other hand,

Year 5 learners do not only negotiate when a repair need arises, but also try to make their interlocutor aware of the absence of such communication difficulties by employing NoM strategies to confirm successful communication (mainly acknowledgements of understanding). As we will discuss, this finding suggests a change in young language learners regarding their ability to interact according to their developmental stage.

The analyses suggest that in terms of NoM, CAFs and L1 use, procedural TR affects these aspects of YLs' oral interaction in different ways. The results reveal that TR has an impact only on two functions of the strategies to negotiate for meaning (namely on those to prevent communication breakdowns and those to repair them), which decrease upon repetition. The other two functions, i.e. to confirm successful communication and to focus on form, remain stable across task performance.

As for the effect of procedural TR on YLs' CAF, we found that the repetition of the PPT used in our study affects only one of the complexity measures (namely the ratio of clauses per AS-unit) and the two fluency measures (speed fluency and breakdown (dis)fluency), which improved significantly at the last testing time. This is an important finding as it corroborates previous research showing trade-off effects between complexity and accuracy. The oral production of the participants became more fluent as well as more structurally complex upon TR, whereas their accuracy did not experience any significant changes and remained apparently stable. The lack of improvement in accuracy may be associated to the trade-off effects identified between complexity and accuracy in Year 3 at the two first testing times, and at T1 in the production of Year 5 learners. Trade-off effects were also found between lexical complexity and breakdown (dis)fluency at T1 (Year 3) and T3 (both groups). That is, the use of more varied vocabulary was related to more L1 use. L1 use also interacted with the production of EFAS by the participants in the two age groups (Year 3: T2 and T3; Year 5: Year 3: T1 and T3).

Our results provide further evidence of the limited use of the L1 by young EFL learners when engaged in an oral communicative task in the language classroom. Furthermore, age seems to play an important role as the younger learners relied on the L1 significantly more frequently than their older counterparts. The two most frequent functions affected by procedural TR were metacognitive talk and vocabulary, which follow different trends: whereas the L1 used to assist learners as they cope with unknown vocabulary significantly diminished at the last performance, the use of the L1 with a metacognitive function

seemed to increase, albeit the differences between T1 and T3 were not statistically significant. With regard to the least common L1 function, discourse markers, no significant differences were found between the production of the two age groups, and it was not significantly affected by procedural TR. Finally, off-task talk only played an anecdotal role in the oral output of these young language learners. The differences between the two age groups will be further discussed below. Table 43 provides a summary of the main findings related to the effect of TR on YLs' oral performance.

Table 43 Summary of the effect of procedural TR on YLs' oral production.

	Time1	Time 2	Time 3
Complexity	↓ Lexical D	=	↑ Clauses/AS-unit
Accuracy	=	=	=
Fluency	↑ Words/minute	↑ Words/minute	↑ Words/minute ↓ Breakdown
NoM strategies		↓ Prevent	↓ Prevent ↓ Repair
L1 use	↓ Metacognitive	↓ Vocabulary	↓ Vocabulary

In the first part of this summary, we have provided the results from analyses that collapsed the data from the two groups and examined the effect of procedural TR on the learners' production, ignoring the group variable. In what follows, we present a summarised account of our findings related to the effect of age on YLs' performance across tasks. Even though the children in the two groups employed all the NoM strategies, some differences were identified between the two age groups, specifically in the use of strategies to confirm successful communication, which were found to be significantly more frequent in the performance of the Year 5 learners. No significant group effects were identified in the analysis of the other three functions served by NoM strategies. The most frequent function in the output of the Year 3 group was repairing communication breakdowns, which was in fact the second most frequent in the production of the older YLs. The least frequent functions in the production of the two groups were preventing communication breakdowns and focus on form. Testing time was found to affect the performance of the two groups in different ways: whereas a drop in the use of strategies to prevent communication breakdowns was found in the production of Year 5 learners between T1 and T3, the decrease in the use of these strategies observed in the younger

learners' performance did not reach statistical significance. Even though the drop in the number of strategies to repair communication breakdowns was not statistically significant for either of the two groups, the comparison of the changes in the use of strategies serving this function each group experienced between T1 and T3 was indeed significant. Additionally, the comparison of the changes from T2 to T3 in the use of strategies to confirm successful communication also reached statistical significance.

The differences between the two groups are more noticeable in the results obtained from the analyses of the learners' general performance, as group effects were found in the three CAF dimensions. Specifically, Year 5 learners' language production was significantly more complex than that of their younger counterparts in two measures: words per AS-unit and clauses per AS-unit. Testing time also influenced the learners' performance: at T1 the ratio of words per AS-unit was significantly higher in the production of Year 5 learners than in that of their younger counterparts, and the comparison of the changes the performance of two groups underwent from T1 to T2 and T1 to T3 also revealed statistically significant differences. As regards the ratio of clauses per AS-unit, a significant increase was identified in the output of the Year 3 group from T2 to T3, while this aspect remained stable in the production of Year 5 learners. Furthermore, significant differences were identified in the changes between T2 and T3 when the performance of the two groups was compared.

One of the measures of accuracy was also significantly affected by age group, our results revealing more EFAS over the total number of AS-units in the oral output of the older learners. On the other hand, no effects of procedural TR were identified for any of the groups.

Regarding fluency, the older learners were found to be more fluent than their Year 3 partners in the two measures examined. The within-subjects analyses revealed significant differences in the increase of words per minute of the two groups from T1 to T2 and T1 to T3. Moreover, the comparisons of the behaviour of each group from task to task indicated that the differences in the changes this measure experienced in each of these two groups were also significant. Regarding the scores of breakdown (dis)fluency, the results indicate significantly different behaviours between T1 and T3 in each group.

Similarly, the study of the learners' L1 use also showed time and age group effects on the two groups' performances, both on general L1 use, indicating a greater reliance of the

Year 3 group on the L1, as well as on the different L1 functions (except for L1 discourse markers). Nevertheless, the ratio of L1 use per AS-unit did not significantly change in the production of either group. In contrast, statistically significant differences were found in the performance of these groups concerning the functions the L1 served: whereas L1 use for metacognitive talk diminished significantly at T2 in the production of the Year 3 group, the use of L1 vocabulary by the Year 5 learners dropped at T3. In addition, the results of the comparison of the changes in the use of this L1 function from T1 to T3 and T2 to T3 by each group also reached statistical significance. Finally, some instances of L1 off-task talk were identified in the output of the younger learners. This use was quite scarce, only used by 4 children, and it has been examined in a more qualitative manner.

The main findings regarding the effect of procedural TR on each of the aspects examined in the current dissertation in the performance of the two age groups are summarized in Table 44. Finally, Table 45 presents the significant differences between the two groups as regards the changes each of them experienced across tasks.

Table 44 *Summary of the effects of procedural TR on the performance of the two age groups.*

		T1-T2	T2-T3	T1-T3
Year 3	Complexity	=	↑ Clauses/AS-unit	=
	Accuracy	=	=	=
	Fluency	↑ Words/minute	=	↑ Words/minute
	NoM strategies	=	=	=
	L1 use	↓ Metacognitive	=	=
Year 5	Complexity	=	=	=
	Accuracy	=	=	=
	Fluency	↑ Words/minute	=	↑ Words/minute
	NoM strategies	=	↓ Prevent communication breakdowns	↓ Prevent communication breakdowns
	L1 use	=	=	↓ Vocabulary

Table 45 *Between-subjects statistically significant differences in the changes derived from procedural TR.*

	T1-T2	T2-T3	T1-T3
Complexity	Words/AS-unit	Clauses/AS-unit	Words/AS-unit
Accuracy	=	=	=
Fluency	Words/minute	Words/minute	Words/minute Breakdown
NoM strategies	Confirm	Confirm	Confirm Repair
L1 use		Vocabulary	Vocabulary

CHAPTER 7 DISCUSSION OF MAIN FINDINGS

The main aim of the current dissertation was to investigate the effect of age and procedural TR on YLs' oral task-supported interactions in an EFL setting. Specifically, we have examined their use of NoM strategies, general performance (CAF) and L1 use. In the following sections, the results from the analyses conducted are discussed. For ease of presentation, each research question will be restated and discussed in tandem with the hypotheses proposed in Chapter 5.

This chapter consists of three sections. In section 7.1, the results concerning YLs' NoM obtained from the analysis of the performance of the two groups at T1 are summarised and discussed in relation to Hypotheses 1, 2, 3 and 4. Section 7.2 presents the findings from the examination of the effects of procedural TR on the participants' oral production (NoM, CAF and L1 use) to explore Hypotheses 5, 6, 7 and 8. Finally, the influence of age on YLs' oral interactions regarding NoM, CAF and L1 use is addressed in Section 7.3 to examine Hypotheses 9, 10, 11 and 12.

7.1 Young language learners' negotiation of meaning

Research questions 1, 2 and 3 addressed the way YLs negotiate for meaning when engaged in an oral collaborative task. The first two research questions focused on YLs' ability to negotiate for meaning during peer-peer interaction and the type of NoM strategies they use, respectively. Hypotheses 1 and 2 were supported as the participants in our study negotiated for meaning employing all the strategies documented in the literature: comprehension checks and mere self-repetitions, acknowledgements and utterance completions, clarification requests and confirmation checks, explicit corrections and recasts. As anticipated in Hypothesis 3, comprehension checks, together with recasts and explicit corrections, were barely found in our database. Concurring with previous research, the most common strategies were clarification requests and repetitions (Gagné and Parks, 2013; Lázaro Ibarrola and Hidalgo, 2017a; Mackey et al., 2007; Oliver, 1998, 2002, 2009). According to Lázaro Ibarrola and Hidalgo (2017a), these strategies are employed to repair communication breakdowns. In the production of the older YLs, these strategies are as frequent as the ones used to confirm successful communication (acknowledgement of understanding).

Clarification requests are the most common NoM strategy in the output of Year 3 learners, followed by repetitions. We believe clarification requests to be the simplest strategy, as our participants mostly used ‘what?’ or even ‘eh?’, as in examples 66 and 67 below. When resorting to these strategies to repair communication breakdowns, the speaker appears not to understand, or to understand very little of what the interlocutor said, and often employs uncomplicated, and even non linguistic expressions, to clarify the meaning of the previous utterance. In fact, in some cases, the speaker seems unable, or at least reluctant, to use the TL (Example 67).

Example 66

1. *CHI2: |you have a girl with something in the sofa?|
2. *CHI1: |**what?**| [Clarification request]
3. *CHI2: |you have a girl with something in the sofa?|

Example 67

1. *CHI1: |{in the in the} the boy it is sitting in the (..) chair?|
2. *CHI2: |**eh?**| [Clarification request]
3. *CHI1: |the boy is sitting in the (..) chair?|
4. *CHI2: |of the park?|
5. *CHI1: |no|

On the other hand, confirmation checks, which were the third most frequent in the output of the Year 3 learners and the second in that of Year 5 learners, are a bit more complex. When employing this strategy, the learners show that they have not fully understood the message their interlocutor intends to communicate and are able to ask about the specific part that remains unclear (Example 68).

Example 68

1. *CHI1: |{you have a girl no} (..) you have a boy in the hall?|
2. *CHI2: |**in the hall?**| [Confirmation check]
3. *CHI1: |yes|
4. *CHI2: |yes|

It is interesting to note that younger learners resort to less complex strategies than older YLs, who have received two more years of instruction in the TL, and whose oral production already presents important differences when compared to that of their younger

peers. In what follows, the differences between younger and older YLs are further discussed.

The most striking finding of this study when compared to the existing research on YLs' NoM comes from the analysis of the effect of age on the use of NoM strategies (Research question 3). Instead of negotiating mostly to repair communication breakdowns by means of clarification requests or repetitions, older learners use strategies that indicate successful communication (particularly acknowledgments of understanding) extensively. Whilst younger learners follow the expected pattern and use NoM strategies mainly to solve communication difficulties, their older counterparts seem to be one stage ahead and make a wider use of strategies. These learners negotiate to confirm that the message is being successfully understood as much as they do to repair breakdowns. The utterances used to fulfil this function range from basic 'ok', which were the most common (see Example 69), to more elaborate sentences (as in Example 70 below).

Example 69

1. *CHI2: |in the classroom is a girl :: eating her sandwich | she {have a}
have a purple jacket +...
2. *CHI1: |**ok**| [Acknowledgment]
3. *CHI2: :: a clock in the hand +...
4. *CHI1: |**ok**| [Acknowledgement]
5. *CHI2: :: and is blonde hair|
6. *CHI1: |**ok**| [Acknowledgement]

In the example above, CHI2 is describing what she sees in her poster and CHI1 offers her support showing that she is following her partners' discourse by saying 'ok' after nearly every meaningful unit. This way, CHI1 does not need to use any strategy to check whether her partner understands the message. Thus, the low appearance of comprehension checks.

Example 70

1. *CHI1: |I have in the park a girl with {with with} brown hair | it's eating a (...)|
2. *CHI2: |a sandwich?| [Utterance completion]
3. *CHI1: |a sandwich yes | and have the t-shirt in red with a decoration of flowers|
4. *CHI2: |**yes I know what**| [Acknowledgement]

Example 70 illustrates an exchange in which CHI2 is clearly interested in her partner's production, and she is even able to complete the interlocutor's unfinished utterance by making use of the other strategy that is used to 'confirm successful communication' (turn 2). Her guess is corroborated and expanded by her partner in turn 3. Finally, in turn 4, CHI2 tells her partner that she knows what he means, implying that it is no longer necessary to continue working on this item and that they can move on to a new one. After this utterance, CHI1 does not need to check his partner's comprehension.

Oliver (2009) hinted at this phenomenon when she observed that younger and older YLs use strategies concerned with 'self' (e.g. clarification requests, confirmation checks) similarly, but reported differences regarding the use of strategies related to 'other' (comprehension checks and other repetitions), which were slightly more frequent in the oral production of the older YLs. In our study, Oliver's (2009) 'other' strategies are still rare in either groups' output, probably because of the different contexts in which the studies were carried out (ESL vs. EFL) or the proficiency level of the participants. Nonetheless, it is evident in our data that older YLs are indeed concerned about their interlocutors' understanding and resort to strategies that let others know that no communication problem has occurred.

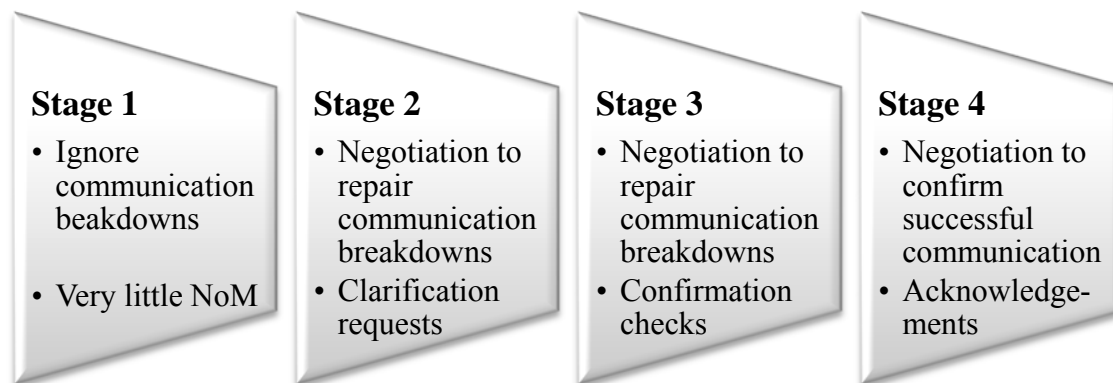
As reported in previous studies (Oliver, 1998, et passim), in younger learners' interactions strategies to repair communication breakdowns prevail, those to confirm successful communication and to prevent communication problems are not very frequent, and focus on form is basically nonexistent. What is revealing of age differences is the fact that the frequency of use of these functions is slightly different in the output of our older YLs, and two functions equally frequent are detected: NoM to repair communication breakdowns and to confirm successful communication. Finally, focus on form is also scant. As for the rates of use of each function, only the frequency of strategies to confirm

successful communication is significantly higher during the performance of the older group. This last finding lends support to previous research pointing to the changes that take place during childhood, and the differences that may be found in the behaviour of children that share the same developmental stage (i.e. middle childhood, ages 7 to 11). Concurring with Pinter's (2007) suggestion, the 10-11 year-olds in our study seem to be more capable of providing support to their interlocutors than our younger participants (8-9 years old), who were more focused on their own understanding of their partners' message. Still, younger learners show some concern about their interlocutors' needs as well, and resorted to strategies to confirm successful communication. Since YLs' (in both groups) attempts to cater for their interlocutors' needs are expressed in the form of strategies that have not been considered in previous research, we believe that YLs' efforts may have gone unnoticed so far.

These findings lead us to propose the first four stages of the acquisition route of the NoM strategies by these children. Just as in other areas of language acquisition, learners seem to go through different stages as they start to negotiate for meaning and use different negotiation strategies. The first stage would be when no negotiation, or very little, takes place. At this stage, when a communication breakdown occurs, it is ignored by the speakers and the conversation continues. Very low proficiency learners would be at this stage and display very low rates of NoM. As Lázaro Ibarrola and Azpilicueta Martínez (2015) proposed, 'a threshold level might be necessary for interaction to occur' (p. 17). The negotiation that took place between our Year 3 participants showcases the second stage: these learners negotiate mainly when there is a communication problem and they need to solve it. In order to do this, learners at this stage mostly use clarification requests, normally by employing uncomplicated formulas. Learners at Stage Three and those at Stage Two share their main reason to negotiate, that is, to repair communication breakdowns. At this third stage, however, learners begin to employ slightly more complicated strategies, i.e. confirmation checks. In the production of the Year 5 participants, this strategy is in fact the second most frequent. Nevertheless, these older YLs would be at Stage Four, when learners do not only negotiate to solve communication problems, but they anticipate potential difficulties and try to prevent them by confirming successful communication (see Figure 7). As learners mature and their command of the TL increases, their negotiation skills improve, as shown in studies with YLs in SL

contexts and with adult learners. However, the possible further stages lie outside the scope of the present dissertation.

Figure 7 *Stages of the acquisition route of NoM strategies.*



7.2 Effect of procedural task repetition on young learners' oral production

Research question 4 was divided into three subquestions, each addressing the impact of procedural TR on a different aspect of the participants' oral production, namely NoM, CAF and L1 use. This structure will be followed in our discussion. Thus, each subquestion will be answered individually in relation to Hypotheses 5 (NoM), 6 (CAF), 7 and 8 (L1 use). The first aspect to be discussed is the impact of procedural TR on YLs' NoM. Then we present our insights into the changes in the three CAF dimensions and, finally, the way L1 use unfolded is commented upon.

7.2.1 Procedural task repetition and negotiation of meaning

Hypothesis 5 was supported: after the repetition of the same task type, with slightly different content but the same procedure, the frequency of two of the functions served by NoM strategies dropped, whereas the other two functions identified remained stable upon TR. Strategies to repair communication breakdowns (i.e. clarification requests and confirmation checks), and to prevent them (i.e. acknowledgements and utterance completions) were significantly influenced by testing time. Strategies to confirm successful communication and to focus on form on the other hand, did not change significantly. Our results reflect those put forward in previous research. As Mackey et al. (2007) and Lázaro Ibarrola and Hidalgo (2017b) found, procedurally familiar tasks trigger fewer CAs since, when learners are familiar with the task procedure, the need to solve communication difficulties diminishes. Regarding the findings in García Mayo and Imaz

Agirre (2016), who reported no changes for any of the NoM strategies, we also found that two of the functions served by the negotiation strategies did not significantly vary from task to task (strategies to confirm successful communication and to focus on form). However, these studies are not fully comparable as their participants only repeated the task twice and with a two-month interval between tasks while the learners in the current study (and in Lázaro Ibarrola and Hidalgo (2017b) and Mackey et al. (2007)) repeated it three times on a weekly basis.

It is also worth mentioning that even though the increase in the use of strategies to confirm successful communication did not reach statistical significance, the raw numbers show that the frequency of NoM strategies serving this function increases across tasks, at the same time strategies to repair communication breakdowns decrease. These results complement the findings reported by Mackey et al. (2007) and Lázaro Ibarrola and Hidalgo (2017b), who also found fewer clarification requests in procedurally familiar tasks, and offer a new perspective from which to understand the drop in the number of CAs: it may not only be that learners' need to solve communication difficulties decreases because they are familiar with the task, but also because they are relying on a different set of strategies which have not been accounted for in the traditional classification as proposed by Oliver (1998, *et passim*): utterance completions and acknowledgements of understanding (see examples 64 and 65 above). Thus, we claim that a more comprehensive inventory of communication strategies should be brought to the field of interaction in order to be able to provide a more complete account of what is going on when learners interact with one another.

7.2.2 Procedural task repetition and CAF

Based on the general claim that TR leads to a more efficient organisation of linguistic resources (Bygate and Samuda, 2005; Kim and Tracey-Ventura, 2013; Lynch and Maclean 2000, 2001; Pinter, 2007) and, consequently, to benefits for general competence (Ahmadian and Tavakoli, 2010; García Mayo et al., *in press*; Kim, 2013), Hypothesis 6 predicted improvements in YLs' overall performance, and fluency was the CAF feature expected to be most affected by procedural TR.

In this regard, the first claim in this study is that procedural TR has a positive effect on fluency and structural complexity. Concerning fluency, the two aspects measured in our

study were benefited: the number of words per minute steadily increased whilst the number of L1 terms per AS-unit dropped at T3. As for structural complexity, the benefits of procedural TR are more modest, as only the ratio of clauses per AS-unit increased at T3, and only in the performance of one of the groups (as discussed in section 7.3.2 below). The two other complexity features analysed (words per AS-unit and lexical variety) were not significantly affected by this type of TR, nor was any of the measurements used to examine accuracy.

Thus, Hypothesis 6 is supported, and can be expanded: the benefits of procedural TR on YLs' general performance are confirmed, above all improvements in fluency (as predicted) and also complexity. Thus, our findings corroborate previous studies with YLs reporting improvements in fluency and complexity after TR. Just as Bret Blasco (2014) found, our participants' production was more fluent and complex, whereas accuracy was not significantly affected. The rest of the studies have mainly reported benefits for YLs' overall performance and significant increases of fluency upon TR (García Mayo et al., in press; Pinter, 2007; Sample and Michel, 2014). To the best of our knowledge, the most recent study which has examined CAF with young EFL learners is Lázaro Ibarrola and Hidalgo (2017b) and, although the improvement in YLs' fluency was not statistically significant, some gains were reported. Studies with adult learners have also concluded that TR leads to greater fluency and complexity in the learners' oral production (e.g. Bygate (2001) and Bygate and Samuda (2005) in ESL settings; and Ahmadian and Tavakoli (2010) and Saeedi and Rahimi Kazerooni (2014) in EFL settings). Altogether, our study provides further evidence of TR as a task feature that brings about beneficial effects for the fluency and structural complexity of young EFL learners.

As regards the underlying relationships between the different dimensions of general performance through the repetition of a task following the same procedure, the only relation that remains present in the output of the two groups at the three testing times is the positive interaction between fluency, measured as words per minute, and structural complexity (words per AS-unit). Hence, improvements in complexity and fluency seem to go hand in hand. This suggests that longer AS-units appear in a more fluent oral production in the TL. Our findings mirror Bret Blasco's (2014), as she also found a simultaneous increase of these two CAF dimensions. This researcher employed fluency measures similar to ours (specifically, words per minute and ratio of L1 words per unit). On the other hand, this finding contrasts to Sample and Michel (2014), who reported

trade-off effects between these two dimensions. However, their analyses correspond to different measures to examine this aspect: in the present study we have not considered the pauses in our participants' speech, but rather the number of words they produced per minute, which visibly increased upon TR. Our conclusion of simultaneous improvement in these two dimensions would thus seem to be justified.

In contrast, some trade-off effects have been detected, which have also been reported on in previous research (e.g. Bret Blasco, 2013; Sample and Michel, 2014). Our data reveal negative interactions among complexity and accuracy: in the first tasks, learners who produced longer and more complex AS-units also made more errors. During the first performances, learners tend to prioritise meaning over form, thus concentrating on getting the message across and keeping the conversation going (Sample and Michel, 2014; Skehan and Foster, 2001). It is most unlikely that learners, when facing a task for the first time, concentrate on both complexity and accuracy simultaneously, above all learners with a low level of proficiency in the TL. This explains why an increase in complexity may bring about more non-targetlike production and, therefore, a decrease in accuracy. Nevertheless, these relationships clearly diminish and disappear upon TR. Even though no significant improvement in accuracy seem to take place, thanks to the familiarity learners achieve through procedural TR, attentional resources can be devoted to the different aspects of general performance but not at the expense of any of them. Providing further support to Skehan and Foster (2012)'s Extended Trade-off Hypothesis, our results reveal that task conditions, specifically procedural TR, contribute to a better distribution of the learners' attentional resources and attenuate the trade-off impact. This finding fits well with previous studies dealing with the relationships among CAF dimensions (e.g. Bret Blasco, 2014; Sample and Michel, 2014). Other trade-off effects were also found, most of them in the performance of the younger learners, and all of them related to L1 use. The different ways in which the interactions among CAF dimensions are affected by age will be discussed in section 7.3.2 below.

7.2.3 Procedural task repetition and L1 use

The ratio of L1 terms per AS-unit decreased at T3. Consequently, Hypothesis 7, predicting fewer L1 terms in the last performance, is confirmed. Our findings concur with those reported by Azkarai and García Mayo (2016), who even after only two performances obtained the same results. Nevertheless, it is surprising that our data do not

reveal a significant decrease between T1 and T2, or T2 and T3; only the drop in the ratio of L1 terms per AS-unit between T1 and T3 is statistically significant. This might be due to the different time interval between tasks in the two studies: in Azkarai and García Mayo (2016), three months passed between each data collection point. This finding suggests new lines of research: Are the results obtained in studies with longer time intervals between repetitions influenced by the learning in between testing times or by the repetition of the tasks? Is it the combination of time span and number of repetitions that enhances the effect of TR? We can only hypothesise that for repetition to show its impact on L1 use, more than two performances are needed when the tasks are administered within a one-week interval.

In any case, the lower L1 use identified was not unexpected since familiarity, achieved through TR, provides learners with more attentional resources to devote to language form. Thus, the need to resort to the L1 to communicate is likely to decrease. Consider example 71 below. The first exchange has been taken from this pair's second task performance, whereas the second one is from their last performance. The two examples illustrate a similar exchange, very frequent in our data: one of the participants asks where one of the pictures is placed, and the other one answers. At T2, CHI2 still relies on the L1 to say that the picture is not on the poster, but a week later, the same learner is able to answer using the TL.

Example 71

Time 2

1. *CHI1: |{where is the boy that have a} (..) where is the girl that :: have a glasses {and a} :: and the hair black?|
2. *CHI2: |**yo lo tengo fuera**| (*I have it outside [the poster]*)

Time 3

1. *CHI1: |where is the boy that :: is reading a book?|
2. *CHI2: |I don't have it in the picture|

Thanks to procedural TR, learners have the opportunity to face similar communicative situations more than once. On these occasions, learners can experiment with their output, produce different versions of the message they want to communicate, modify their initial production and, eventually, make it more targetlike.

We have also examined how procedural TR influences the functions the L1 serves. In light of our results, we can state that Hypothesis 8 is not fully confirmed: even though, as predicted, the vocabulary function is the most common at the three data collection times, its frequency decreased significantly at T3. Our findings partly support Azkarai and García Mayo's (2016) claim that the use of the L1 to cover for unknown TL vocabulary is the most frequent L1 function in the oral production of young EFL learners. However, instead of no changes in the use of this L1 function across tasks, we found a significant drop after repetition. Thanks to procedural TR, our participants seem to have gained a better access to the language needed to complete the task. Since they mainly had to describe children and places, by T3 they had probably encountered most of the necessary TL vocabulary in the previous tasks, or, as in example 72 below, they were able to ask for unknown vocabulary (appeal for help) in the TL.

Example 72

Time 2

1. *CHI2: |{where is the boy that :: {have a have a} (.)| ¿cómo se decía
|{cortina} **no esto** (..) {a (..)} **que tiene un disfraz de dalmata?** |
(how did we say {curtain} no I mean (..) who is wearing a
Dalmatian costume?)
2. *CHI1: |next to the door eh <on the> [//] in the second plant|

Time 3

1. *CHI1: |where is the girl that :: is eating a sandwich?|
2. *CHI2: |{the girl is in the} (.)| how do you say sentada en la cama?|
3. *CHI1: |{in the in the} (..) on the bed|

The example above illustrates the development the oral production our young language learners' experienced. At T2, CHI2 still uses the L1 to appeal for help, one of the functions included under the 'vocabulary' category. The same learner, a week later, is able to use the TL for the same purpose. Additionally, in turn 3 CHI1 gives her interlocutor the expression he was looking for, something that did not happen at T2, when the vocabulary request was ignored and information about the game was offered instead.

The second most frequent L1 function is metacognitive talk and it also changed across tasks, decreasing significantly at T2 to then slightly increase at T3. The differences between T1 and T3 were, however, non significant. The decrease at T2 came as no

surprise, as the learners should already be familiar with the task procedure. The subsequent increase was unexpected. If in the second performance the participants did not need to rely on the L1 as frequently as the first time they faced the game, theoretically, the need would become even smaller at T3. The new tendency of the frequency of this function may be due to a drop in the learners' motivation to complete the task, as other authors have already suggested (García Mayo and Hidalgo, 2017; García Mayo and Lázaro Ibarrola, 2015). After performing the same task twice, the learners may have not been as motivated as at T1 and T2 to give their best. Consequently, they shifted their attention to meaning and the language they needed to complete the task, which may be the reason why L1 vocabulary decreased. The metacognitive function, as well as L1 discourse markers, may be considered as language not belonging to the task discourse, hence likely to be pushed aside when concentrating on a difficult task (T1) or when the motivation to complete a task drops (T3 possibly). Nevertheless, since no motivation questionnaires were administered, this hypothesis needs to be considered with caution: we can only speculate on the basis of our results that a change in the learners' attitude towards the task took place at T3. Nonetheless, instances of L1 metacognitive talk turning into TL expressions have been identified in our learners' performance, as in example 73.

Example 73

Time 1

1. *CHI2: |ok | **te toca** Armando| (*It is your turn Armando.*)

Time 3

1. *CHI2: |you start|

As for discourse markers, which we had hypothesised would decrease, these were not significantly affected by procedural TR. Discourse markers are the least frequent L1 function at the three data collection times and, even though the raw numbers slightly decrease at T3, the changes do not reach statistical significance. This result may be also due to the low overall number of instances of this L1 use present in our dataset, as L1 discourse markers only represent 11.27% of the total L1 use by the two groups at the three data collection points. This finding supports previous research that has shown that this aspect remains constant at even higher levels of proficiency (see Lázaro Ibarrola (2016)). In relation with the line of thought regarding metacognitive talk, L1 discourse markers may not be considered part of the 'task-related discourse', and remain overlooked by the young participants. Moreover, this language use is mainly linked to expressions of

excitement during task completion (Example 74) or used when adding new elements to a previous utterance (Example 75), and may be internalised in the learners' identity and L1 discourse. Thus, the second part of Hypothesis 8 was not confirmed: as discussed, discourse markers remained stable upon TR.

Example 74

Time 1

1. *CHI2: |it is a girl in a school?|
2. *CHI1: |yes|
3. *CHI1: |¡bien!| (*great!*)

Example 75

Time 3

1. *CHI2: |{ehm} the girl that she has a t-shirt purple :: that is eating a snack :: is in the left of the sofa?|
2. *CHI1: |{yes} ah no!|
3. *CHI2: |o sea in the sofa :: in the bed| (*I mean*)

To conclude, the use of the L1 is scant and it decreases upon procedural TR, as observed in previous studies (Azkarai and García Mayo, 2016; Bret Blasco, 2014). Additionally, our participants mainly use the L1 to facilitate task completion: to cover for unknown vocabulary and for metacognitive talk. Other L1 functions, such as discourse markers appear to play a minor role. These positive findings encourage the use of oral collaborative tasks in the language classroom.

7.3 Effect of age on young learners' oral production

The effect of age on YLs' oral performance has been addressed in Research question 5, which consists of three subquestions, each targeting the impact of age on one distinct aspect of YLs' oral performance: Section 7.3.1 deals with the answer to Research question 5 in relation to Hypothesis 9, section 7.3.2 discusses the question as regards CAF and Hypotheses 10 and 11, and finally, section 7.3.3 considers the answer to the second part of Research question 6, YLs' L1 use in regard to the hypothesis posed (Hypothesis 12).

7.3.1 Age and negotiation of meaning

The two groups of young participants in our study used strategies to negotiate for meaning in a similar way. Only the strategies used to confirm successful communication (i.e. acknowledgements of understanding and utterance completions) were significantly affected by the learners' age. In this case, it was the older YLs the ones who made a more extensive use of these strategies at the three testing times. Thus, Hypothesis 9 is not supported as our finding clashes with García Mayo and Lázaro Ibarrola (2015) and García Mayo and Imaz Agirre (2016) who reported a greater use of NoM strategies by the younger learners. This may be due to the different set of strategies examined in these studies. Oliver's (1998) classification taken in isolation may divert the attention from other equally important strategies which have, therefore, not been considered in previous studies. García Mayo and Lázaro Ibarrola (2015) and García Mayo and Imaz Agirre (2016) did not consider the use of strategies that confirm successful communication (i.e. utterance completions and acknowledgements), which were in fact the most frequent in the production of our older YLs. In this regard, we can only hypothesise that YLs in the latter age range are in a different stage of acquisition of the NoM strategies in which learners become more prone to offering support to their interlocutor during oral interaction, as already indicated by Pinter (2007). As reasoned in section 7.1 above, the most basic function NoM strategies serve seems to be repairing communication difficulties, and that is the main reason why younger learners negotiate. It appears that as they become older, they start to be aware of the need to assist communication and, in order to do so, older YLs resort to strategies that let their interlocutor know that the message is being successfully communicated (as examples 64 and 65 illustrate).

When looking at the impact of procedural TR on each of the age groups separately, only the drop at T3 in the use of strategies to prevent communication difficulties was statistically significant and only in the production of the Year 5 group. It is interesting how strategies serving this function, which is one of the least frequent, only followed by focus on form, are in fact the only ones that significantly diminish upon TR. Older learners may be aware of their interlocutors' better understanding of the PPT procedure by the last performance, and therefore, of the gratuitous nature of the use of strategies to prevent problems that were unlikely to happen. This is reinforced by the fact that these learners make an extensive use of strategies that confirm that the message has been successfully communicated.

As seen in section 7.2.1, the other function that was significantly affected by procedural TR when the performance of all the participants was analysed as a whole, without considering the two individual groups, were strategies to repair communication breakdowns. The comparison of the output of the two age groups revealed that the changes each group experienced between T1 and T3 were statistically significant, that is, even though the decrease in the number of strategies serving this function did not reach significance in the performance of the groups independently, it became evident that procedural TR affected each group differently. Therefore, the second part of Hypothesis 9 is not supported either, as some differences have been found regarding the effect of repetition on each age group.

These findings highlight the fact that childhood is a period of many changes and that children in the same developmental stage, middle-childhood in the case of our participants (Berk, 2006), behave and interact differently. Thus, their language learning process is affected in different ways by different task conditions.

7.3.2 Age and CAF

Age-related differences in the performance of the two groups of participants in the current study were more pronounced than the differences produced by the effect of procedural repetition. Whereas only two CAF dimensions (complexity and fluency) were affected by procedural TR, differences between the two age groups were identified in relation to all three CAF aspects.

The analyses revealed significant differences between the two groups for two of the complexity measures: the oral production of the older YLs contained significantly more words and clauses per AS/unit than that of their younger counterparts. Even though the output of the Year 5 learners was more structurally complex, no significant differences were found in regard to lexical complexity. This finding comes as a surprise, since we would have expected the older YLs' slightly better command of the TL to be displayed in the three measures of complexity. The fact that the oral output of the two groups presented a similar lexical variety may be attributed to the PPT used in our study, as it may not allow for richer vocabulary. After all, the number of items in the pictures was quite similar in the three tasks, as was the type of language the learners' used to complete the each task. Consider Example 76:

*Example 76**Time 1*

1. *CHI2: |do you have the girl :: eating a sandwich with a purple jacket?|
2. *CHI1: |yes where?|
3. *CHI1: |in the class?|
4. *CHI2: |where|

Time 2

1. *CHI1: |do you have a girl with glasses?|
2. *CHI2: |with glasses? | no in the|
3. *CHI1: |in the poster?|
4. *CHI2: |in the poster no|

Time 3

1. *CHI2: |ok| {and do you have (...)} eh do you have eh the girl with the glasses?|
2. *CHI1: |where {where} I have to put it?|
3. *CHI2: |you have to put it in the bedroom in the left side of the girl|
4. *CHI1: |ok the left {side} side|

The extracts of conversation in this example come from the output of the same pair at the three data collection times. As can be observed, the questions and answers employed are very similar, as is the vocabulary used. In the three tasks, the learners had to choose among four pictures of children, and to do that they had to describe, or ask about, the clothes they were wearing, or any other characteristic of the photos in the game. Then, they had to give directions to their partner about where to place the pictures on the posters. Thus, as our results show, the vocabulary needed is rather limited. However, the exchange from T3 contains longer utterances than in the instances from the previous two performances, illustrating how structural complexity was affected even though the lexical variety was quite similar.

As discussed above, procedural TR exerted a moderate effect on complexity, as only the ratio of clauses per AS-unit significantly increased through repetition, and only in the performance of the younger learners, whereas that of the Year 5 learners seemed to remain stable. This finding offers further evidence for the different stages within childhood, and how different procedures may affect different age YLs in different ways. According to

our results, procedural TR does not significantly assist complexity in the performance of the older children, whereas the benefits are statistically significant in the output of the younger learners. Taking into consideration the already more complex language identified in the performance of the Year 5 group, absence of significant improvement in the production of this group after TR could be attributed to the possibility that the level of complexity they employed at T1 was enough to successfully complete the task. In contrast, Year 3 learners took advantage of the benefits of procedural TR and were able to produce more complex language after only two task performances.

The oral production of the Year 5 learners was significantly more accurate, as regards the ratio of EFAS per AS-unit, than that of their younger counterparts. No differences between the two groups were found in the analyses of the other two accuracy measures. Our results partially corroborate those obtained by García Mayo et al. (in press), as our older learners, even though a year older than theirs, produced more EFAS per AS-unit than the younger participants. Nevertheless, as already seen, no task time effects were identified in our data, as opposed to their findings. Again these results come as no surprise since the older learners' command of the TL was slightly higher than that of the Year 3 learners. Furthermore, they had had two more years of instruction in the TL, and their developmental level may also have been slightly different.

As regards fluency, our findings differ to some extent from those in García Mayo et al. (in press). According to our results, procedural TR assists fluency not only in younger learners' oral output, but also their fellow learners' fluency improved upon TR. Additionally, the oral production of the two age groups developed in different ways from task to task: even though the two groups' fluency increases significantly at each testing time, Year 5 learners were overall more fluent. Regarding the second fluency measure, breakdown (dis)fluency, a parallel pattern is observed: the output of the Year 5 group contained significantly fewer L1 terms per AS-unit. However, when the development of this measure from task to task is examined, no significant changes are observed in either of the two age groups. Nevertheless, the comparison of the changes between T1 and T3 of Year 3 and Year 5 are indeed significant. Taking our discussion back to the fact that Year 5 learners have received more instruction in the TL, the need to rely on the L1 to communicate a message is lower.

Altogether, our results strengthen Hypothesis 10, which predicted statistically significant differences in favour of the Year 5 learners as regards general performance (CAF). The older YLs' (age 10-11) oral production was significantly more complex, accurate and fluent. Nevertheless, these results need to be interpreted with care as not all of the measures of the CAF dimensions examined in this study reached statistical significance. As mentioned previously, lexical complexity was not significantly affected in the performance of the older group. Concerning accuracy, only the ratio of EFAS per AS-unit was significantly greater in the output of this group.

Hypothesis 11, which anticipated procedural TR would improve Year 5 learners' accuracy and Year 3 participants' fluency was not confirmed: our analyses reveal a significant increase of structural complexity in the performance of the younger group (clauses per AS-unit) whilst accuracy remained stable in the two groups. Fluency (words per minute) improved in the two groups.

As already discussed, a positive relationship takes place between fluency and complexity in the performance of the two age groups, which is maintained through TR. According to these results, YLs who produce longer and more complex sentences are also more fluent in the TL. On the other hand, trade-off effects have also been identified in our data. In the performance of the older YLs, trade-off effects have been found only between complexity and accuracy and only at T1. In the output of the younger learners, apart from the negative relationship between complexity and accuracy, some more negative interactions were identified between accuracy and fluency and between complexity and fluency.

As in Sample and Michel (2014), complexity-accuracy trade-off effects diminished upon TR. Nevertheless, some differences were found in the production of our two age groups: whereas trade-off effects between these two dimensions disappear in the performance of the Year 5 group upon TR already in these learners' second performance, the negative interactions among some of the measures of these CAF dimensions continue to be present in the output of Year 3 also at T2. Still, the negative interaction between complexity and accuracy disappear completely at T3, as reported by Sample and Michel (2014). The first time YLs face the PPT, improvements in one of the aspects of general performance seem to come at the expense of others. Thanks to task familiarity, achieved through procedural TR, these negative effects disappear by the last time they carry out the task, and learners are able to focus on all the aspects of general performance at the same time. Hence, by

the last performance, tasks demands appeared to no longer exceed the learners' linguistic resources, at least as regards some of the CAF dimensions.

In the younger learners' production, a negative relationship was also identified between lexical complexity and breakdown (dis)fluency, which disappeared at T2 but reappeared surprisingly at T3. When younger learners tried to use richer vocabulary, they had to rely on their L1 more frequently. This finding can be related to the trade-off effects between lexical complexity and accuracy. Attempts to employ a more varied vocabulary bring about either more errors or more L1 terms per AS-unit. Also in relation to L1 use, interaction effects were found between accuracy (EFAS per AS-unit) and breakdown (dis)fluency (L1 terms per AS-unit) at T2 and T3. This relationship was also present in the production of the Year 5 group, although only at T1 and T3. These results reveal some negative effects of the use of the L1 during L2 interaction as it seems to bring about fewer EFAS and less rich TL vocabulary. Our findings add a new perspective to research on the impact of L1 use for FL learning which, as we will discuss below, also provides benefits to language learning.

In light of these findings, we can claim that CAF developed in different ways in the performance of these two age groups (ages 8-9 and 10-11).

7.3.3 Age and L1 use

In line with previous research, the two age groups of YLs made a limited use of the L1 (Azkarai and García Mayo, 2016; García Mayo and Hidalgo, 2017; Pinter, 2007). The younger children employed the L1 in less than 23.21% of the total number of AS-units they produced. The highest ratio of L1 use per AS-unit by older the group was 16.16%. When considering other studies which have addressed this aspect with YLs in EFL settings, the children in our study employed the L1 less than Azkarai and García Mayo (2016) reported (below 36%). The more limited use of the L1 by the learners in the current study may be due to the different methodologies followed by the schools the participants attended (e.g. Azkarai and García Mayo, 2016; Azkarai and Imaz Agirre, 2016; García Mayo and Lázaro Ibarrola, 2015): whereas the school in the present study followed a CLIL approach, with half the school subjects taught through the TL, in Azkarai and García Mayo (2016) the learners received only five hours of TL instruction. Research addressing the L1 use of CLIL learners has demonstrated a lower frequency of use, e.g. Lázaro Ibarrola and Hidalgo (2017a) (10.49%) and García Mayo and Hidalgo (2017)

(11.9%), closer to the ratio of use of our participants. In any case, our findings also show that younger learners, who had a slightly lower command of the TL, as seen in the analyses of CAF, need to rely more on the L1. This finding contrasts with García Mayo and Imaz Agirre (2017), who attributed a greater use of the L1 by Year 6 learners (one year older than our Year 5 participants) to the task not being engaging enough for this age group. At this point, we can only speculate that the PPT we employed was motivating for our Year 5 learners, maybe because they were a year younger than the participants in the study by García Mayo and Imaz Agirre (2017).

As seen in section 7.2.3 above, the L1 serves functions that assist task completion, mainly to assist learners as they cope with unknown TL vocabulary and for metacognitive talk. This first finding mirrors previous research studies (e.g. Azkarai and García Mayo, 2016; García Mayo and Hidalgo, 2017; García Mayo and Lázaro Ibarrola, 2015; Lázaro Ibarrola and Azpilicueta Martínez, 2015; Lázaro Ibarrola and Hidalgo, 2017a; Storch and Aldosari, 2010). These two L1 functions were significantly more frequent in the output of the younger learners. The differences in the use of discourse markers by our participants were not statistically significant. Another difference between these two age groups is the fact that younger learners exhibit a more childish behaviour. This fact was displayed in how some learners drifted away from the task at hand and started talking about some unrelated topic and resorted to the L1 for off-task talk. Nevertheless, as seen in the results section, this happened seldom. To the best of our knowledge, off-task talk has not been considered in studies with young EFL learners, and it rarely appears in adult studies. Azkarai and García Mayo (2015) reported off-task talk as the least frequent L1 function in the oral production of college participants while carrying out oral communicative tasks. This result has further strengthened the claim that YLs make a limited use of their L1 and that, with a few exceptions, the L1 serves functions that facilitate task completion. Further L1 use with functions that are unrelated to the task discourse are rare.

The most common L1 function in the two age groups was to compensate for unknown vocabulary. In the production of the Year 5 learners, the vocabulary function was the most frequent in the three task performances. In the output of the younger learners, vocabulary was also the most common L1 function at T1 and T2. At T3 however, the percentage of use of the L1 for metacognitive talk was higher than that of vocabulary. As discussed in section 7.2.3 during the examination of the impact of procedural TR on the

performance of the two groups together, metacognitive talk shows an increasing tendency across task performance, whereas L1 vocabulary decreases significantly. At T3, the use of the L1 to cover for unknown vocabulary dropped in the oral production of the two groups while metacognitive talk increased. However, when the data of the two groups was examined separately, the fall in the use of L1 vocabulary was statistically significant only in the Year 5 data. The second most frequent function, metacognitive talk decreased, but only in the output of Year 3 and only at T2, to then increase again at T3. Since the differences between these two functions in the production of the Year 3 group were not as big as in that of the older learners', Year 3 metacognitive talk surpassed L1 vocabulary at T3. As indicated above, the opposite tendencies of these two functions may stem from a drop in the learners' motivation and a shift of their attention to a greater focus on the task-related discourse that took place, presumably, in order to complete the task with as little delay as possible. Despite this, differences between the two groups were evident: the rise in the frequency of the metacognitive function at each task performance was not significant in the performance of Year 5 whilst the decrease in the use of L1 vocabulary did not reach significance in the production of the Year 3 group.

Therefore, the main effect of age on the changes in the use of L1 is a drop in the L1 vocabulary in the oral production of Year 5 group: older YLs appear to use familiarity with the task to concentrate on TL vocabulary. Year 3 learners, on the other hand, seem to need more practice, i.e. more repetitions, for the use of the L1 to assist them with unknown TL vocabulary to decrease. No statistically significant differences in the frequency of use of the L1 functions by the two groups were identified. Even though the output of the younger learners contained a significantly greater ratio of L1 per AS-unit, the distribution of the L1 uses is not significantly different between the two groups.

Our results support Hypothesis 12: younger YLs make a greater overall use of the L1 than their older counterparts. As we anticipated, the most frequent purpose of L1 use was to compensate for unknown vocabulary and metacognitive talk. Taking this into consideration, it is expected that by the end of primary school, the oral production of YLs will contain very few L1 terms, and that the diminished L1 use will assist task completion and, eventually, language acquisition.

7.4 Conclusion

This chapter has provided a discussion of our findings in light of each of the hypotheses and the research questions posed in Chapter 5. Figure 8 illustrates the most notable findings related to the YLs' use of NoM strategies, and the functions these serve. We have seen how the ability to provide support to the interlocutor seems to increase with age, as the older YLs use NoM strategies that serve this function abundantly whereas our younger participants negotiate mostly to solve communication problems. Moreover, younger learners repair breakdowns by using clarification requests, while older YLs resort to confirmation checks, which are slightly more complex forms. Based on previous studies, and on our own observations, we have proposed the four first developmental stages of the acquisition of the NoM strategies by young language learners, as illustrated in Figure 7.

Figure 8 *YLs' NoM.*



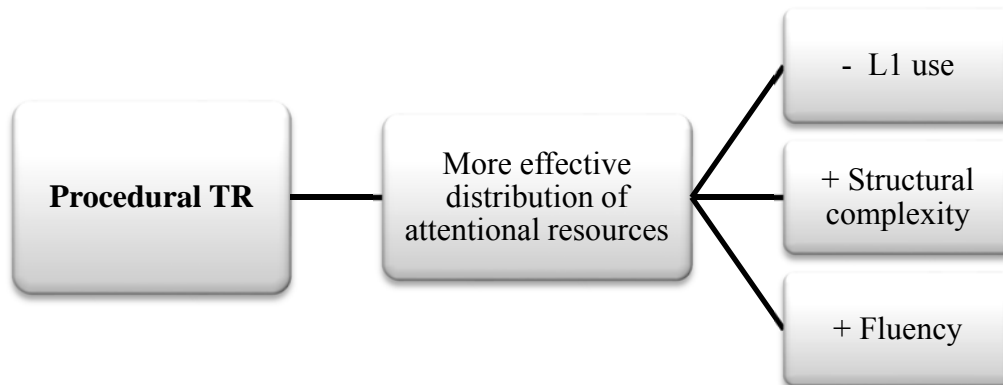
More differences between the two age groups examined are related to the learners' general performance: Year 5 learners' (age 10-11) oral production during peer-peer interaction is more complex, accurate, fluent and contains fewer L1 terms than that of their younger counterparts (age 8-9).

From the examination of the effect of procedural TR on the functions of the NoM strategies, another interesting phenomenon between the two most common functions (i.e. repairing communication breakdowns and confirming successful communication) is observed in the performance of the two age groups across tasks: as the amount of NoM strategies to confirm successful communication increases, the number of strategies to repair communication breakdowns diminishes.

As regards the effect of procedural TR on the learners' general performance, our results indicate that the learners achieved a more effective organisation of attentional resources

and, consequently, their performance became more structurally complex, more fluent and contained fewer L1 terms, as Figure 9 depicts.

Figure 9 *Effect of procedural TR on YLs' oral production.*



Finally, YLs use their shared L1 wisely and with functions that facilitate task performance, mainly to assist them as they cope with unknown TL vocabulary or to deal with task procedure (metacognitive talk).

Several conclusions can be drawn from these findings, as well as implications for SLA theory and methodology. The next chapter provides some concluding remarks and pedagogical implications derived from our findings. We are aware of the limitations of our study and these will be presented together with directions for future research.

PART III
CONCLUSIONS AND
CONTRIBUTIONS

CHAPTER 8 Conclusion: implications, limitations and future research

This final chapter summarises the aims of the current dissertation as well as the main conclusions derived from the results obtained in the data analyses. The major implications of our findings will be underlined and the limitations of our study will be acknowledged. Additionally, future research directions on child task-supported interaction will be outlined.

8.1 Conclusions

The present study set out to examine the effects of procedural TR on the oral production of young EFL learners of two different age groups (8-9 and 10-11). We have focused on the NoM, general performance (CAF) and L1 use across three task performances. We have sought to contribute to our understanding of how peer-peer interaction and the repetition of a task affect YLs' oral interactions. Even though research addressing YLs' SLA is growing, this study is one of the first attempts to investigate the impact of procedural TR and age on the members of this cohort from an interactionist perspective. We hope to have shed light on the emergent, though still relatively scant, research on YLs' peer-peer oral interaction.

The results from this study indicate that YLs negotiate mostly to repair communication breakdowns. Nevertheless, as children mature their ability to provide support to their interlocutors increases. This growth is recognisable in the employment of different NoM strategies to let others know that there are no communication difficulties, that is, by confirming successful communication. Older YLs negotiate mainly with this aim. This evidence suggests different developmental stages of the acquisition of strategies to negotiate for meaning. These stages range from nearly nonexistent negotiation, to the use of strategies concerned with 'self' and finally to those concerned with 'other'. NoM strategies to repair communication breakdowns and those used to confirm successful communication seem to be complementary as, apparently, one gives way to the other. The development in the use of these two functions is visible at a smaller scale thanks to the effect of procedural TR: while the former decrease, the latter increase.

More differences between the two age groups were found when examining the learners' general performance: the oral production of the 10-11 year-olds is more complex, accurate and fluent than that of the 8-9 years old learners. In addition, as shown in Table

16, the influence of age on task completion time is also significant and younger learners need significantly more time to carry out each task.

Procedural TR has been proven to have positive effects on YLs' fluency and structural complexity, which improve simultaneously. Both groups' fluency significantly improved at T3, although the increase in complexity in the last performance did not reach statistical significance for either group individually. The time the learners in the two groups employed to complete the task decreases significantly (Table 17). On the other hand, trade-off effects between complexity and accuracy were also spotted. However, consistent with the extended Trade-off Hypothesis (Skehan and Foster, 2012), it is evident that procedural TR attenuates these negative effects. TR facilitates a more effective distribution of the learners' attentional resources, as it provides opportunities to face similar communicative situation more than once.

Finally, L1 use is limited, has functions that facilitate task completion, and decreases across tasks. In this respect, indicators of a change in the learners' behaviour have been observed at T3, and a potential differentiation between 'task-related functions' and 'not task-related' is outlined. Vocabulary is 'task-related' and it decreases upon procedural TR, whereas metacognitive talk and discourse markers may be considered not part of the language necessary to deal with the task content, and they remain apparently unaffected. Finally, younger learners, who had a lower command of the TL, make a more extensive use of the L1.

In what follows, we reflect on the implications of the findings reported above for the different aspects examined in the current study: oral collaborative tasks, procedural TR and age-related differences among young FL learners.

8.2 Pedagogical implications

This study has enhanced our understanding of how young EFL learners interact orally with age- and level-matched peers. Our findings have several implications for research into SLA. With the current dissertation we hope to inform teaching practices by exploring which aspects of FL learning are benefited by tasks that promote interaction and NoM among YLs.

Normally, due to the ratio of having only one instructor in a class of approximately 25 pupils, the learners' speaking opportunities are quite limited and largely dependent on the

teachers' turn-allocation, their types of questions, class activities and feedback. This dissertation has provided evidence supporting the use of collaborative tasks, and shown that pair work benefits young language learners' oral production in different ways: not only because our participants were able to complete the task autonomously, or with little intervention on the part of the researcher, but their general performance also improved. In spite of their young age and limited command of the TL, YLs successfully participated in a conversation in English with their partners and fulfilled their role as interlocutors. YLs worked on their own in the TL, making a limited use of their shared L1, which mainly served functions that facilitate task completion. L1 use has been found to play a positive role in FL learning, as it is generally used as a tool to help learners when difficulties with the TL arise.

Consequently, oral collaborative tasks have been once more confirmed as a very valuable tool to assist learners in the language learning process by creating occasions in which learners can use the TL in a meaningful context. This is particularly noteworthy in FL learning contexts where the opportunities to interact in the TL are quite limited outside the language classroom. Likewise, based on our results, practitioners should not consider L1 use as detrimental for language learning. They should be aware of the benefits L1 use can potentially lead to, and try to take advantage of a sensible use while working with the L2.

In addition, we have obtained results that demonstrate procedural TR offers YLs advantages in terms of general competence: improvements have been identified in structural complexity and fluency. Moreover, the initial trade-off between some CAF dimensions disappears, or becomes attenuated, upon procedural TR. Overall L1 use has decreased through the repetition of the PPT, although the decrease was not as noticeable in the examination of the performance of the two groups individually. According to our results, a better organisation of the linguistic resources, achieved through more language practice, leads to less L1 use. A not so positive side of L1 use during TL interaction are the trade-off effects between L1 use and some accuracy measures, such as the number of EFAS. On the other hand, when the participants used a more varied vocabulary, L1 use increased, together with the number of errors. Thus, the learners' L1 may also serve as an element of support that affords YLs the self-confidence to try and stretch their interlanguage to the maximum by, for instance, producing a more varied vocabulary. As

described above, L1 use should not be considered as negative for language learning, as it may represent a necessary tool at a specific stage of the language learning process.

Therefore, a pedagogical recommendation would be that teachers employ tasks with the same procedure more than once in the language classroom, as YLs' interlanguage develops positively thanks to the opportunities to face similar communicative situations repetitively. Another practical implication of our findings is the advantage for teachers of having the possibility to recycle and re-use tasks. This aspect should not be underestimated given that preparation time is usually limited among school teachers.

On the downside, one of the CAF dimensions, accuracy, did not seem to significantly benefit from procedural TR in the oral output of any of the two age groups. This should be considered when implementing this type of task in the classroom: if the objective of the lesson is to improve oral accuracy, complementary activities should then be included. In addition, many of the errors identified in our data were related to L1 structural transfers, which remained unnoticed and uncorrected. Thus, we recommend the inclusion of focus on form activities to maximise the potential of collaborative tasks for YLs that share the L1. Further, some NoM strategies decreased at T3. Accordingly, if the objective of the lesson is to promote negotiation, this type of TR has been proven not to be the most suitable.

The results of this study endorse previous research suggesting that the ability to collaborate and understand our partner's needs increases with age (e.g. Oliver, 2009; Pinter, 2007). The older YLs provided more support to their interlocutors, as opposed to their younger counterparts who were more focused on conveying their own message. Older learners' also displayed a higher command of the TL, revealing a more structurally complex, accurate and fluent oral production. Through procedural TR, some CAF dimensions (i.e. complexity) improved in the performance of the younger learners, whilst they did not show as much improvement in the production of the older group. The younger learners resorted more frequently to the L1, using it mainly to help them to complete the task. Nevertheless, some learners in this group also used it for off-task talk, evidencing their different developmental stage and less mature behaviour. These findings highlight the importance of taking into account the differences between children of different ages. Finally, as stated above, task completion times varied significantly between the two age groups, with the younger learners devoting more time to perform

each task. As reported in previous research studies, childhood is a period of many changes and different age children have different needs and display different behaviours. Consequently, tasks, and task procedures, which may be appropriate for a specific age group, may not be appropriate for another age group.

We expect that the results obtained and the guidelines proposed above will be valuable for pedagogical practice and will help teachers to implement new methods that will offer YLs more learning opportunities, particularly more oral production opportunities, in the FL classroom. The following section details the limitations that should be considered alongside our findings, and outlines several areas for future research on child SLA.

8.3 Limitations and future research directions

We are aware that our research has some shortcomings that need to be acknowledged. The first one is that we have only considered a single task, albeit one frequently employed in FL lessons and considered the most suitable to trigger interaction. Additionally, the fact that the data were collected under laboratory conditions may also have influenced our results. Thus, more experimental research is required to determine if this is, in fact, the case. These limitations evidence the difficulty of collecting data on YLs' performance. Access to schools is quite restricted and we consider ourselves fortunate to have been able to work with these children and their teachers.

Replications with a wider range of tasks, also with different levels of complexity, would help to assess the influence of task type on YLs' peer-peer interaction. Special attention should be paid to the vocabulary needed to complete the task, as lexical complexity was the only complexity measure that was not affected by procedural TR. Future studies should also target learners with a different level of the TL and ages further apart to help gain a deeper understanding of the role these two variables play on YLs' oral performance. We have obtained encouraging results with regard to peer-peer oral interaction and we believe that they should be validated by studies taking place in a classroom setting. Concerning the procedure followed, it is recommended that investigations of the impact of the time span between task performances as well as the number of repetitions required for procedural TR to show its positive effects (e.g. increased fluency and complexity, less reliance on the L1) be increased.

One drawback related to our methodology is that the classification we have used to examine NoM has not been widely employed by other researchers so far, thus hindering the comparability of our findings. Our classification broadens previous methods by expanding the number of strategies learners employ during oral interaction, as well as offering a new perspective on NoM strategies. Of particular interest is the fact that this classification considers a new strategy which has turned out to be one of the most common in the output of the older YLs (acknowledgements). More data collection is required to determine exactly how this new strategy, and the function it serves, affects previous claims of a lower NoM by older YLs. In order to further examine the patterns of NoM among young EFL learners, it would be interesting to explore the underlying reasons why interaction strategies have been reported to decline. Does NoM decrease because older participants have a better command of the TL and are able to understand each other and resolve the task without difficulties, or because they are in fact relying on a different set of strategies not considered in previous research (i.e. strategies to confirm successful communication).

Given that the focus of our study was on a specific set of measures used to examine the three CAF dimensions, there is some likelihood that different evaluations could have arisen if the focus had been on other measures. Therefore, more research using the same set of measures that allows for comparability is needed in order to confirm the reliability of our findings.

In relation to the task used, tasks with different levels of complexity would perhaps lead to different results. Our tasks might have been too easy for the older learners to reach their maximum learning potential. Since the repetition of the PPT seems to produce no significant effects on accuracy, replicating the current study employing a task with a greater focus on language structures, or including FonF sessions between tasks, may be helpful to clarify the impact of procedural TR on the accuracy of YLs' oral production. Additionally, further work that includes different tasks and controls for vocabulary would probably shed some light on the effect of procedural TR on the development of lexical complexity.

Finally, the administration of motivation questionnaires would have definitely determined potential reasons for the changes in the frequency of L1 use (e.g. drops and rises in metacognitive talk) and whether the task we used is age-appropriate. Stimulated recall

interviews may also have explained why learners resort to their L1 in different situations during each task performance. In addition, although L1 terms were not abundant, manifold L1 structural transfers were found in our data, but lie beyond the scope of this dissertation. Pronunciation errors have not been considered either. Had the learners not shared the L1, structural transfers and pronunciation problems may have interfered with comprehension and led to communication difficulties. All our participants shared Spanish as their L1, which may have been the reason for the relatively few instances of communication breakdowns and little focus on form, hence limited feedback. Therefore, future research with language learners with different L1s is desirable. Another interesting aspect that would contribute to the understanding of child L2 oral interaction is the consideration of the extent to which YLs use the TL for the same purposes as they use their L1 (e.g. discourse markers, metacognitive talk, unknown vocabulary).

As demonstrated by the current dissertation, age is a crucial factor to be considered when examining FL acquisition. Even within the same childhood stage (middle childhood), and with learners of ages not so far apart, distinct traits in each group have been identified. NoM is one of the aspects found to be approached differently by the two age groups, particularly with regard to the use of a wider set of strategies. The learners' general performance also differed greatly, and even the procedure followed in our study (i.e. procedural TR) affected some dimensions of the learners' performance depending on their age. This study has shed light on how best to address YLs' FL learning and the benefits of oral collaborative task-supported interaction.

The YLs in this study carried out the tasks autonomously, tasks which have triggered NoM and brought about various benefits for language acquisition. We would like to underline that, in spite of the young age of the participants, their limited command of the TL, and the differences between the two age groups, the whole experience has demonstrated that it is definitely worth letting the children do the talking.

REFERENCES

- Ackerl, C. (2007). Lexico-grammar in the essays of CLIL and non-CLIL students: Error analysis of written production. *VIEWS (Vienna English Working Papers)*, 16, (3), 6-11.
- Adams, R. (2007). Do second language learners benefit from interacting with each other? In A. Mackey (Ed.), *Conversational Interaction in Second Language Acquisition* (pp. 29-52). Oxford: Oxford University Press.
- Adams, R. and Newton, J. (2009). TBLT in Asia: Constraints and opportunities. *Asian Journal of English Language Teaching*, 19, 1-17.
- Agustín-Llach, M.P. (2009). The role of lexical errors in L2 reading proficiency of young EFL learners from a longitudinal perspective. *Revista de Filología Inglesa*, 30, 7-22.
- Admiraal, W., Westhoff, G. and de Bot, K. (2006). Evaluation of bilingual secondary education in the Netherlands: Students' language proficiency in English. *Educational Research and Evaluation*, 12, (1), 75-93.
- Ahmadian, J. (2012). Task repetition in ELT. *ELT Journal*, 66, (3), 380-382.
- Ahmadian, J. and Tavakoli, M. (2011). The effect of simultaneous use of careful on-line planning and task repetition on accuracy, complexity and fluency in EFL learners' oral production. *Language Teaching Research*, 15, (1), 35-59.
- Airey, J. and Linder, C. (2006). Language and the experience of learning university physics in Sweden. *European Journal of Physics*, 27, (3), 553-560.
- Alegría de la Colina, A. and García Mayo, M.P. (2009). Oral interaction in task-based EFL learning: The use of the L1 as a cognitive tool. *International Review of Applied Linguistics (IRAL)*, 47, (3), 325-345.
- Alcón Soler, E. and García Mayo, M.P. (2008). Focus on form in and learning outcomes in the foreign language classroom. In J. Philp, R. Oliver & A. Mackey (Eds.). *Child's Play? Second Language Acquisition and the Younger Learner* (pp. 173-192). Amsterdam: John Benjamins.
- Antón, M. and DiCamilla, F. (1998). Socio-cognitive functions of L1 collaborative interaction in the L2 classroom. *Canadian Modern Language Review*, 54, 314-342.
- Azkarai, A. (2013). *Gender and Task Modality in EFL Task-based Interaction*. Unpublished Doctoral Dissertation. University of the Basque Country, Spain.
- Azkarai, A. and García Mayo, M.P. (2012). Does gender influence task performance in EFL? Interactive tasks and language related episodes. In E. Alcón Soler & M.P. Safont Jordá (Eds.), *Language Learners' Discourse Across L2 Instructional Settings* (pp. 249-278). Amsterdam: Rodopi.

REFERENCES

- Azkarai, A. and García Mayo, M.P. (2015). Task-modality and L1 use in EFL oral interaction. *Language Teaching Research*, 19, (5), 550-571.
- Azkarai, A. and García Mayo, M.P. (2016). Task repetition effects on L1 use in EFL child task-based interaction. *Language Teaching Research*. doi: 10.1177/1362168816654169
- Azkarai, A., and Imaz Agirre, A. (2016). Negotiation of meaning strategies in child EFL mainstream and CLIL settings. *TESOL Quarterly*, 50, 844-870.
- Bagheri, M., Rahimi, F. and Riasati, M. (2012). Communicative interaction in language learning tasks among EFL learners. *Journal of Language Teaching and Research*, 3, (5), 948-952.
- Berk, L. (2006). *Child Development*. Boston: Pearson Education.
- Birdsong, D. (2005a). Interpreting age effects in second language acquisition. In J.F. Kroll & A.M.B. DeGroot (Eds.), *Handbook of Bilingualism: Psycholinguistic Perspectives* (pp. 109-127). New York: Oxford University Press.
- Birdsong, D. (2005b). Native-likeness and non-nativeness in L2. A research. *International Review of Applied Linguistics*, 43, 319-328.
- Bley-Vroman, R. (2009). The evolving context of the Fundamental Difference Hypothesis. *Studies in Second Language Acquisition*, 31, (2), 175-198.
- Bret Blasco, A. (2014). *L2 English Young Learners' Oral Production Skills in Instructed Second Language Contexts: A Longitudinal Study of CLIL and EFL Settings*. Unpublished Doctoral Dissertation. Universitat Autònoma de Barcelona, Spain.
- Bruton, A. (2002). From tasking purposes to purposing tasks. *ELT Journal*, 56, (3), 280-288.
- Bruton, A. (2011a). Are the differences between CLIL and non-CLIL groups in Andalusia due to CLIL? A reply to Lorenzo, Casal and Moore (2010). *Applied Linguistics*, 32, (2), 236-241.
- Bruton, A. (2011b). Is CLIL so beneficial, or just selective? Re-evaluating some of the Research. *System*, 39, (4), 523-532.
- Butler, Y. G. and Zeng, W. (2014). Young foreign language learners' interactions during task-based paired assessment. *Language Assessment Quarterly*, 11, (1), 45-75.
- Bygate, M. (1996). Effects of task repetition: Appraising the development language of learners. In J. Willis and D. Willis (Eds.), *Challenge and Change in Language Teaching* (pp. 134-146). Oxford: Heinemann.
- Bygate, M. (1999a). Quality of language and purpose of task: Patterns of learners' language on two oral communication tasks. *Language Teaching Research*, 3, (3), 185-214.

- Bygate, M. (1999b). Task as context for the framing, reframing and unframing of language. *System*, 27, (1), 33-48.
- Bygate, M. (2001). Effects of task repetition on the structure and control of language. In M. Bygate, P. Skehan & M. Swain (Eds.), *Task-based Learning: Language Teaching, Learning and Assessment* (pp. 23-48). London: Longman.
- Bygate, M. (2006). Areas of research that influence L2 speaking instruction. In E. Uso-Juan & A. Martinez-Flor (Eds.), *Current Trends in the Development and Teaching of the Four Language Skills* (pp. 159-186). Berlin: Mouton de Gruyter.
- Bygate, M. (2009). Effects of task repetition on the structure and control of oral language. In K. Van den Branden, M. Bygate, & J. Norris (Eds.), *Task-based Language Teaching: A Reader* (pp. 249-274). Philadelphia, PA: John Benjamins
- Bygate, M. and Samuda, V. (2005). Integrative planning through the use of task-repetition. In R. Ellis (Ed.), *Planning and Task Performance in a Second Language* (pp. 37-74). Amsterdam: John Benjamins.
- Bygate, M., Skehan, P. and Swain, M. (2001). Introduction. In M. Bygate, P. Skehan & M. Swain. (Eds.), *Task-based Learning: Language Teaching, Learning and Assessment* (pp. 1-20). London: Longman.
- Cameron, L. (2001). *Teaching Language to Young Learners*. Cambridge: Cambridge University Press.
- Cameron, L. (2003). Challenges for ELT from the expansion in teaching children. *ELT Journal*, 57, (2), 105-112.
- Cammarata, L. and Tedick, D. (2012). Balancing content and language in instruction: The experience of immersion teachers. *The Modern Language Journal*, 96, (2), 251-269
- Carless, D. (2004). Issues in teachers' reinterpretation of a task-based innovation in primary schools. *TESOL Quarterly*, 38, (4), 639-662.
- Carless, D. (2012). TBLT in EFL settings: looking back and moving forward. In A. Shehadeh & C.A. Coombe (Eds.), *Task-Based Language Teaching in Foreign Language Contexts: Research and Implementation* (pp. 187-214). Amsterdam: John Benjamins.
- Cenoz, J. (2013). Towards an educational perspective in CLIL language policy and pedagogical practice [Special issue]. *International Journal of Bilingual Education and Bilingualism*, 16, (3), 389-394.
- Cenoz, J. (2015). Content-based instruction and content and language integrated learning: the same or different? *Language, Culture and Curriculum*, 28, (1), 8-24.

REFERENCES

- Cenoz, J. Genesee, F. and Gorter, D. (2014). Critical analysis of CLIL: taking stock and looking forward. *Applied Linguistics*, 35, 243-262.
- Chacón, C.T. (2012). Task-based language teaching through film-oriented activities in a teacher education program in Venezuela. In A. Shehadeh & C.A. Coombe (Eds.), *Task-Based Language Teaching in Foreign Language Contexts: Research and Implementation* (pp. 241-266). Amsterdam: John Benjamins.
- Chan, S.P. (2012). Qualitative differences in novice teachers' enactment of task-based language teaching in Hong Kong primary schools. In A. Shehadeh & C.A. Coombe (Eds.), *Task-based Language Teaching in Foreign Language Contexts: Research and Implementation* (pp. 137-161). Amsterdam: John Benjamins.
- Chomsky, N. (1986). *Knowledge of Language: Its Nature, Origin and Use*. New York: Praeger.
- Cohen, J.W. (1988). *Statistical Power Analysis for the Behavioural Sciences*. Hillsdale, NJ: Lawrence Erlbaum Associates.
- Collins, L. and Muñoz, C. (2016). The foreign language classroom: Current perspectives and future considerations. *The Modern Language Journal*, 100, 133-147.
- Comunidad de Madrid (2015). *Madrid, Comunidad Bilingüe / Madrid, a Bilingual Community 2014-15*. Retrieved the 01/06/15 from <http://www.madrid.org/bvirtual/BVCM016212.pdf>
- Council of Europe. *Common European Framework of Reference for Languages: learning, teaching, assessment*. Retrieved the 05/05/15 from http://www.coe.int/t/dg4/linguistic/Source/Framework_EN.pdf
- Council of the European Union (2002). *Presidency Conclusions, Barcelona European Council, 15 – 16 March 2002*. Retrieved the 05/04/2015 from http://ec.europa.eu/invest-in-research/pdf/download_en/barcelona_european_council.pdf
- Coyle, D. (2007). Content and language integrated learning: Towards a connected research agenda for CLIL pedagogies. *The International Journal of Bilingual Education and Bilingualism*. 10, (5), 543-562.
- Coyle, D. (2010). Foreword. In Y. Ruiz de Zarobe and D. Lasagabaster (Eds.), *CLIL in Spain. Implementation, Results and Teacher Training* (pp. vii-viii). Newcastle upon Tyne: Cambridge Scholars.
- Coyle, D., Hood, P. and Marsh, D. (2010). *CLIL: Content and Language Integrated Learning*. New York: Cambridge University Press.
- Dalton-Puffer, C. (2007). *Discourse in Content and Language Integrated Learning (CLIL) Classrooms*. Amsterdam: John Benjamins.

- Dalton-Puffer, C. (2011). Content and language integrated learning: from practice to principles. *Annual Review of Applied Linguistics*, 31, 182-204.
- Dalton-Puffer, C., Llinares, A., Lorenzo, F. and Nikula, T. (2014). FORUM “You can stand under my umbrella”: Immersion, CLIL and bilingual education. A response to Cenoz, Genesee & Gorter (2013). *Applied Linguistics*, 35, (2), 213-218.
- Dalton-Puffer, C., Nikula, T. and Smit, U. (2010). Charting policies, premises and research on content and language integrated learning. In C. Dalton-Puffer, T. Nikula, & U. Smit (Eds.), *Language Use and Language Learning in CLIL Classrooms* (pp. 1-19). Amsterdam: John Benjamins.
- Dalton-Puffer, C. and Smit, U. (2007). Introduction. In C. Dalton-Puffer & U. Smit (Eds.), *Empirical Perspectives on CLIL Classroom Discourse* (pp. 7-23). Frankfurt: Peter Lang.
- Dalton-Puffer, C. and Smit, U. (2013). Content-and-language-integrated learning: a research agenda. *Language Teaching*, 46, (4), 545-559.
- DeKeyser, R. (2000). The robustness of critical period effects in second language acquisition. *Studies in Second Language Acquisition*, 22, 493-533.
- DeKeyser, R. (2003). Implicit and explicit learning. In C. Doughty & M. Long (Eds.), *Handbook of Second Language Acquisition* (pp. 313-348). Oxford: Blackwell.
- DiCamilla, F. and Antón, M. (2012). Functions of L1 in the collaborative interaction of beginning and advanced second language learners. *International Journal of Applied Linguistics*, 22, (2), 160-188.
- Dimroth, C. (2008). Perspectives on second language acquisition at different ages. In J. Philp, R. Oliver & A. Mackey (Eds.), *Second Language Acquisition and the Younger Learner: Child's Play?* (pp. 53-79). Amsterdam: John Benjamins.
- Donato, R. (1994). Collective scaffolding in second language learning. In J.P. Lantolf & G. Appel (Eds.), *Vygotskian Approaches to Second Language Research* (pp. 33-56). Norwood, NJ: Ablex.
- Doughty, C. and Varela, E. (1998). Communicative focus on form. In C. Doughty & J. Williams (Eds.), *Focus on form in classroom second language acquisition* (pp. 114-138). Cambridge: Cambridge University Press.
- Doughty, C. and Williams, J. (1998). *Focus on Form in Classroom Second Language Acquisition*. New York: Cambridge University Press.
- Dufficy, P. (2004). Predisposition to choose: The language of an information gap task in a multilingual primary classroom. *Language Teaching Research*, 8, (3), 241-261.
- Ellis, R. (2003). *Task Based Language Learning and Teaching*. Oxford University Press: Oxford.

REFERENCES

- Ellis, R. (Ed.) (2005). *Planning in Task-based Performance*. Amsterdam: John Benjamins.
- Ellis, R., Basturkmen, H. and Loewen, S. (2002). Doing focus-on-form. *System*, 30, 419-432.
- Enever, J. (Ed.) (2011). *ELLie: Early Language Learning in Europe*. London: British Council.
- Enever, J. (2016). Primary ELT: issues and trends. In G. Hall (Ed.), *The Routledge Handbook of English Language Teaching* (pp. 353-362). Abingdon: Routledge.
- European Union Commission (1995). *Teaching and learning – Towards the learning society. White Paper on Education and Training*: Brussels. Retrieved the 13/03/15 from: http://europa.eu/documents/comm/white_papers/pdf/com95_590_en.pdf
- European Union Commission (1995). Council Resolution of 31 March 1995 on improving and diversifying language learning and teaching within the education systems of the European Union. *Official Journal C 207*, 12.08.1995, (pp.1-5). Retrieved the 13/03/15 from: <http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=CELEX:31995Y0812%2801%29:EN:HTML>
- European Union Commission (2008). Communication from the Commission to the European Parliament, the Council, the European Economic and Social Committee and the Committee of the Regions. *Multilingualism: an Asset for Europe and a Shared Commitment*. Retrieved the 23/03/15 from: <http://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:52008DC0566&from=EN>
- Eurydice European Network (2006a). *Content and Language Integrated Learning (CLIL) at School in Europe*. Brussels: European Commission. Retrieved the 15/03/15 from <http://bookshop.europa.eu/en/content-and-language-integrated-learning-clil-at-school-in-europe-pbNCX106001/>
- Eurydice European Network (2006b). *Content and Language Integrated Learning (CLIL) at School in Europe. Spain. National Description 2004/05*, Eurydice Unit. Brussels: European Commission.
- Eurydice European Network (2012). *Key Data on Teaching Languages at School in Europe*. Brussels: European Commission. Retrieved the 07/05/2015 from http://eacea.ec.europa.eu/education/eurydice/documents/key_data_series/143en.pdf
- Fernández Dobao, A. (2014). Vocabulary learning in collaborative tasks: A comparison of pair and small group work. *Language Teaching Research*, 18, (4), 497-520.
- Fernández Fontecha, A. (2009). Spanish CLIL: Research and official actions. In Y. Ruiz de Zarobe & Jiménez Catalán (Eds.), *Content and Language Integrated Learning: Evidence from Research in Europe* (pp. 3-21). Bristol: Multilingual Matters.

- Foster, P. (1998). A classroom perspective on the negotiation of meaning. *Applied Linguistics*, 19, 1-23.
- Foster, P. and Ohta, A.S. (2005). Negotiation for meaning and peer assistance in second language classrooms. *Applied Linguistics*, 26, 402-430.
- Foster, P. and Skehan, P. (1996). The influence of planning on performance in task-based learning. *Studies in Second Language Acquisition*, 18, (3), 299-324.
- Foster, P., Tonkyn, A. and Wigglesworth, G. (2000). Measuring spoken language: A unit for all reasons. *Applied linguistics*, 21, (3), 354-375.
- Gagné, N. and Parks, S. (2013). Cooperative learning tasks in a Grade 6 intensive ESL class: Role of scaffolding. *Language Teaching Research*, 17, (2), 188-209.
- García Mayo, M.P. (Ed.) (2007). *Investigating Tasks in Formal Second Language Learning*. Clevedon: Multilingual Matters.
- García Mayo, M.P. (2011). The relevance of attention to L2 form in communicative classroom contexts. *Estudios de Lingüística Aplicada (ELIA)*, 11, 11-45.
- García Mayo, M.P. (Ed.) (2017a). *Learning Foreign Languages in Primary School. Research Insights*. Bristol: Multilingual Matters.
- García Mayo, M.P. (2017b). Focus and unfocused tasks. Entrada en J. Liantas (Ed.), *The TESOL Encyclopedia of English Language Teaching* (pp.in press). New York: Wiley.
- García Mayo, M.P. and Alcón Soler, E. (2013). Input, output. The interactionist framework. In J. Herschensohn & M. Young-Scholten (Eds.), *The Handbook of Second Language Acquisition* (pp. 209-229). Cambridge: Cambridge University Press.
- García Mayo, M.P. and García Lecumberri, M.L. (Eds.) (2003). *Age and the Acquisition of English as a Foreign Language*. Clevedon: Multilingual Matters.
- García Mayo, M.P. and Hidalgo, M.A. (2017) L1 use among young EFL mainstream and CLIL learners in task-supported interaction. *System*, 67, 132-145.
- García Mayo, M.P. and Imaz Agirre, A. (2016). Task repetition and its impact on EFL children's negotiation of meaning strategies and pair dynamics: an exploratory study. *The Language Learning Journal*, 44, (4), 451-466.
- García Mayo, M.P. and Imaz Agirre, A. (2017). Child EFL interaction: Age, instructional setting and development. In J. Enever & E. Lindgren (eds.), *Early Language Learning: Complexity and Mixed Methods* (pp. 249-268). Bristol: Multilingual Matters.
- García Mayo, M.P., Imaz Agirre, A., and Azkarai, A. (in press). Task repetition in EFL child oral interaction. In A.M. Ahmadian & M.P. García Mayo (eds.), *Recent Perspectives on Task-based Language Learning and Teaching*. Berlin: Mouton de Gruyter.

REFERENCES

- García Mayo, M.P. and Lázaro Ibarrola, A. (2015). Do children negotiate for meaning in task-based interaction? Evidence from CLIL and EFL settings. *System*, 54, 40-54.
- García Mayo, M.P. and Pica, T. (2000). Interaction among proficient learners: Are input, feedback and output needs addressed in a foreign language context? *Studia Linguistica: A Journal of General Linguistics*, 54, (2), 272-279.
- Gass, S. (1997). *Input, Interaction and the Second Language Learner*. Mahwah, NJ: Lawrence Erlbaum Associates.
- Gass, S. (2007). Input and interaction. In C.J. Doughty & M.H. Long (Eds.), *Handbook of Second Language Acquisition* (pp. 224-255). Oxford: Blackwell.
- Gass, S. and Mackey, A. (2006). Input, interaction and output: An overview. *AILA Review*, 19, 3-17.
- Gass, S. and Mackey, A. (2007). Input, interaction, and output in second language acquisition. In B. VanPatten & J. Williams (Eds.), *Theories in Second Language Acquisition: An Introduction* (pp. 175-199). Mahwah, NJ: Lawrence Erlbaum Associates.
- Gass, S. and Varonis, E. (1985). Task variation and nonnative/nonnative negotiation of meaning. In S.M. Gass & C. Madden (Eds.), *Input in Second Language Acquisition* (pp.149-161). Rowley, MA: Newbury House.
- Gass, S., Mackey, A., Alvarez-Torres, M.J. and Fernández-García, M. (1999). The effects of task repetition on linguistic output. *Language Learning*, 49, (4), 549-481.
- Gass, S., Mackey, A. and Ross-Feldman, L. (2005). Task-based interactions in classroom and laboratory settings. *Language Learning*, 55, (4), 575-661.
- Genesee, F. (1987). *Learning through Two Languages: Studies of Immersion and Bilingual Education*. Cambridge, MA: Newbury House.
- Gurzynski-Weiss, L. and Baralt, M. (2014). Exploring learner perception and use of task-based interactional feedback in face-to-face and computer-mediated modes. *Studies in Second Language Acquisition*, 36, (1), 1-37.
- Halliday, M.A.K. (1994). *An Introduction to Functional Grammar* (Second Edition). London: Edward Arnold.
- Hawkes, M. (2012) Using task repetition to direct learner attention and focus on form. *ELT Journal*, 66, (3), 327-335.
- Heras, A. and Lasagabaster, D. (2015). The impact of CLIL on affective factors and vocabulary learning. *Language Teaching Research*, 19, 70-88.
- Hellekjær, G.O. (2010). Language matters: Assessing lecture comprehension in Norwegian English medium higher education. In C. Dalton-Puffer, T. Nikula & U. Smit

(Eds.), *Language Use and Language Learning in CLIL Classrooms* (pp. 233-258). Amsterdam/Philadelphia: John Benjamins.

Housen, A., and Kuiken, F. (2009). Complexity, accuracy and fluency in second language acquisition. *Applied Linguistics*, 30, (4), 461-473.

Housen, A., Kuiken, F. and Vedder, I. (2012). Complexity, accuracy and fluency: Definitions, measurement and research. In Housen et al. (Eds.), *Dimensions of L2 Performance and Proficiency: Complexity, Accuracy and Fluency*. (pp. 1-20). Amsterdam: John Benjamins, 2012.

Hüttner, J. and Rieder-Bünemann, A. (2007). The effect of CLIL instruction on children's narrative competence. *VIEWS (Vienna English Working Papers)*, 16, (3), 20-27.

Iwashita, N. and Li, H. (2012). Patterns of corrective feedback in a task-based adult EFL classroom setting in China. In A. Shehadeh & C.A. Coombe (Eds.), *Task-based Language Teaching in Foreign Language Contexts: Research and Implementation* (pp. 137-161). Amsterdam: John Benjamins.

Jäppinen, A.K. (2005). Thinking and content learning of mathematics and science as cognitive development in content and language integrated learning (CLIL): Teaching through a foreign language in Finland. *Language and Education*, 19, (2), 147-168.

Jexenflicker, S. and Dalton-Puffer, C. (2010). The CLIL differential: Comparing the writing of CLIL and non-CLIL students in higher colleges of technology. In C. Dalton-Puffer, T. Nikula, & U. Smit (Eds.), *Language Use and Language Learning in CLIL Classrooms* (pp. 169-190). Amsterdam/Philadelphia: John Benjamins.

Junta de Andalucía (2004). *Plan de Fomento de Plurilingüismo*. Sevilla: Consejería de Educación. Retrieved the 04/05/2015 from:
<http://cms.ual.es/idc/groups/public/@vic/@vinternacional/documents/documento/jc80302.pdf>

Keck, C.M., Iberri-Shea, G., Tracy-Ventura, N. and Wa-Mbaleka, S. (2006). Investigating the empirical link between task-based interaction and acquisition: A quantitative meta-analysis. In J.M. Norris & L. Ortega (Eds.), *Synthesizing Research on Language Learning and Teaching* (pp. 91-131). Amsterdam: John Benjamins.

Kim, Y. (2009). The effects of task complexity on learner-learner interaction. *System*, 37, 254-268.

Kim, Y. (2013). Effects of pretask modelling on attention to form and question development. *TESOL Quarterly*, 47, 8-35.

Kim, Y. and Tracy-Ventura, N. (2013). The role of task repetition in L2 performance development: What needs to be repeated during task-based interaction? *System*, 41, 829-840.

REFERENCES

- Krashen, S. (1982). *Principles and Practice in Second Language Acquisition*. New York: Pergamon Institute of English.
- Krashen, S. (1985). *The Input Hypothesis: Issues and Implications*. Oxford: Pergamon.
- Lasagabaster, D. (2011). English achievement and student motivation in CLIL and EFL settings. *Innovation in Language Learning and Teaching*, 5, 3-18.
- Lasagabaster, D. and Ruiz de Zarobe, Y. (Eds.). (2010). *CLIL in Spain: Implementation, Results and Teacher Training*. Newcastle upon Tyne, UK: Cambridge Scholars.
- Lasagabaster, D. and Sierra, J.M. (2010). Immersion and CLIL in English: more differences than similarities. *ELT Journal*, 64, (4), 367-375.
- Lázaro Ibarrola, A. (2016). Are CLIL learners simply faster or also different? Evidence from L1 use in the repair sequences and discourse markers of CLIL and EFL learners. *Vigo International Journal of Applied Linguistics*, 13, 127-145.
- Lázaro Ibarrola, A. and Azpilicueta Martínez, R. (2015). Investigating negotiation of meaning in EFL children with very low levels of proficiency. *International Journal of English Studies*, 15, (1), 1-21.
- Lázaro Ibarrola, A. and García Mayo M.P. (2012). L1 use and morphosyntactic development in the oral production of EFL learners in a CLIL context. *International Review of Applied Linguistics*, 50, 135-160.
- Lázaro Ibarrola, A. and Hidalgo, M.A. (2017a). Benefits and limitations of conversational interactions among young learners of English in a CLIL context. In M.P. García Mayo (Ed.), *Learning Foreign Languages in Primary School: Research Insights* (pp.86-102). Bristol: Multilingual Matters.
- Lázaro Ibarrola, A. and Hidalgo, M.A. (2017b). Repetition in task-based interaction among young EFL learners: Does it make a difference? *International Journal of Applied Linguistics*, 168, (2), 182-201.
- Leaver, B. and Willis, R.M. (Eds.) (2004). *Task-Based Instruction in Foreign Language Education. Practices and Program*. Washington, DC: Georgetown University Press.
- Lenneberg, E. (1967). *Biological Foundations of Language*. New York: Wiley.
- Li, S. (2010). The effectiveness of corrective feedback in SLA: A meta-analysis. *Language Learning*, 60, 309-365.
- Lightbown, P. and Spada, N. (2006). *How Languages Are Learned*. Oxford: Oxford University Press.
- Lim, L. and Low, E.L. (Eds.). (2009). Multilingual, globalizing Asia: Implications for policy and education. *AILA Review*, 22, 52-71. Amsterdam: John Benjamins Publishing.

- Llinares, A. (2015). Integration in CLIL: a proposal to inform research and successful pedagogy. *Language, Culture and Curriculum*, 28, (1), 58-73.
- Llinares, A. and Dafouz, E. (2010). Content and language integrated programmes in the Madrid region: Overview and research findings. In D. Lasagabaster and Y. Ruiz de Zarobe (Eds.), *CLIL in Spain: Implementation, Results and Teacher Training* (pp.95-114). Newcastle: Cambridge Scholars Publishing.
- Llinares, A. and Peña, I. (2014). A genre approach to the effect of academic questions on CLIL students' language production. *Language and Education*, 23, (1), 15-30.
- Llinares, A. and Whittaker, R. (2010). Writing and speaking in the history class: data from CLIL and first language contexts. In C. Dalton-Puffer, T. Nikula & U. Smit (Eds.), *Language Use in Content-and-Language Integrated Learning (CLIL)* (pp. 125-143). AILA Applied linguistic series (AALS). Amsterdam: John Benjamins Publishers.
- Loewen, S. (2004). Uptake in incidental focus on form and second language learning. *Language Learning*, 54, (1), 153-188.
- Loewen, S. (2005). Incidental focus on form and second language learning. *Studies in Second Language Acquisition*, 27, (3), 361-386.
- Loewen, S. (2015). *Introduction to Instructed Second Language Acquisition*. New York: Routledge.
- Long, M.H. (1981). Input, interaction and second language acquisition. In H. Winitz, (Ed.), *Native Language and Foreign Language Acquisition. Annals of the New York Academy of Sciences*, 379: 259-278.
- Long, M.H. (1983). Native speaker/non-native speaker conversation and the negotiation of comprehensible input. *Applied Linguistics* 4, (2), 126-141.
- Long, M.H. (1985). Input and second language acquisition theory. In S. Gass & C. Madden (Eds.), *Input and Second Language Acquisition* (pp. 377-393). Rowley, Mass.: Newbury House.
- Long, M.H. (1991). Focus on form: A design feature in language teaching methodology. In K. de Bot, D. Coste, R. Ginsberg & C. Kramsch (Eds.), *Foreign Language Research in Cross-Cultural Perspectives* (pp. 39-52). Amsterdam: John Benjamins.
- Long, M.H. (1996). The role of the linguistic environment in second language acquisition. In W.C. Ritchie & T.K. Bhatia (Eds.), *Handbook of Language Acquisition: Vol. 2. Second Language Acquisition* (pp. 413-468). New York: Academic Press.
- Long, M.H. (2007). *Problems in SLA*. Mahwah, NJ: Erlbaum.
- Long, M.H. and Crookes, G. (1992). Three approaches to task-based syllabus design. *Tesol Quarterly*, 26, (1), 27-56.

REFERENCES

- Long, M.H. and Porter, P.A. (1985). Group work, interlanguage talk, and second language acquisition. *TESOL Quarterly*, 19, (2), 207-228.
- Long, M.H. and Robinson, P. (1998). Focus on form: Theory, research and practice. In C. Doughty & J. Williams (Eds.), *Focus on Form in Classroom Second Language Acquisition* (pp. 15-41). Cambridge: Cambridge University Press.
- Lorenzo, F., Casal, S. and Moore, P. (2010). The effects of content and language integrated learning in European education: key findings from the Andalusian bilingual sections evaluation project. *Applied Linguistics*, 31, (3), 418-442.
- Loschky, L. and Bley-Vroman, R. (1993). Grammar and task-based methodology. In G. Crookes & S. Gass (Eds.), *Tasks and Language Learning: Integrating Theory and Practice* (pp. 123-167). Clevedon, England: Multilingual Matters Ltd.
- Lynch T. and Maclean J. (2000). Exploring the benefits of repetition and recycling of a classroom task. *Language Teaching Research*, 4, (3), 21-250.
- Lynch T. and Maclean J. (2001). "A case of exercising": effects of immediate task repetition on learners' performance. In M. Bygate, P. Skehan & M. Swain (Eds.), *Researching Pedagogic Tasks: Second Language Learning, Teaching and Testing* (pp. 141-162). Addison Wesley Longman.
- Lyster, R. (1998). Negotiation of form, recasts, and explicit correction in relation to error types and learner repair in immersion classrooms. *Language Learning*, 48, (2), 183-218.
- Lyster, R. (2001). Negotiation of form, recasts, and explicit correction in relation to error types and learner repair in immersion classrooms. *Language Learning*, 51, (1), 265-301.
- Lyster, R. (2004). Differential effects of prompts and recasts in form-focused instruction. *Studies in Second Language Acquisition*, 26, 399-432.
- Lyster, R. (2007). *Learning and Teaching Languages through Content: A Counterbalanced Approach*. Amsterdam, the Netherlands: John Benjamins.
- Lyster, R. and Izquierdo, J. (2009). Prompts versus recasts in dyadic interaction. *Language Learning*, 59, 453-498.
- Lyster, R. and Saito, K. (2010). Interactional feedback as instructional input: A synthesis of classroom SLA research. *Language, Interaction and Acquisition*, 1, 276-297.
- Macaro, E. (2005). Codeswitching in the L2 classroom: A communication and learning strategy. In E. Llorca (Ed.), *Nonnative Language Teachers: Perceptions, Challenges and Contributions to the Profession* (pp. 63-84). New York, NY: Springer.
- Mackey, A. (1999). Input, interaction and second language development: An empirical study of question formation in ESL. *Studies in Second Language Acquisition*, 21, (4), 557-587.

- Mackey, A. (Ed.) (2007). *Conversational Interaction in Second Language Acquisition: A Collection of Empirical Studies*. Oxford: Oxford University Press.
- Mackey, A. (2012). *Input, Interaction and Corrective Feedback in L2 Learning*. Oxford: Oxford University Press.
- Mackey, A., Abbuhl, R. and Gass, S.M. (2012). Interactionist approaches. In S.M. Gass & A. Mackey (Eds.). *Handbook of Second Language Acquisition* (pp. 7-23). New York: Routledge.
- Mackey, A., Gass, S.M. and McDonough, K. (2000). How do learners perceive interactional feedback? *Studies in Second Language Acquisition*, 22, (4), 471-497.
- Mackey, A. and Goo, J. (2007). Interaction research in SLA: A meta-analysis and research synthesis. In A. Mackey (Ed.), *Conversational Interaction in Second Language Acquisition: A Series of Empirical Studies*. (pp. 407-453). (Oxford applied linguistics). Oxford: Oxford University Press.
- Mackey, A., Kaganas, A.P. and Oliver, R. (2007). Task familiarity and interactional feedback in child ESL classrooms. *Tesol Quarterly*, 41, (2), 285-312.
- Mackey, A. and Oliver, R. (2002). Interactional feedback and children's L2 development. *System*, 30, 459-477.
- Mackey, A., Oliver, R. and Leeman, J. (2003). Interactional input and the incorporation of feedback: An exploration of NS-NNS and NNS-NNS adult and child dyads. *Language Learning*, 53, (1), 35-66.
- Mackey, A. and Polio, C. (Eds.) (2009). *Multiple Perspectives on Interaction: Second Language Research in Honor of Susan M. Gass*. Taylor & Francis/Routledge.
- Mackey, A. and Silver, R.E. (2005). Interactional tasks and English L2 learning by immigrant children in Singapore. *System*, 33, 239-260.
- MacWhinney, B. (2000). *The CHILDES Project: Tools for Analyzing Talk*. 3rd Edition. Mahwah, NJ: Lawrence Erlbaum Associates.
- Manchón, R. (2014). The distinctive nature of task repetition in writing. implications for theory, research, and pedagogy. *ELIA*, 14, 13-42.
- Martínez Adrián, M. and Gutierrez-Mangado, M.J. (2015). Is CLIL instruction beneficial both in terms of general proficiency and specific areas of grammar? *Journal of Immersion and Content-Based Language Education*, 3, (1), 51-76.
- McAllister, J., Narcy-Combes, M. and Starkey-Perret, R. (2012). Language teachers' perceptions of a task-based learning programme in a French University. In A. Shehadeh & C.A. Coombe (Eds.), *Task-based Language Teaching in Foreign Language Contexts: Research and Implementation* (pp. 313-342). Amsterdam: John Benjamins.

REFERENCES

- McDonough, K. (2005). Identifying the impact of negative feedback and learners' responses on ESL question development. *Studies in Second Language Acquisition*, 27, (1), 79-103.
- McDonough, K. and Mackey, A. (2000). Communicative tasks, conversational interaction and linguistic form: An empirical study of Thai. *Foreign Language Annals*, 33, (1), 82-92.
- McDonough, K. and Mackey, A. (2006). Responses to recasts: Repetitions, primed production and linguistic development. *Language Learning*, 54, (4), 693-720.
- McDonough, K. and Mackey, A. (2008). Syntactic priming and ESL question development. *Studies in Second Language Acquisition*, 30, (1), 31-47.
- Mehisto, P., Marsh, D. and Frigols, M.J. (2008). *Uncovering CLIL*. Macmillan Education.
- Met, M. (1998). Curriculum decision-making in content-based language teaching. In J. Cenoz and F. Genesse (Eds.), *Beyond Bilingualism. Multilingualism and Multilingual Education*. (pp. 35-63). Clevedon: Multilingual Matters.
- Ministerio de Educación, Cultura y Deporte. *Integrated Curriculum for Secondary Education English - Years 1 and 2*. Retrieved the 01/06/15 from: <http://www.educacion.navarra.es/documents/27590/40045/Curr%C3%ADculo+Secundaria/86843b3d-064d-4770-b0bb-6cacba1b2e1b>
- Muñoz, C. (2006). The effects of age on foreign language learning: The BAF Project. In C. Muñoz (Ed.), *Age and the Rate of Foreign Language Learning* (pp. 1-40). Clevedon: Multilingual Matters.
- Muñoz, C. (2007a). Cross-linguistic influence and language switches in L4 oral production. *Vigo International Journal of Applied Linguistics*, 4, 73-94.
- Muñoz, C. (2007b). Age-related differences and second language learning practice. In R. DeKeyser (Ed.), *Practice in a Second Language. Perspectives from Applied Linguistics and Cognitive Psychology* (pp. 229-255). Cambridge: Cambridge University Press.
- Muñoz, C. (2014). Contrasting effects of starting age and input on the oral performance of foreign language learners. *Applied Linguistics*, 35, 463-482.
- Muñoz, C. and Singleton, D. (2011) A critical review of age-related research on L2 ultimate attainment. *Language Teaching*, 44, (1), 1-35.
- Muranoi, H. (2007). Output practice in the L2 classroom. In R. DeKeyser (Ed.), *Practice in a Second Language: Perspectives from Applied Linguistics and Cognitive Psychology* (pp. 51-84). Cambridge: Cambridge University Press.
- Murphy, V.A. (2014). *Second Language Learning in the Early School Years. Trends and Contexts*. Oxford University Press, USA.

- Neokleous, G. (2016). Closing the gap: Student attitudes toward first language use in monolingual EFL classrooms. *TESOL Journal*. doi: 10.1002/tesj.272
- Nicholas, H. and Lightbown, P.M. (2008). Defining child second language acquisition, defining roles for L2 instruction. In J. Philp, R. Oliver, & A. Mackey (Eds.), *Second Language Acquisition and the Younger Learner: Child's Play?* (pp. 27-51). Amsterdam: John Benjamins.
- Nikolov, M. (2009). The age factor in context. In M. Nikolov (Ed.), *The Age Factor and Early Language Learning* (pp. 1-38). Berlin: Mouton de Gruyter.
- Nikolov, M. and Mihaljevic-Djigunovic, J. (2006). Recent research on age, second language acquisition, and early foreign language learning. *Annual Review of Applied Linguistics*, 26, 234-260.
- Nikolov, M. and Mihaljevic-Djigunović, J. (2011). All shades of every color: An overview of early teaching and learning of foreign languages. *Annual Review of Applied Linguistics*, 31, 95-119.
- Nikula, T. (2007). The IRF pattern and space for interaction: Observations on EFL and CLIL classrooms. In C. Dalton-Puffer & U. Smit (Eds.), *Empirical Perspectives on CLIL Classroom Discourse* (pp. 179-204). Frankfurt: Peter Lang.
- Nuevo, A. (2006). *Task Complexity and Interaction: L2 Learning Opportunities and Development*. Unpublished Doctoral Dissertation, Georgetown University, Washington.
- Nunan, D. (1991) Communicative tasks and the language curriculum. *TESOL Quarterly*, 25, 279-295.
- Nunan, D. (2004). *Task-Based Language Teaching*. Cambridge: Cambridge University Press.
- Oliver, R. (1998). Negotiation of meaning in child interactions. *Modern Language Journal*, 82, (3), 372-386.
- Oliver, R. (2002). The patterns of negotiation for meaning in child interactions. *The Modern Language Journal*, 86, (1), 97-111.
- Oliver, R. (2009): "How young is too young? Investigating negotiation of meaning and feedback in children aged five to seven years". In A. Mackey & C. Polio (Eds.), *Multiple Perspectives on Interaction. Second Language Research in Honor of Susan M. Gass* (pp. 135-156). New York, Routledge.
- Oliver, R. and Mackey, A. (2003). Interactional context and feedback in child ESL classrooms. *The Modern Language Journal*, 87, (4), 519-533.
- Oliver, R., Philp, J. and Mackey, A. (2008). The impact of teacher input, guidance and feedback on ESL children's task-based interactions. In J. Philp, R. Oliver & A. Mackey

REFERENCES

- (Eds.), *Second Language Acquisition and the Young Learner: Child's Play?* (pp.131-147). Amsterdam: John Benjamins.
- Patanasorn, C. (2010). *Effects of procedural content and task repetition on accuracy and fluency in an EFL context*. Unpublished Doctoral Dissertation, Northern Arizona University.
- Pérez-Cañado, M.L. (2012). CLIL research in Europe: Past, present and future. *International Journal of Bilingual Education and Bilingualism*, 15, (3), 315-341.
- Philp, J. and Duchesne, S. (2008). When the gate opens: the interaction between social and linguistic goals in child second language development. In J. Philp, R. Oliver & A. Mackey (Eds.), *Second Language Acquisition and the Younger Learner: Child's Play?* (pp. 83-104). Amsterdam: John Benjamins.
- Philp, J., Oliver, R. and Mackey, A. (2006). The impact of planning time on children's task-based interactions. *System*, 34, (4), 547-565.
- Philp, J., Mackey, A. and Oliver (2008). Introduction. In J. Philp, R. Oliver and A. Mackey (Eds.), *Second Language Acquisition and Younger Learner: Child's Play?* (pp. 1-23). Amsterdam: John Benjamins.
- Philp, J. and Tognini, R. (2009). Language acquisition in foreign language contexts and the differential benefits of interaction. *International Review of Applied Linguistics*, 47, 245-66.
- Pica, T. (1994). Research on negotiation: What does it reveal about second language learning conditions, processes, and outcomes? *Language Learning*, 44, 493-527.
- Pica, T. (2002). Subject matter content: How does it assist the interactional and linguistic needs of classroom language learners? *Modern Language Journal*, 86, 1-19.
- Pica, T. (2005). Classroom learning, teaching, and research: A task-based perspective. *The Modern Language Journal*, 89, (3), 339-352.
- Pica, T. (2012). Foreword. In A. Shehadeh & C.A. Coombe (Eds.), *Task-based Language Teaching in Foreign Language Contexts: Research and Implementation* (pp. xv-xx). Amsterdam: John Benjamins.
- Pica, T. (2013). From input, output and comprehension to negotiation, evidence and attention: An overview of theory and research on learner interaction and SLA. In M.P. García Mayo, M.J. Gutierrez-Mangado & M. Martínez Adrián (Eds.), *Contemporary Approaches to Second Language Acquisition* (pp.49-70). Amsterdam: John Benjamins.
- Pica, T. and Doughty, C. (1985). Non-native speaker interaction in the ESL classroom. In S.M. Gass & C. Madden (Eds.), *Input in Second Language Acquisition* (pp. 115-132). Rowley, MA: Newbury House.

Pica, T., Kanagy, R. and Falodun, J. (1993). Choosing and using communication tasks for second language instruction and research. In G. Crookes & S.M. Gass (Eds.), *Tasks and Language Learning: Integrating Theory and Practice* (pp. 9-34). Clevedon: Multilingual Matters.

Pica, T., Kang, H. and Sauro, S. (2006). Information gap tasks: Their multiple roles and contributions to interaction research methodology. *Studies in Second Language Acquisition*, 28, 301-338.

Pica, T., Lincoln-Porter, F., Paninos, D. and Linnell, J. (1996). Language learners' interaction: How does it address the input, output and feedback needs of L2 learners? *TESOL Quarterly*, 30, 59-84.

Pienemann, M. and Johnston, M. (1986). An acquisition based procedure for second language assessment (ESL). *Australian Review of Applied Linguistics*, 9, (1), 92-122.

Pinter, A. (2006). Verbal evidence of task related strategies: Child versus adult interactions. *System*, 24, 615-630.

Pinter, A. (2007). Some benefits of peer-peer interaction: 10 year-old children practicing with a communicative task. *Language Teaching Research*, 11, (2), 189-207.

Pinter, A. (2011). *Children Learning Second Languages*. London: Palgrave Macmillan.

Pinter, A. (2017). *Teaching Young Language Learners* (Second Edition). Oxford: Oxford University Press.

Pladevall-Ballester, E. (2015). Exploring primary school CLIL perceptions in Catalonia: students', teachers' and parents' opinions and expectations. *International Journal of Bilingual Education and Bilingualism*, 18, (1), 45-59.

Pladevall-Ballester, E. and Vraciu, A. (2017). Exploring early EFL: L1 use in oral narratives by CLIL and non-CLIL primary school learners. In M.P. García Mayo (Ed.), *Learning Foreign Languages in Primary School: Research Insights* (pp. 124-148). Bristol: Multilingual Matters.

Révész, A. (2011). Task complexity, focus on L2 constructions, and individual differences: A classroom based study. *The Modern Language Journal*, 95, (1), 162-181.

Robinson, P. (2001). Task complexity, task difficulty, and task production: Exploring interactions in a componential framework. *Applied Linguistics*, 22, 27-57.

Robinson, P. (2003). The Cognition Hypothesis, task design and adult task-based language learning. *Second Language Studies*, 21, (2), 45-107.
<http://www.hawaii.edu/sls/wp-content/uploads/2014/09/Robinson.pdf>

REFERENCES

- Robinson, P. (2005). Cognitive complexity and task sequencing: Studies in a Componential Framework for second language task design. *International Review of Applied Linguistics in Language Teaching (IRAL)*, 43, (1), 1-32.
- Robinson, P. (2007). Criteria for classifying and sequencing pedagogic tasks. In M.P. García Mayo (Ed.), *Investigating Tasks in Formal Second Language Learning. Multilingual Matters* (in SLA series) (pp. 7-27). Clevedon, Avon: Multilingual Matters.
- Robinson, P. (2011). Task-based language learning: A review of issues. *Language Learning*, 61, (1), 1-36.
- Ross-Feldman, L. (2007). Interaction in the L2 classroom: Does gender influence learning opportunities? In A. Mackey (Ed.), *Conversational Interaction in Second Language Acquisition: A Collection of Empirical Studies* (pp. 53-77). Oxford: Oxford University Press.
- Ruiz de Zarobe, Y. (2008). CLIL and foreign language learning: A longitudinal study in the Basque Country. *International CLIL Research Journal*, 1, (1), 60-73.
- Ruiz de Zarobe, Y. (2011). Which language competencies benefit from CLIL? An insight into Applied Linguistics Research. In Y. Ruiz de Zarobe, J.M. Sierra, & F. Gallardo del Puerto, F. (Eds.), *Content and Foreign Language Integrated Learning: Contributions to Multilingualism in European Contexts* (pp. 129-154). Bern: Peter Lang.
- Ruiz de Zarobe, Y. (2015). The effects of implementing CLIL in education. In M. Juan-Garau and J. Salazar-Noguera (Eds.), *Content-based Language Learning in Multilingual Educational Environments* (pp. 51-68). Educational Linguistics 23. Switzerland: Springer International Publishing. doi: 10.1007/978-3-319-11496-5_4
- Ruiz de Zarobe, Y. and Lasagabaster, D. (2010). CLIL in a Bilingual Community: The Basque Autonomous Country. In D. Lasagabaster and Y. Ruiz de Zarobe (Eds.), *CLIL in Spain. Implementation, Results and Teacher Training* (pp.12-29). Newcastle upon Tyne: Cambridge Scholars Publishing.
- Russell, J. and Spada, N. (2006). The effectiveness of corrective feedback for second language acquisition: A meta-analysis of the research. In J. Norris & L. Ortega (Eds.), *Synthesizing Research on Language Learning and Teaching* (pp. 131-164). Amsterdam: Benjamins.
- Saeedi, M. and Rahimi Kazerooni, S. (2014). The influence of task repetition and task structure on EFL learners' oral narrative retellings. *Innovation in Language Learning and Teaching. Innovation in Language Learning and Teaching*, 8, (2), 116-131.
- Sample, E. and Michel, M. (2014). An exploratory study into trade-off effects of complexity, accuracy and fluency in young learners' oral task repetition. *TESL Canada Journal*, 31(Special Issue 8), 23-46.

- Samuda, V. (2001). Guiding relationships between form and meaning during task performance: the role of the teacher. In M. Bygate, P. Skehan & M. Swain (Eds.). *Researching Pedagogic Tasks: Second Language Learning, Teaching and Testing* (pp. 119-134). Harlow: Pearson Education.
- Samuda, V. and Bygate, M. (2008). *Tasks in Second Language Learning*. (Research and practice in applied linguistics). Basingstoke: Palgrave Macmillan.
- San Isidro, X. (2010). An insight into Galician CLIL. In D. Lasagabaster & Y. Ruiz de Zarobe (Eds.), *CLIL in Spain: Implementation, Results and Teacher Training* (pp. 55-78). Newcastle upon Tyne: Cambridge Scholars Publishing.
- San Isidro, X. (2011). Análisis del modelo CLIL gallego. In C. Escobar Urmeneta, N. Evnitskaya, E. Moore & A. Patiño (Eds.), *AICLE/CLIL/EMILE. Educació Plurilingüe: Experiències, Research and Politiques* (pp. 327-332). Servei de Publicacions, Universitat Autònoma de Barcelona.
- Sato, M. (2016). Interaction mindsets, interactional behaviors, and L2 development: An affective-social-cognitive model. *Language Learning*. doi:10.1111/lang.12214
- Sato, M. and Lyster, R. (2007). Modified output of Japanese EFL learners: Variable effects of interlocutor vs. feedback types. In A. Mackey (Ed.), *Conversational Interaction in Second Language Acquisition: A Series of Empirical Studies* (pp. 123-142). Oxford: Oxford University Press.
- Schiffrin, D. (2006). Discourse marker research and theory: Revisiting *and*. In K. Fischer (ed.), *Approaches to Discourse Particles* (pp. 315-338). Oxford: Elsevier.
- Schmidt, R. (1990). The role of consciousness in second language learning. *Applied Linguistics*, 11, 129-158.
- Schmidt, R. (1995). Consciousness and foreign language learning: A tutorial on the role of attention and awareness in learning. In R. Schmidt (Ed.), *Attention and Awareness in Foreign Language Learning* (pp. 1-64). Manoa, HI: University of Hawaii.
- Schmidt, R. (2001). Attention. In P. Robinson (Ed.), *Cognition and Second Language Instruction* (pp. 3-32). Cambridge: Cambridge University Press.
- Schmidt, R. and Frota, S. (1986). Developing basic conversational ability in a second language. A case study of an adult learner of Portuguese. In R.R. Day (Ed.), *Talking to Learn: Conversation in Second Language Acquisition* (237-326). Rowley, MA: Newbury House.
- Seedhouse, P. (2005). "Task" as research construct. *Language Learning*, 55, (3), 533-570. doi: 10.1111/j.0023-8333.2005.00314.x
- Seikkula-Leino, J. (2007). CLIL learning: Achievement levels and affective factors. *Language and Education*, 21, 328-341.

REFERENCES

- Shehadeh, A. (2012). Broadening the perspective of task-based language teaching scholarship: The contribution of research in foreign language contexts. In A. Shehadeh & C.A. Coombe (Eds.), *Task-based Language Teaching in Foreign Language Contexts: Research and Implementation* (pp. 1-20). Amsterdam: John Benjamins.
- Shehadeh, A. and Coombe, C.A. (Eds.) (2012). *Task-based Language Teaching in Foreign Language Contexts: Research and Implementation*. Amsterdam: John Benjamins.
- Singleton, D. (2005). The critical period hypothesis: A coat of many colours. *International Review of Applied Linguistics*, 43, (4) 269-285.
- Singleton, D. and Ryan, L. (Eds.) (2004). *Language Acquisition: The Age Factor*. Clevedon: Multilingual Matters.
- Skehan, P. (1996). A framework for the implementation of task-based instruction. *Applied Linguistics*, 17, 38-62.
- Skehan, P. (1998). *A Cognitive Approach to Language Learning*. Oxford: Oxford University Press.
- Skehan, P. (2001). Task and language performance assessment. In M. Bygate, P. Skehan & M. Swain. (Eds.), *Researching Pedagogic Tasks: Second Language Learning, Teaching and Testing* (pp. 167-185). London: Longman.
- Skehan, P. (2003). Task-based instruction. *Language Teaching*, 36, 1-14.
- Skehan, P. (2009). Modelling second language performance: Integrating complexity, accuracy, fluency, and lexis. *Applied Linguistics*, 30, (4), 510-532.
- Skehan, P. and Foster, P. (2001). Cognition and tasks. In P. Robinson (Ed.), *Cognition and Second Language Instruction* (pp. 183-205). Cambridge, UK: Cambridge University.
- Skehan, P. and Foster, P. (2012). Complexity, accuracy, fluency and lexis in task-based performance: a synthesis of the Ealing research. In A. Housen, F. Kuiken and I. Vedder (Eds.), *Dimensions of L2 Performance and Proficiency: Complexity, Accuracy and Fluency* (pp. 199-220). Amsterdam: John Benjamins.
- Skehan, P., Bei, X., Li, Q. and Wang, Z. (2012). The task is not enough: Processing approaches to task-based performance. *Language Teaching Research*, 16, (2), 170-187.
- Smith, B. (2009). Revealing the nature of SCMC interaction. In A. Mackey & C. Polio (Eds.), *Multiple Perspectives on Interaction in SLA* (pp. 197-225). Mahwah, NJ: Lawrence Erlbaum Associates.
- Spada, N. (2011). Beyond form-focused instruction: Reflections on past, present and future research. *Language Teaching Research*, 44, 225-236.

- Spada, N. and Lightbown, P. M. (2008). Interaction research in second/foreign language classrooms. In C. Polio & A. Mackey (Eds.), *Multiple Perspectives on Interaction: Second Language Research in Honor of Susan M. Gass* (pp. 157-175). New York: Taylor and Francis.
- Storch, N. (2016). Collaborative writing. In R. Manchón & P.K. Matsuda (Eds.), *Handbook of Second and Foreign Language Writing* (pp. 387-406). De Gruyter Mouton.
- Storch, N. and Aldosari, A. (2010). Learners' use of first language (Arabic) in pair work in an EFL class. *Language Teaching Research*, 14, (4), 355-375.
- Storch, N. and Wigglesworth, G. (2003). Is there a role for the use of the L1 in an L2 setting? *TESOL Quarterly*, 37, (4), 760-770.
- Swain, M. (2005). Legislation by hypothesis: The case of task-based instruction. *Applied Linguistics*, 26, (3), 376-401.
- Swain, M. (1985). Communicative competence: Some roles of comprehensible input and comprehensible output in its development. In S. Gass & C. Madden (Eds.), *Input in Second Language Acquisition* (pp. 235-252). New York: Newbury House.
- Swain, M. (1993). The output hypothesis: Just speaking and writing aren't enough. *The Canadian Modern Language Review*, 50, 158-164.
- Swain, M. (2005). The output hypothesis: Theory and research. In E. Hinkel (Ed.), *Handbook of Research in Second Language Teaching and Learning* (pp. 471-483). Mahwah, NJ: Lawrence Erlbaum.
- Swain, M. and Lapkin, S. (1995). Problems in output and the cognitive processes they generate: A step towards second language learning. *Applied Linguistics*, 16, 371-391.
- Swain, M. and Lapkin, S. (1998). Interaction and second language learning: Two adolescent French immersion students working together. *Modern Language Journal*, 82, 320-337.
- Swain, M. and Lapkin, S. (2000). Task-based second language learning: The uses of the first language. *Language Teaching Research*, 4, 251-274.
- Swain, M. and Lapkin, S. (2001). Focus on form through collaborative dialogue: Exploring task effects. In M. Bygate, P. Skehan & M. Swain (Eds.), *Researching Pedagogic Tasks: Second Language Learning, Teaching and Testing* (pp. 99-118). London, UK: Pearson Education.
- Sylvén, L.K. and Ohlander, S. (2014). The CLISS Project: receptive vocabulary in CLIL versus non-CLIL groups. *Moderna Språk*, 2, 80-114.

REFERENCES

- Tognini, R. and Oliver, R. (2012). L1 use in primary and secondary foreign language classrooms and its contribution to learning. In E. Alcon & M.P. Safont (Eds.), *Language Learners' Discourse in Instructional Settings* (pp. 53-78). Amsterdam: Rodopi.
- Valeo, A. (2013). Language awareness in a content-based language programme. *Language Awareness*, 22, (2), 126-145.
- Van den Branden, K. (1997). Effects of negotiation on language learners' output. *Language Learning*, 47, 589-636.
- Van den Branden, K. (2006). *Task-based Language Education: From Theory to Practice*. Cambridge: Cambridge University Press.
- Van den Branden, K., Bygate, M. and Norris, J. (Eds.) (2009). *Task-based Language Teaching: A Reader*. Amsterdam: John Benjamins.
- VanPatten, B. and Williams, J. (2007) (Eds.). *Theories in Second Language Acquisition. An Introduction*. Mahwah, NJ: Lawrence Erlbaum Associates.
- Varonis, E. and Gass, S. (1985). Non-native/non-native conversations: A model for negotiation of meaning. *Applied Linguistics*, 6, (1), 71-90.
- Wang, Z. (2009). Strategic planning, on-line planning, and repetition. In P. Skehan, X. Bei, Q. Li & Z. Wang (2012), *The task is not enough: Processing approaches to task-based performance*. *Language Teaching Research*, 16, (2), 170-187.
- Willis, J. (1996). *A Framework for Task-based Learning*. Essex, UK, Longman.
- Wolff, D. (2007). CLIL: Bridging the gap between school and working life. In D. Marsh and D. Wolff, *Diverse Contexts - Converging Goals. CLIL in Europe* (pp. 15-25). Frankfurt: Peter Lang.
- Wong, W. (2001). Modality and attention to meaning and form in the input. *Studies in Second Language Acquisition*, 23, (3), 345-368.
- Xanthou, M. (2011). The impact of CLIL on L2 vocabulary development and content knowledge. *English Teaching: Practice and Critique*, 10, (4), 116-126.

APPENDIXES

Appendix A

HOJA DE INFORMACIÓN Y CONSENTIMIENTO INFORMADO

El presente formulario tiene como objeto proporcionarle la información necesaria para que decida libre y voluntariamente la participación de su hijo/a en esta prueba. Es necesario que lea detenidamente la siguiente información y que pregunte si tiene alguna duda al respecto.

CONTACTO:

Investigadora principal: María del Pilar García Mayo

Dirección: Paseo de la Universidad 5

Centro: Facultad de Letras

Teléfono: 945 013036

Correo electrónico: mariapilar.garciamayo@ehu.eus

DATOS RELATIVOS AL PROYECTO:

- Título del proyecto: LA INTERACCIÓN ORAL ENTRE JÓVENES APRENDICES DE INGLÉS COMO LENGUA EXTRANJERA: ESTRATEGIAS DE NEGOCIACIÓN Y RETROALIMENTACIÓN EN TAREAS COMUNICATIVAS Y SU IMPACTO EN EL APRENDIZAJE.
- Financiado por el Ministerio de Economía y Competitividad (Ref. FFI2012-32212)

DESCRIPCIÓN DEL PROCEDIMIENTO:

El principal objetivo de la grabación es determinar si los alumnos de Educación Primaria que aprenden el inglés como lengua extranjera son capaces de interactuar y negociar el significado y la forma lingüística con sus compañeros. Un segundo objetivo será analizar los efectos de la retroalimentación interaccional (*interactional feedback*) y el uso de la primera lengua (L1) en el desarrollo de la interlengua de estos aprendices. El análisis de estas estrategias (negociación, retroalimentación interaccional y uso de la L1) nos ayudará a determinar su posible impacto en el aprendizaje de la lengua extranjera y a proporcionar guías pedagógicas para el profesorado de esta etapa educativa.

DESCRIPCIÓN DEL PROCEDIMIENTO:

- Tipo de procedimiento: el participante completará una tarea de comunicación oral con un compañero de clase.
- La tarea será repetida tres veces, una vez a la semana.
- Datos personales anónimos: los datos personales serán tratados de forma totalmente anónima así como los resultados de todas las pruebas.
- **Número de intervenciones:** la recogida de datos se realizará en tres sesiones de 10 minutos de duración aproximada en un aula del centro de educación primaria durante las horas de clase con instrucción en inglés bajo la supervisión de la investigadora.
- **Descripción de riesgos:** no existe ningún riesgo.

DERECHOS DEL PARTICIPANTE:

- La participación en este estudio es **voluntaria** y podrá dejar de participar en cualquier momento, sin que ello suponga ningún perjuicio, comunicando la intención de abandono, a la IP mediante correo electrónico será suficiente.
- Si usted colabora en este proyecto, una vez haya finalizado, usted tendrá a su **disposición** toda la información relativa a los resultados obtenidos en el mismo, respetando la confidencialidad de los participantes. Puede obtener los datos poniéndose en contacto con el IP.

Deseo ser informado NO deseo ser informado

- Las pruebas pueden incluir también la recogida de datos de vídeo y voz (grabaciones):
 - Doy el consentimiento para la grabación de vídeo y voz
 - NO doy el consentimiento para la grabación de vídeo y voz
- Los datos personales que nos ha facilitado para este proyecto de investigación serán tratados con absoluta **confidencialidad** de acuerdo con la Ley de Protección de Datos.
- Los datos recogidos se utilizarán solamente para fines de investigación y únicamente con fines de mejora de la práctica educativa y de los aprendizajes de los propios alumnos.

IDENTIFICACION DE LA PERSONA QUE PRESTA EL CONSENTIMIENTO

Yo (nombre y apellidos) con D.N.I., madre/padre/tutor de

EXPONGO

que he sido debidamente **INFORMADO/A** por el investigador con D.N.I. donde he recibido la información necesaria sobre la naturaleza y propósitos del procedimiento del estudio, habiendo tenido ocasión de aclarar las dudas que me han surgido.

MANIFIESTO

que he entendido que este consentimiento puede ser revocado por mí en cualquier momento. Asimismo he entendido las explicaciones y aclaraciones recibidas sobre el estudio y **OTORGO MI CONSENTIMIENTO** para la participación de mi hijo/a en este estudio.

(Fecha) (Firma del padre/madre/tutor)

(Fecha) (Firma del padre/madre/tutor)

Appendix B

Picture Placement

General instructions:

1. Students work in pairs. They should not see each other while performing the task. The reason is that we want them to rely exclusively on oral English. They could sit with a table/piece of cardboard/folding screen between them.
2. Make sure you have the material: Two similar posters and 6 pictures of children (we will refer to these pictures as **a**, **b**, **c**, **d**, **e**, **f**).

Poster:



Six pictures of children:



3. Distribute the material:

APPENDIXES

Student A needs: One poster with pictures **a** and **b** placed on it and pictures **c**, **d**, **e** and **f** outside the poster.

Poster for student A:



Pictures of children for Student A:



Student B needs: One poster with pictures **c** and **d** placed on it and pictures **a**, **b**, **e** and **f** outside the poster.

Poster for student B:



Pictures of children for Student B:



4. Start recording (camera and digital recorder at the same time):
 - a. First of all say your name, the date, and the name of the students.
 - b. Second, record the two posters:
 - The one you will use for student A.
 - The one you will use for student B.When you record them say out loud:
 - Poster for student A + the name of student A
 - Poster for student B + the name of student B
 - c. Now, focus on the students. Ask them to say their names. Make sure that you can see both students on the screen.

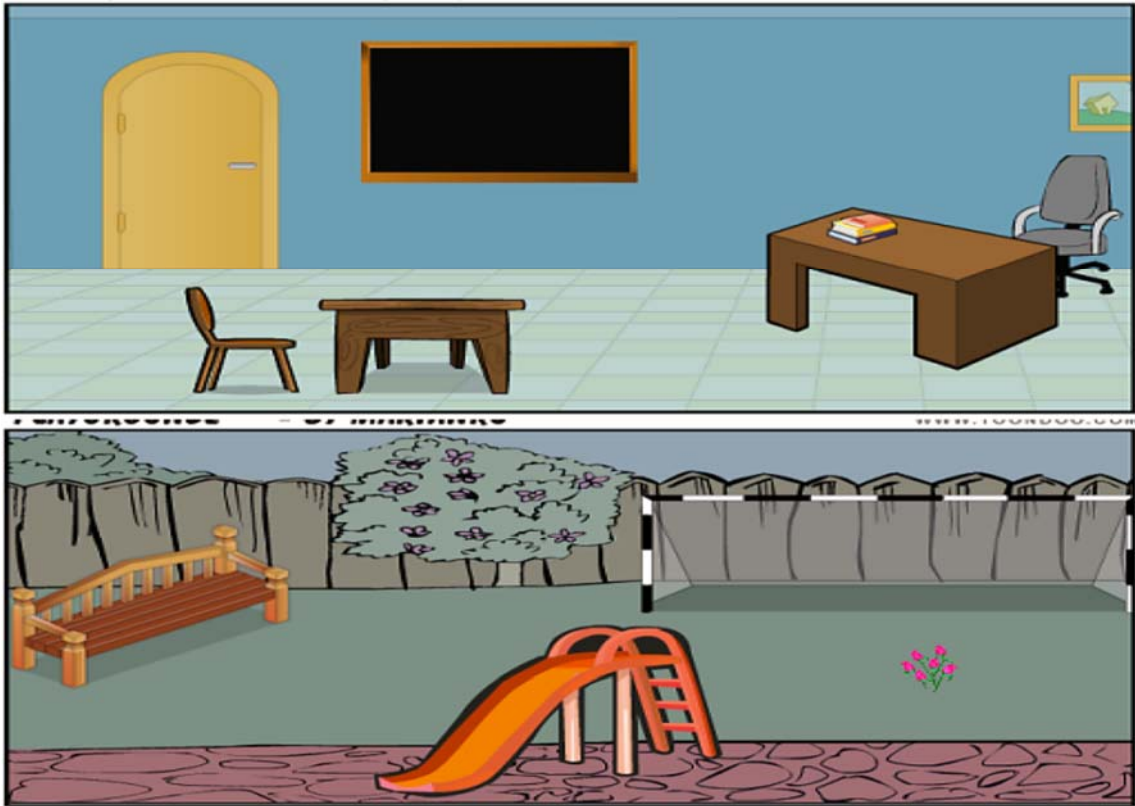
5. Explain the game: The objective is that the students manage to complete their posters so that both have the same children and in the same positions. Don't tell them the number of children on every poster or any information.

APPENDIXES

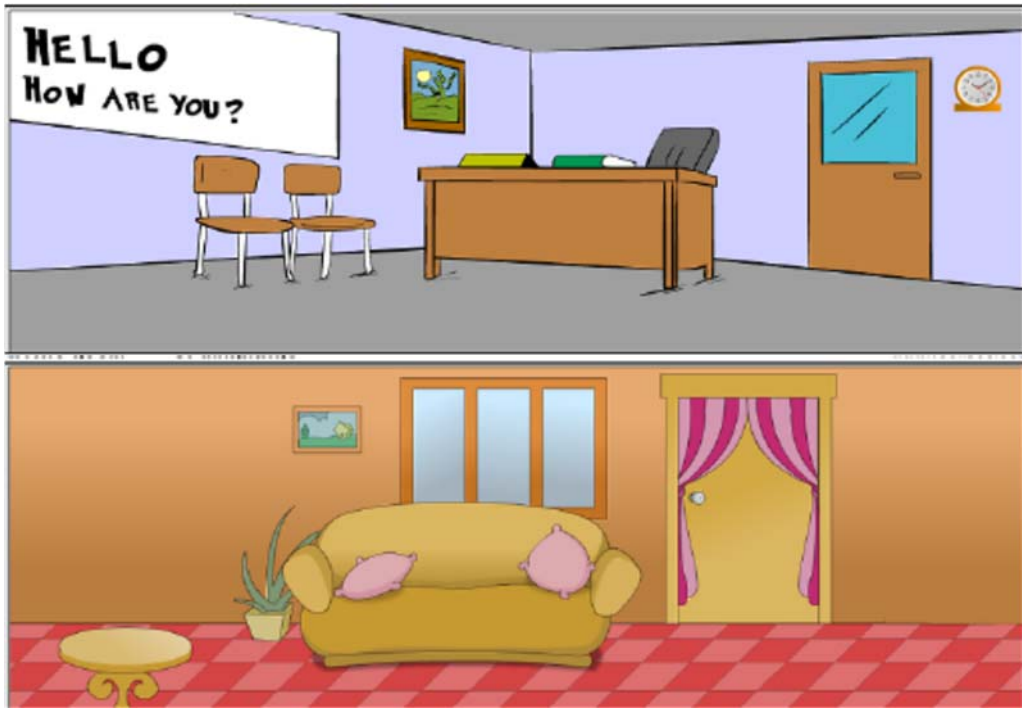
6. Emphasize that it is **team work** and they have to help each other. Try to motivate them so that they do their best. You can say things like “it is difficult but I think you can do it” or “I bet you won’t be able to do it”.
7. Allow them to speak freely. Interrupt only if there is no interaction. In that case, help them with questions such as the following:
 - a. Why don’t you ask your partner: *How many students do you have in the park? How many students do you have in the classroom? Do you have any girls? Do you have a boy with a ball/ a book/ laughing...? Do you have a student next to the tree/ bin/...?*
 - b. Why don’t you describe what you see? *I can see a park, I can see a classroom, I can see two children in the classroom....*
8. Ask them to tell you when they think they have finished (regardless of whether they have succeeded). Compare both posters together and ask them to comment on their result.
9. Finally, ask them briefly the following two questions:
 - a. Did you find the game difficult? Why?
 - b. Would you like to play again?

Appendix C

Task 1



Task 2



Task 3



Table 1 Results of the within- and between-subjects comparisons of the NoM strategies to confirm successful communication.

Differences of Group*time Least Squares Means Adjustment for Multiple Comparisons: Tukey-Kramer						
	Estimate	SD	DF	t Value	Pr > t 	Adj P
Group 3 * Time 1-2	-0.1431	0.2505	196	-0.57	0.5684	0.9928
Group 3 * Time 2-3	-0.2624	0.2273	196	-1.15	0.2497	0.8577
Group 3 * Time 1-3	-0.4055	0.2676	196	-1.52	0.1314	0.6550
Group 5 * Time 1-2	-0.02083	0.1349	196	-0.15	0.8774	1.0000
Group 5 * Time 2-3	0.04211	0.1356	196	0.31	0.7565	0.9996
Group 5 * Time 1-3	0.02128	0.1541	196	0.14	0.8903	1.0000
Groups 3-5 * Time 1	-1.2784	0.3103	196	-4.12	<.0001	0.0008*
Groups 3-5 * Time 1-2	-1.2993	0.3099	196	-4.19	<.0001	0.0006*
Groups 3-5 * Time 1-3	-1.2571	0.3108	196	-4.05	<.0001	0.0010*
Groups 3-5 * Time 2	-1.1562	0.2997	196	-3.86	0.0002	0.0021*
Groups 3-5 * Time 2-3	-1.1140	0.3006	196	-3.71	0.0003	0.0037*
Groups 3-5 * Time 3	-0.8517	0.2847	196	-2.99	0.0031	0.0365*

Table 2 Results of the within- and between-subjects comparisons of the NoM strategies to prevent communication breakdowns.

Differences of Group*time Least Squares Means Adjustment for Multiple Comparisons: Tukey-Kramer						
	Estimate	SD	DF	t Value	Pr > t 	Adj P
Group 3 * Time 1-2	0.2076	0.2920	196	0.71	0.4778	0.9804
Group 3 * Time 2-3	0.6190	0.3688	196	1.68	0.0949	0.5478
Group 3 * Time 1-3	0.8267	0.3941	196	2.10	0.0372	0.2928
Group 5 * Time 1-2	-0.2877	0.2440	196	-1.18	0.2398	0.8464
Group 5 * Time 2-3	1.7918	0.4408	196	4.06	<.0001	0.0010*
Group 5 * Time 1-3	1.5041	0.4828	196	3.12	0.0021	0.0254*
Groups 3-5 * Time 1	0.02492	0.4323	196	0.06	0.9541	1.0000
Groups 3-5 * Time 1-2	-0.2628	0.4193	196	-0.63	0.5316	0.9889
Groups 3-5 * Time 1-3	1.5290	0.5846	196	2.62	0.0096	0.0983
Groups 3-5 * Time 2	-0.4704	0.4328	196	-1.09	0.2784	0.8861
Groups 3-5 * Time 2-3	1.3214	0.5943	196	2.22	0.0273	0.2318
Groups 3-5 * Time 3	0.7023	0.6369	196	1.10	0.2715	0.8798

Table 3 Results of the within- and between-subjects comparisons of the NoM strategies to repair communication breakdowns.

Differences of Group*time Least Squares Means Adjustment for Multiple Comparisons: Tukey-Kramer						
	Estimate	SD	DF	t Value	Pr > t 	Adj P
Group 3 * Time 1-2	0.2231	0.1613	196	1.38	0.1680	0.7367
Group 3 * Time 2-3	0.2036	0.1793	196	1.14	0.2576	0.8660
Group 3 * Time 1-3	0.4267	0.1811	196	2.36	0.0195	0.1773
Group 5 * Time 1-2	0.09382	0.1854	196	0.51	0.6133	0.9959
Group 5 * Time 2-3	0.4220	0.2134	196	1.98	0.0494	0.3591
Group 5 * Time 1-3	0.5158	0.2217	196	2.33	0.0210	0.1885
Groups 3-5 * Time 1	0.3501	0.2306	196	1.52	0.1306	0.6529
Groups 3-5 * Time 1-2	0.4439	0.2345	196	1.89	0.0599	0.4097
Groups 3-5 * Time 1-3	0.8659	0.2564	196	3.38	0.0009	0.0113*
Groups 3-5 * Time 2	0.2208	0.2414	196	0.91	0.3616	0.9425
Groups 3-5 * Time 2-3	0.6427	0.2627	196	2.45	0.0153	0.1458
Groups 3-5 * Time 3	0.4391	0.2697	196	1.63	0.1051	0.5809

Table 4 Results of the within- and between-subjects comparisons of the NoM strategies to focus on form.

Differences of Group*time Least Squares Means Adjustment for Multiple Comparisons: Tukey-Kramer						
	Estimate	SD	DF	t Value	Pr > t 	Adj P
Group 3 * Time 1-2	0.6931	0.8707	196	0.80	0.4270	0.9679
Group3 * Time 2-3	-0.4055	0.9188	196	-0.44	0.6595	0.9979
Group 3 * Time 1-3	0.2877	0.7468	196	0.39	0.7005	0.9989
Group 5 * Time 1-2	0.6931	1.2314	196	0.56	0.5742	0.9933
Group 5 * Time 2-3	-1.0986	1.1586	196	-0.95	0.3442	0.9333
Group 5 * Time 1-3	-0.4055	0.8926	196	-0.45	0.6502	0.9975
Groups 3-5 * Time 1	0.6931	0.8504	196	0.82	0.4160	0.9645
Groups 3-5 * Time 1-2	1.3863	1.0967	196	1.26	0.2077	0.8041
Groups 3-5 * Time 1-3	0.2877	0.7505	196	0.38	0.7019	0.9989
Groups 3-5 * Time 2	0.6931	1.2011	196	0.58	0.5645	0.9924
Groups 3-5 * Time 2-3	-0.4055	0.8962	196	-0.45	0.6514	0.9976
Groups 3-5 * Time 3	-178E-16	0.8020	196	-0.00	1.0000	1.0000

Table 5 Overall changes across TR in the production of words, clauses and AS-units by YLs.

Differences of Time Least Squares Means							
Adjustment for Multiple Comparisons: Tukey-Kramer							
	Effect	Estimate	SD	DF	t Value	Pr > t 	Adj P
Words	Time 1-2	0.009495	0.05348	196	0.18	0.8593	0.009495
	Time 1-3	0.06818	0.05429	196	1.26	0.2106	0.06818
	Time 2-3	0.05869	0.05441	196	1.08	0.2821	0.05869
	Group 3-5	-0.01481	0.1206	38	-0.12	0.9030	-0.01481
Clauses	Time 1-2	0.04379	0.04355	196	1.01	0.3159	0.5742
	Time 1-3	0.06802	0.04386	196	1.55	0.1226	0.2696
	Time 2-3	0.02424	0.04433	196	0.55	0.5852	0.8483
	Group 3-5	0.1189	0.1146	38	1.04	0.3063	0.3063
AS-units	Time 1-2	0.04421	0.04524	196	0.98	0.3297	0.5922
	Time 1-3	0.1157	0.04613	196	2.51	0.0129	0.0344*
	Time 2-3	0.07149	0.04659	196	1.53	0.1266	0.2772
	Group 3-5	0.1783	0.1196	38	1.49	0.1442	0.1442

Table 6 Overall time and group comparisons of the complexity measures.

Differences of Least Squares Means								
	Effect	Estimate	SD	DF	t Value	Pr > t	Adjustment	Adj P
Words/ AS-unit	Time 1-2	-0.00917	0.1571	196	-0.06	0.9535	Tukey- Kramer	0.9981
	Time 1-3	-0.2142	0.1679	196	-1.28	0.2036	Tukey- Kramer	0.4106
	Time 2-3	-0.2050	0.1760	196	-1.16	0.2455	Tukey- Kramer	0.4756
	Groups 3-5	-0.9409	0.3054	38	-3.08	0.0038	Tukey	0.0038*
Clauses/ AS-unit	Time 1-2	0.02219	0.01496	196	1.48	0.1396	Tukey- Kramer	0.3011
	Time 1-3	-0.04898	0.02333	196	-2.10	0.0370	Tukey- Kramer	0.0925
	Time 2-3	-0.07117	0.01997	196	-3.56	0.0005	Tukey- Kramer	0.0013*
	Groups 3-5	-0.07117	0.03217	38	-2.21	0.0330	Tukey	0.0330*
Lexical D	Time 1-2	1.8250	0.7932	191	2.30	0.0225	Tukey- Kramer	0.0581*
	Time 1-3	1.5581	1.0763	191	1.45	0.1493	Tukey- Kramer	0.3187
	Time 2-3	-0.2668	0.9060	191	-0.29	0.7687	Tukey- Kramer	0.9533
	Groups 3-5	-3.5824	1.9039	38	-1.88	0.0676	Tukey- Kramer	0.0676

Table 7 Results of the within- and between-subjects comparisons of the words per AS-unit scores.

Differences of Least Squares Means							
Effect	Estimate	SD	DF	t Value	Pr > t	Adjustment	Adj P
Group 3 * Time 1-2	-0.2878	0.2222	196	-1.30	0.1967	Tukey-Kramer	0.7873
Group 3 * Time 2-3	-0.2706	0.2489	196	-1.09	0.2784	Tukey-Kramer	0.8862
Group 3 * Time 1-3	-0.5584	0.2375	196	-2.35	0.0197	Tukey-Kramer	0.1789
Group 5 * Time 1-2	0.2695	0.2222	196	1.21	0.2267	Tukey-Kramer	0.8302
Group 5 * Time 2-3	-0.1395	0.2489	196	-0.56	0.5759	Tukey-Kramer	0.9934
Group 5 * Time 1-3	0.1300	0.2375	196	0.55	0.5848	Tukey-Kramer	0.9941
Groups 3-5 * Time 1	-1.3562	0.3683	196	-3.68	0.0003	Tukey-Kramer	0.0040*
Groups 3-5 * Time 1-2	-1.0867	0.3509	196	-3.10	0.0022	Tukey-Kramer	0.0268*
Groups 3-5 * Time 1-3	-1.2262	0.3748	196	-3.27	0.0013	Tukey-Kramer	0.0158*
Groups 3-5 * Time 2	-0.7989	0.3326	196	-2.40	0.0172	Tukey-Kramer	0.1606
Groups 3-5 * Time 2-3	-0.9384	0.3578	196	-2.62	0.0094	Tukey-Kramer	0.0966
Groups 3-5 * Time 3	-0.6678	0.3813	196	-1.75	0.0815	Tukey-Kramer	0.4997

Table 8 Results of the within- and between-subjects comparisons of the clauses per AS-unit scores.

Differences of Least Squares Means							
Effect	Estimate	SD	DF	t Value	Pr > t	Adjustment	Adj P
Group 3 * Time 1-2	0.006422	0.02116	196	0.30	0.7618	Tukey-Kramer	0.9997
Group 3 * Time 2-3	-0.09350	0.02824	196	-3.31	0.0011	Tukey-Kramer	0.0139*
Group 3 * Time 1-3	-0.08708	0.03299	196	-2.64	0.0090	Tukey-Kramer	0.0928
Group 5 * Time 1-2	0.03796	0.02116	196	1.79	0.0743	Tukey-Kramer	0.4719
Group 5 * Time 2-3	-0.04884	0.02824	196	-1.73	0.0854	Tukey-Kramer	0.5143
Group 5 * Time 1-3	-0.01088	0.03299	196	-0.33	0.7419	Tukey-Kramer	0.9995
Groups 3-5 * Time 1	-0.1071	0.03864	196	-2.77	0.0061	Tukey-Kramer	0.0665
Groups 3-5 * Time 1-2	-0.06913	0.03551	196	-1.95	0.0530	Tukey-Kramer	0.3771
Groups 3-5 * Time 1-3	-0.1180	0.04261	196	-2.77	0.0062	Tukey-Kramer	0.0669
Group 3-5 * Time 2	-0.07555	0.03207	196	-2.36	0.0195	Tukey-Kramer	0.1774
Groups 3-5 * Time 2-3	-0.1244	0.03979	196	-3.13	0.0020	Tukey-Kramer	0.0246*
Groups 3-5 * Time 3	-0.03089	0.04624	196	-0.67	0.5049	Tukey-Kramer	0.9852

Table 9 Results of the within- and between-subjects comparisons of lexical complexity *D*.

Differences of Least Squares Means							
Effect	Estimate	SD	DF	t Value	Pr > t 	Adjustment	Adj P
Group 3 * Time 1-2	2.1549	1.1343	191	1.90	0.0590	Tukey- Kramer	0.4055
Group 3 * Time 2-3	-0.3026	1.2874	191	-0.24	0.8144	Tukey- Kramer	0.9999
Group 3 * Time 1-3	1.8523	1.5363	191	1.21	0.2294	Tukey- Kramer	0.8337
Group 5 * Time 1-2	1.4950	1.1091	191	1.35	0.1793	Tukey- Kramer	0.7578
Group 5 * Time 2-3	-0.2310	1.2750	191	-0.18	0.8564	Tukey- Kramer	1.0000
Group 5 * Time 1-3	1.2640	1.5078	191	0.84	0.4029	Tukey- Kramer	0.9600
Groups 3-5 * Time 1	-3.1663	2.2700	191	-1.39	0.1647	Tukey- Kramer	0.7301
Groups 3-5 * Time 1-2	-1.6713	2.1911	191	-0.76	0.4465	Tukey- Kramer	0.9733
Groups 3-5 * Time 1-3	-1.9023	2.2344	191	-0.85	0.3956	Tukey- Kramer	0.9573
Groups 3-5 * Time 2	-3.8262	2.0957	191	-1.83	0.0695	Tukey- Kramer	0.4517
Groups 3-5 * Time 2-3	-4.0573	2.1410	191	-1.90	0.0596	Tukey- Kramer	0.4084
Groups 3-5 * Time 3	-3.7546	2.1925	191	-1.71	0.0884	Tukey- Kramer	0.5253

Table 10 Overall time and group comparisons of the accuracy measures.

Differences of Least Squares Means								
	Effect	Estimate	SD	DF	t Value	Pr > t	Adjustment	Adj P
Errors/ AS-unit	Time 1-2	-0.05015	0.03476	196	-1.44	0.1507	Tukey- Kramer	0.3211
	Time 1-3	0.03247	0.03722	196	0.87	0.3840	Tukey- Kramer	0.6583
	Time 2-3	0.08263	0.03693	196	2.24	0.0264	Tukey- Kramer	0.0675
	Groups 3-5	-0.02093	0.05385	38	-0.39	0.6998	Tukey	0.6998
EFAS/ AS-unit	Time 1-2	0.7678	1.8162	196	0.42	0.6729	Tukey- Kramer	0.9062
	Time 1-3	-0.9821	1.9972	196	-0.49	0.6234	Tukey- Kramer	0.8753
	Time 2-3	-1.7499	1.8360	196	-0.95	0.3417	Tukey- Kramer	0.6074
	Groups 3-5	-7.7718	3.4157	38	-2.28	0.0286	Tukey	0.0286 *
Correct verbs/ total number of verbs	Time 1-2	-3.2962	2.7454	196	-1.20	0.2313	Tukey- Kramer	0.4542
	Time 1-3	-2.7547	1.7804	196	-1.55	0.1234	Tukey- Kramer	0.2712
	Time 2-3	0.5415	2.0792	196	0.26	0.7948	Tukey- Kramer	0.9633
	Groups 3-5	2.8450	4.4088	38	0.65	0.5226	Tukey	0.5226

Table 11 *Within- and between-subjects comparisons of the ratio of errors per AS-unit.*

Differences of Least Squares Means							
Effect	Estimate	SD	DF	t Value	Pr > t 	Adjustment	Adj P
Group 3* Time 1-2	-0.1226	0.04916	196	-2.49	0.0134	Tukey- Kramer	0.1308
Group 3 * Time 2-3	0.1040	0.05223	196	1.99	0.0479	Tukey- Kramer	0.3515
Group 3 * Time 1-3	-0.01867	0.05264	196	-0.35	0.7233	Tukey- Kramer	0.9993
Group 5 * Time 1-2	0.02232	0.04916	196	0.45	0.6503	Tukey- Kramer	0.9975
Group 5 * Time 2-3	0.06129	0.05223	196	1.17	0.2420	Tukey- Kramer	0.8489
Group 5 * Time 1-3	0.08362	0.05264	196	1.59	0.1138	Tukey- Kramer	0.6071
Groups 3-5 * Time 1	-0.1033	0.06448	196	-1.60	0.1106	Tukey- Kramer	0.5978
Groups 3-5 * Time 1-2	-0.08101	0.06894	196	-1.18	0.2414	Tukey- Kramer	0.8483
Groups 3-5 * Time 1-3	-0.01972	0.06568	196	-0.30	0.7643	Tukey- Kramer	0.9997
Groups 3-5 * Time 2	0.04162	0.07314	196	0.57	0.5700	Tukey- Kramer	0.9929
Groups 3-5 * Time 2-3	0.1029	0.07007	196	1.47	0.1435	Tukey- Kramer	0.6846
Groups 3-5 * Time 3	-0.00105	0.06686	196	-0.02	0.9874	Tukey- Kramer	1.0000

Table 12 *Within- and between-subjects comparisons of the ratio of EFAS per AS-unit.*

Differences of Least Squares Means							
Effect	Estimate	SD	DF	t Value	Pr > t 	Adjustment	Adj P
Group 3 * Time 1-2	3.0668	2.5685	196	1.19	0.2339	Tukey- Kramer	0.8393
Group 3 * Time 2-3	-2.8323	2.5965	196	-1.09	0.2767	Tukey- Kramer	0.8846
Group 3 * Time 1-3	0.2345	2.8244	196	0.08	0.9339	Tukey- Kramer	1.0000
Group 5 * Time 1-2	-1.5312	2.5685	196	-0.60	0.5518	Tukey- Kramer	0.9912
Group 5 * Time 2-3	-0.6675	2.5965	196	-0.26	0.7974	Tukey- Kramer	0.9998
Group 5 * Time 1-3	-2.1988	2.8244	196	-0.78	0.4372	Tukey- Kramer	0.9709
Groups 3-5 * Time 1	-5.4280	3.9430	196	-1.38	0.1702	Tukey- Kramer	0.7410
Group 3-5 * Time 1-2	-6.9592	4.0399	196	-1.72	0.0865	Tukey- Kramer	0.5186
Group 3-5 * Time 1-3	-7.6267	4.0074	196	-1.90	0.0585	Tukey- Kramer	0.4033
Group 3-5 * Time 2	-10.0260	4.1346	196	-2.42	0.0162	Tukey- Kramer	0.1528
Groups 3-5 * Time 2-3	-10.6936	4.1028	196	-2.61	0.0099	Tukey- Kramer	0.1005
Groups 3-5 * Time 3	-7.8612	4.0707	196	-1.93	0.0549	Tukey- Kramer	0.3863

Table 13 *Within- and between-subjects comparisons of the ratio of correct verbs over the total number of verbs.*

Differences of Least Squares Means							
Effect	Estimate	SD	DF	t Value	Pr > t 	Adjustment	Adj P
Group 3 * Time 1-2	-2.4212	3.8826	196	-0.62	0.5336	Tukey- Kramer	0.9892
Group 3 * Time 2-3	-1.1451	2.9404	196	-0.39	0.6974	Tukey- Kramer	0.9988
Group 3 * Time 1-3	-3.5663	2.5178	196	-1.42	0.1582	Tukey- Kramer	0.7171
Group 5 * Time 1-2	-4.1713	3.8826	196	-1.07	0.2840	Tukey- Kramer	0.8911
Group 5 * Time 2-3	2.2281	2.9404	196	0.76	0.4495	Tukey- Kramer	0.9741
Group 5 * Time 1-3	-1.9432	2.5178	196	-0.77	0.4412	Tukey- Kramer	0.9720
Groups 3-5 * Time 1	2.8874	5.1389	196	0.56	0.5749	Tukey- Kramer	0.9933
Groups 3-5 * Time 1-2	-1.2839	5.2674	196	-0.24	0.8077	Tukey- Kramer	0.9999
Groups 3-5 * Time 1-3	0.9442	4.9633	196	0.19	0.8493	Tukey- Kramer	1.0000
Groups 3-5 * Time 2	1.1372	5.3928	196	0.21	0.8332	Tukey- Kramer	0.9999
Groups 3-5 * Time 2-3	3.3653	5.0962	196	0.66	0.5098	Tukey- Kramer	0.9860
Groups 3-5 * Time 3	4.5104	4.7812	196	0.94	0.3467	Tukey- Kramer	0.9347

Table 14 *Within- and between-subjects comparisons of the ratio of words per minute.*

Differences of Least Squares Means							
Effect	Estimate	SD	DF	t Value	Pr > t 	Adjustment	Adj P
Group 3 * Time 1-2	-3.2894	0.9683	196	-3.40	0.0008	Tukey-Kramer	0.0105*
Group 3 * Time 2-3	-2.4971	1.0660	196	-2.34	0.0202	Tukey-Kramer	0.1823
Group 3 * Time 1-3	-5.7865	1.3097	196	-4.42	<.0001	Tukey-Kramer	0.0002*
Group 5 * Time 1-2	-3.6873	0.9683	196	-3.81	0.0002	Tukey-Kramer	0.0026*
Group 5 * Time 2-3	-1.8013	1.0660	196	-1.69	0.0927	Tukey-Kramer	0.5402
Group 5 * Time 1-3	-5.4885	1.3097	196	-4.19	<.0001	Tukey-Kramer	0.0006*
Groups 3-5 * Time 1	-6.3182	2.3476	196	-2.69	0.0077	Tukey-Kramer	0.0816
Groups 3-5 * Time 1-2	-10.0054	2.3553	196	-4.25	<.0001	Tukey-Kramer	0.0005*
Groups 3-5 * Time 1-3	-11.8067	2.4897	196	-4.74	<.0001	Tukey-Kramer	<.0001*
Groups 3-5 * Time 2	-6.7160	2.3630	196	-2.84	0.0050	Tukey-Kramer	0.0551
Groups 3-5 * Time 2-3	-8.5172	2.4970	196	-3.41	0.0008	Tukey-Kramer	0.0101*
Groups 3-5 * Time 3	-6.0201	2.6241	196	-2.29	0.0228	Tukey-Kramer	0.2014

Table 15 *Within- and between-subjects comparisons of the percentage of AS-units containing LI use.*

Differences of Least Squares Means							
Effect	Estimate	SD	DF	t Value	Pr > t 	Adjustment	Adj P
Group 3 * Time 1-2	0.007994	0.02092	196	0.38	0.7029	Tukey-Kramer	0.9989
Group 3 * Time 2-3	0.02742	0.02768	196	0.99	0.3231	Tukey-Kramer	0.9205
Group 3 * Time 1-3	0.03542	0.02668	196	1.33	0.1859	Tukey-Kramer	0.7695
Group 5 * Time 1-2	0.04200	0.02092	196	2.01	0.0461	Tukey-Kramer	0.3420
Group 5 * Time 2-3	0.01337	0.02768	196	0.48	0.6298	Tukey-Kramer	0.9967
Group 5 * Time 1-3	0.05536	0.02668	196	2.07	0.0393	Tukey-Kramer	0.3047
Groups 3-5 * Time 1	0.06834	0.04088	196	1.67	0.0962	Tukey-Kramer	0.5522
Groups 3-5 * Time 1-2	0.1103	0.04114	196	2.68	0.0079	Tukey-Kramer	0.0835
Groups 3-5 * Time 1-3	0.1237	0.04250	196	2.91	0.0040	Tukey-Kramer	0.0458*
Groups 3-5 * Time 2	0.1023	0.04140	196	2.47	0.0143	Tukey-Kramer	0.1376
Groups 3- 5 * Time 2-3	0.1157	0.04275	196	2.71	0.0074	Tukey-Kramer	0.0785
Groups 3-5 * Time 3	0.08828	0.04406	196	2.00	0.0465	Tukey-Kramer	0.3440

Table 16 Spearman's rank-order correlation between CAF measures for Year 3 group at the three testing times ($N = 40$).

		Complexity			Accuracy			Fluency	
		Words/ AS-unit	Clauses/ AS-unit	D	Correct Verbs/ Total Verbs	EFAS/ AS-unit	Erros/ AS-unit	Words/ Minute	L1/ AS- unit
Time 1									
Complexity									
Words/AS-unit	Correlation Coefficient	1.000	.633**	.014	-.017	.010	.351*	.462**	.247
	Sig. (2-tailed)		.000	.934	.917	.950	.027	.003	.124
Clauses/AS-unit	Correlation Coefficient	.633**	1.000	.152	.081	.302	.111	.525**	.157
	Sig. (2-tailed)	.000		.363	.620	.059	.495	.000	.333
D	Correlation Coefficient	.014	.152	1.000	-.177	-.210	-.118	.358*	.368*
	Sig. (2-tailed)	.934	.363		.288	.205	.482	.027	.023
Accuracy									
Correct Verbs/ Total Verbs	Correlation Coefficient	-.017	.081	-.177	1.000	.325*	-.288	.069	-.173
	Sig. (2-tailed)	.917	.620	.288		.040	.072	.673	.285
EFAS/AS-unit	Correlation Coefficient	.010	.302	-.210	.325*	1.000	-.542**	.043	-.277
	Sig. (2-tailed)	.950	.059	.205	.040		.000	.792	.084
Erros/AS-unit	Correlation Coefficient	.351*	.111	-.118	-.288	-.542**	1.000	.058	-.119
	Sig. (2-tailed)	.027	.495	.482	.072	.000		.722	.464
Fluency									
Words/Minute	Correlation Coefficient	.462**	.525**	.358*	.069	.043	.058	1.000	.320*
	Sig. (2-tailed)	.003	.000	.027	.673	.792	.722		.044
L1/AS-unit	Correlation Coefficient	.247	.157	.368*	-.173	-.277	-.119	.320*	1.000
	Sig. (2-tailed)	.124	.333	.023	.285	.084	.464	.044	

		Complexity			Accuracy			Fluency	
		Words/ AS-unit	Clauses/ AS-unit	D	Correct Verbs/ Total Verbs	EFAS/ AS-unit	Erros/ AS-unit	Words/ Minute	L1/ AS- unit
Time 2									
Complexity									
Words/AS-unit	Correlation Coefficient	1.000	.474**	.267	-.108	-.303	.472**	.427**	.232
	Sig. (2-tailed)		.002	.096	.509	.058	.002	.006	.149
Clauses/AS-unit	Correlation Coefficient	.474**	1.000	.338*	-.336*	-.239	.164	.212	.186
	Sig. (2-tailed)	.002		.033	.034	.137	.312	.190	.250
D	Correlation Coefficient	.267	.338*	1.000	-.272	-.433**	-.011	.406**	.714**
	Sig. (2-tailed)	.096	.033		.089	.005	.945	.009	.000
Accuracy									
Correct Verbs/ Total Verbs	Correlation Coefficient	-.108	-.336*	-.272	1.000	.531**	-.468**	.018	-.134
	Sig. (2-tailed)	.509	.034	.089		.000	.002	.911	.409
EFAS/AS-unit	Correlation Coefficient	-.303	-.239	-.433**	.531**	1.000	-.679**	-.161	-.505**
	Sig. (2-tailed)	.058	.137	.005	.000		.000	.321	.001
Erros/AS-unit	Correlation Coefficient	.472**	.164	-.011	-.468**	-.679**	1.000	.159	.002
	Sig. (2-tailed)	.002	.312	.945	.002	.000		.327	.990
Fluency									
Words/Minute	Correlation Coefficient	.427**	.212	.406**	.018	-.161	.159	1.000	.267
	Sig. (2-tailed)	.006	.190	.009	.911	.321	.327		.095
L1/AS-unit	Correlation Coefficient	.232	.186	.714**	-.134	-.505**	.002	.267	1.000
	Sig. (2-tailed)	.149	.250	.000	.409	.001	.990	.095	

Time 3	Complexity			Accuracy			Fluency		
	Words/ AS-unit	Clauses/ AS-unit	D	Correct Verbs/ Total Verbs	EFAS/ AS-unit	Erros/ AS-unit	Words/ Minute	L1/ AS- unit	
Complexity									
Words/AS-unit	Correlation Coefficient	1.000	.708**	.019	.115	.036	.308	.354*	-.057
	Sig. (2-tailed)	.	.000	.908	.480	.828	.053	.025	.727
Clauses/AS-unit	Correlation Coefficient	.708**	1.000	-.053	.133	.145	.062	.257	-.107
	Sig. (2-tailed)	.001	.	.751	.414	.371	.704	.109	.512
D	Correlation Coefficient	.019	-.053	1.000	-.146	-.149	-.445**	.068	.588**
	Sig. (2-tailed)	.908	.751	.	.383	.373	.005	.687	.000
Accuracy									
Correct Verbs/ Total Verbs	Correlation Coefficient	.115	.133	-.146	1.000	.278	-.375*	.346*	-.265
	Sig. (2-tailed)	.480	.414	.383	.	.082	.017	.029	.099
EFAS/AS-unit	Correlation Coefficient	-.036	.145	-.149	.278	1.000	-.404**	-.132	-.616**
	Sig. (2-tailed)	.838	.371	.373	.082	.	.010	.415	.000
Erros/AS-unit	Correlation Coefficient	.005	.062	-.445**	-.375*	-.404**	1.000	.032	.138
	Sig. (2-tailed)	.976	.704	.005	.017	.010	.	.844	.396
Fluency									
Words/Minute	Correlation Coefficient	.354*	.257	.068	.346*	-.132	-.032	1.000	.032
	Sig. (2-tailed)	.025	.109	.687	.029	.415	.844	.	.845
L1/AS-unit	Correlation Coefficient	-.057	-.107	.588**	-.265	-.616**	-.138	.032	1.000
	Sig. (2-tailed)	.727	.512	.000	.099	.000	.396	.845	.

** . Correlation is significant at the 0.01 level (2-tailed).

* . Correlation is significant at the 0.05 level (2-tailed).

Table 17 Spearman's rank-order correlation between CAF measures for Year 5 group at the three testing times ($N = 40$).

		Complexity			Accuracy			Fluency	
		Words/ AS-unit	Clauses/ AS-unit	D	Correct Verbs/ Total Verbs	EFAS/ AS-unit	Erros/ AS-unit	Words/ Minute	L1/ AS- unit
Time 1									
Complexity									
Words/AS-unit	Correlation Coefficient	1.000	.725**	.031	.013	-.287	.394*	.422**	-.115
	Sig. (2-tailed)	.	.000	.847	.937	.073	.012	.007	.440
Clauses/AS-unit	Correlation Coefficient	.725**	1.000	.079	-.139	-.114	.289	.314*	-.253
	Sig. (2-tailed)	.000	.	.629	.393	.485	.071	.049	.116
D	Correlation Coefficient	.031	.079	1.000	.202	.183	-.269	.100	-.074
	Sig. (2-tailed)	.847	.629	.	.211	.259	.094	.540	.651
Accuracy									
Correct Verbs/ Total Verbs	Correlation Coefficient	.013	-.139	.202	1.000	.191	-.533**	.336*	-.018
	Sig. (2-tailed)	.937	.393	.211	.	.238	.000	.034	.911
EFAS/AS-unit	Correlation Coefficient	-.287	-.114	.183	.191	1.000	-.756**	.282	-.459**
	Sig. (2-tailed)	.073	.485	.259	.238	.	.000	.078	.003
Erros/AS-unit	Correlation Coefficient	.394*	.289	-.269	-.533**	-.756**	1.000	-.156	.126
	Sig. (2-tailed)	.012	.071	.094	.000	.000	.	.336	.437
Fluency									
Words/Minute	Correlation Coefficient	.422**	.314*	.100	.336*	.282	-.156	1.000	-.402*
	Sig. (2-tailed)	.007	.049	.540	.034	.078	.336	.	.010
L1/AS-unit	Correlation Coefficient	-.115	-.253	-.074	-.018	-.459**	.126	-.402*	1.000
	Sig. (2-tailed)	.480	.116	.651	.911	.003	.437	.010	.

		Complexity			Accuracy			Fluency	
		Words/ AS-unit	Clauses/ AS-unit	D	Correct Verbs/ Total Verbs	EFAS/ AS-unit	Erros/ AS-unit	Words/ Minute	L1/ AS- unit
Time 2									
Complexity									
Words/AS-unit	Correlation Coefficient	1.000	.416**	-.096	.121	-.084	.230	.446**	-.030
	Sig. (2-tailed)	.	.008	.554	.456	.607	.153	.004	.856
Clauses/AS-unit	Correlation Coefficient	.416**	1.000	.198	-.162	-.051	.209	.448**	-.191
	Sig. (2-tailed)	.008	.	.220	.319	.754	.195	.004	.231
D	Correlation Coefficient	-.096	.198	1.000	-.065	.107	-.049	.209	.032
	Sig. (2-tailed)	.554	.220	.	.691	.511	.765	.195	.846
Accuracy									
Correct Verbs/ Total Verbs	Correlation Coefficient	.121	-.162	-.065	1.000	.526**	-.412**	.184	-.296
	Sig. (2-tailed)	.456	.319	.691	.	.000	.008	.256	.063
EFAS/AS-unit	Correlation Coefficient	-.084	-.051	.107	.526**	1.000	-.791**	.136	-.114
	Sig. (2-tailed)	.607	.754	.511	.000	.	.000	.403	.483
Erros/AS-unit	Correlation Coefficient	.230*	.209	-.049	-.412**	-.791**	1.000	.062	.047
	Sig. (2-tailed)	.153	.195	.765	.008	.000	.	.703	.774
Fluency									
Words/Minute	Correlation Coefficient	.446**	.448**	.209	.184	.136	.062	1.000	-.237
	Sig. (2-tailed)	.004	.004	.195	.256	.403	.703	.	.141
L1/AS-unit	Correlation Coefficient	-.030	-.191	.032	-.296	-.114	.047	-.237	1.000
	Sig. (2-tailed)	.856	.231	.846	.063	.483	.774	.141	.

	Complexity			Accuracy			Fluency		
	Words/ AS-unit	Clauses/ AS-unit	D	Correct Verbs/ Total Verbs	EFAS/ AS-unit	Erros/ AS-unit	Words/ Minute	L1/ AS-unit	
Time 3									
Complexity									
Words/AS-unit	Correlation Coefficient	1.000	.492**	.144	.209	-.090	.308	.572**	-.001
	Sig. (2-tailed)	.	.001	.381	.196	.579	.053	.000	.995
Clauses/AS-unit	Correlation Coefficient	.492**	1.000	.013	.086	.051	.174	.134	-.225
	Sig. (2-tailed)	.001	.	.938	.596	.755	.283	.411	.164
D	Correlation Coefficient	.144	.013	1.000	.253	.234	-.132	.212	-.169
	Sig. (2-tailed)	.381	.938	.	.120	.152	.422	.196	.305
Accuracy									
Correct Verbs/ Total Verbs	Correlation Coefficient	.209	.086	.253	1.000	.371*	-.377*	.032	-.370*
	Sig. (2-tailed)	.196	.596	.120	.	.018	.016	.843	.019
EFAS/AS-unit	Correlation Coefficient	-.090	.051	.234	.371*	1.000	-.785**	.050	-.502**
	Sig. (2-tailed)	.579	.755	.152	.018	.	.000	.762	.001
Erros/AS-unit	Correlation Coefficient	.308	.174	-.132	-.377*	-.785**	1.000	.043	.307
	Sig. (2-tailed)	.053	.283	.422	.016	.000	.	.792	.054
Fluency									
Words/Minute	Correlation Coefficient	.572**	.134	.212	.032	.050	.043	1.000	.011
	Sig. (2-tailed)	.000	.411	.196	.843	.762	.792	.	.946
L1/AS-unit	Correlation Coefficient	-.001	-.225	-.169	-.370*	-.502**	.307	.011	1.000
	Sig. (2-tailed)	.995	.164	.305	.019	.001	.054	.946	.

** . Correlation is significant at the 0.01 level (2-tailed).

* . Correlation is significant at the 0.05 level (2-tailed).

Table 18 *Within- and between-subjects comparisons of the ratio of L1 use for metacognitive talk.*

Differences of Group*time Least Squares Means Adjustment for Multiple Comparisons: Tukey-Kramer						
Effect	Estimate	SD	DF	t Value	Pr > t 	Adj P
Group 3 * Time 1-2	0.6694	0.2127	196	3.15	0.0019	0.0231*
Group 3 * Time 2-3	-0.4055	0.2930	196	-1.38	0.1680	0.7368
Group 3 * Time 1-3	0.2640	0.2392	196	1.10	0.2711	0.8794
Group 5 * Time 1-2	0.3124	0.3231	196	0.97	0.3349	0.9279
Group 5 * Time 2-3	0.1054	0.4843	196	0.22	0.8280	0.9999
Group 5 * Time 1-3	0.4177	0.4394	196	0.95	0.3429	0.9326
Groups 3-5 * Time 1	1.1147	0.4995	196	2.23	0.0268	0.2280
Group 3-5 * Time 1-2	1.4271	0.5421	196	2.63	0.0092	0.0944
Group 3-5 * Time 1-3	1.5325	0.5649	196	2.71	0.0073	0.0772
Group 3-5 * Time 2	0.7577	0.5698	196	1.33	0.1852	0.7682
Groups 3-5 * Time 2-3	0.8630	0.5916	196	1.46	0.1462	0.6907
Groups 3-5 * Time 3	1.2685	0.5791	196	2.19	0.0297	0.2469

Table 19 *Within- and between group comparisons of the ratio of L1 use with vocabulary function.*

Differences of Group*time Least Squares Means Adjustment for Multiple Comparisons: Tukey-Kramer						
Effect	Estimate	SD	DF	t Value	Pr > t 	Adj P
Group 3 * Time 1-2	0.05972	0.1853	196	0.32	0.7476	0.9995
Group 3 * Time 2-3	0.4249	0.1818	196	2.34	0.0205	0.1845
Group 3 * Time 1-3	0.4846	0.2384	196	2.03	0.0435	0.3279
Group 5 * Time 1-2	0.4378	0.2281	196	1.92	0.0563	0.3932
Group 5 * Time 2-3	0.4785	0.2504	196	1.91	0.0575	0.3987
Group 5 * Time 1-3	0.9163	0.3040	196	3.01	0.0029	0.0342*
Groups 3-5 * Time 1	0.2268	0.2943	196	0.77	0.4419	0.9721
Group 3-5 * Time 1-2	0.6646	0.3021	196	2.20	0.0290	0.2426
Group 3-5 * Time 1-3	1.1431	0.3352	196	3.41	0.0008	0.0101*
Group 3-5 * Time 2	0.6049	0.2965	196	2.04	0.0427	0.3237
Groups 3-5 * Time 2-3	1.0833	0.3302	196	3.28	0.0012	0.0153*
Groups 3-5 * Time 3	0.6585	0.3452	196	1.91	0.0579	0.4006

Table 20 *Within- and between group comparisons of the ratio of L1 use with discourse markers function.*

Differences of Group*time Least Squares Means Adjustment for Multiple Comparisons: Tukey-Kramer						
Effect	Estimate	SD	DF	t Value	Pr > t 	Adj P
Group 3 * Time 1-2	0.2136	0.3278	196	0.65	0.5155	0.9868
Group 3 * Time 2-3	0.1001	0.4123	196	0.24	0.8084	0.9999
Group 3 * Time 1-3	0.3137	0.4765	196	0.66	0.5111	0.9862
Group 5 * Time 1-2	0.1252	0.3948	196	0.32	0.7515	0.9996
Group 5 * Time 2-3	-0.4274	0.4241	196	-1.01	0.3148	0.9149
Group 5 * Time 1-3	-0.3023	0.4931	196	-0.61	0.5406	0.9900
Groups 3-5 * Time 1	0.4249	0.5218	196	0.81	0.4164	0.9646
Group 3-5 * Time 1-2	0.5500	0.5447	196	1.01	0.3138	0.9143
Group 3-5 * Time 1-3	0.1226	0.5162	196	0.24	0.8125	0.9999
Group 3-5 * Time 2	0.3365	0.5685	196	0.59	0.5546	0.9915
Groups 3-5 * Time 2-3	-0.09097	0.5412	196	-0.17	0.8667	1.0000
Groups 3-5 * Time 3	-0.1911	0.5846	196	-0.33	0.7441	0.9995