L2 English Young Learners' Oral Production Skills in CLIL and EFL Settings: A Longitudinal Study.

Anna Bret Blasco

DOCTORAL DISSERTATION

SUPERVISED BY: Dr Elisabet Pladevall Ballester

November 2014

Departament de Filologia Anglesa i Germanística
Facultat de Filosofia i Lletres
Universitat Autònoma de Barcelona

Universitat Autònoma de Barcelona

ACKNOWLEDGEMENTS

Foremost, I would like to express my deepest gratitude to my supervisor, Dr Elisabet Pladevall, for her invaluable and encouraging support. This has been the longest and most exciting learning journey I have ever made and I thank her for having given me the chance to experience it by her side. I could not have had a better travelling companion for such a journey. I really enjoyed listening to her thoughtful and critical comments. They were extremely revealing. Moreover, the completion of this study would not have been possible without her constant and motivating messages of support.

I would also like to express my gratitude to the late Mia Victori. I would not be presenting this study if she had not seen something in me during my undergraduate studies at UAB.

A huge thanks to all the members of the CLIL-SLA Project, in particular to Elisabet Pladevall, Anna Vallbona, Alexandra Vraciu, Marc Miret, and Patricia Martínez, for their help in the data collection sessions. A big thanks to Mary Louise Walsh and Berta Torras too for helping me with the transcriptions.

A very special thanks goes to the schools and teachers who allowed us to collect the data and interrupt their sometimes frenetic school life. Collaborations between researchers and school teachers like this are essential for the improvement and development of foreign language teaching pedagogy.

I also want to thank Oliver Valero, in the Servei d'Estadistica Aplicada in the UAB, for his help in the statistical analyses and for his extremely helpful support at all times.

A special thanks to all those scholars who have shown interest in this study, among whom I would like to name Professor Nikolov, Professor García Mayo and Dr Gilabert. Their opinions and unexpected interest were a great source of motivation in the toughest moments.

Per últim, vull expressar un agraïment molt especial al Xevi per la seva paciència, comprensió i recolzament incondicional en els moments més difícils. El seu positivisme i sentit de l'humor han estat claus en l'elaboració d'aquesta tesi. No podré oblidar mai el que hem passat junts en el tram final d'aquesta tesi, i és per això que li agraeixo cada gest i paraula. També vull donar les gràcies als meus pares, Maria i Toni, per haver cregut en mi durant tot aquest temps i per alimentar el meu esperit crític ja des de ben petita. Un agraïment molt especial pel meu germà Marc també, per fer-me riure quan més ho necessitava. Vull agrair la comprensió i paciència dels meus familiars i a mics. Espero tornar-vos tot el temps que us mereixeu. Finalment, vull expressar el meu més sentit agraïment al Josep per ensenyar-me a no rendir-me mai.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ABBREVIATIONS

ADJR	Adjective ratio
AS-unit	Analysis of speech unit
BE	Bilingual Education
BICS	Basic interpersonal communication skills
CAF	Complexity, accuracy and fluency
CALP	Cognitive academic language proficiency
CBI	Content-based Instruction
CLAN	Computerized language analysis
CLIL	Content and Language Integrated Learning
CPH	Critical Period Hypothesis
CU	Coordinate units
CV	Correct verb forms
EAL	English as an Additional Language
EFL	English as a Foreign Language
EFU	Error-free unit
ESL	English as a Second Language
FFI	Form-focused instruction
FonF	Focus on Form
FL	Foreign Language
ISLA	Instructed Second Language Acquisition
LEP	Limited English Proficiency
L1WR	L1 word ratio
L2	Second Language
MFI	Meaning-focused instruction
NR	Noun ratio
SLA	Second Language Acquisition
SPRW	Speech rate in words
SU	Subordinate units
T	Time
TBL	Task-based Learning
TL	Target language
TNADJ	Total number of adjectives
TNCU	Total number of coordinate units
TNCV	Total number of correct verb forms

TNN	Total number of nouns
TNSU	Total number of subordinate units
TNU	Total number of units
TNV	Total number of lexical verbs
TNW	Total number of words
TNWE	Total number of words in English
TTT	Total task time
VR	Verb ratio
YLL	Young Language Learners

CHAPTER 1 Introduction and research questions

Content and Language Integrated Learning (CLIL) has been defined as an umbrella term which refers to "all types of provision in which a second language is used to teach certain subjects in the curriculum other than language lessons themselves" (Eurydice European Unit, 2006: 8). The implementation of CLIL programmes started to become popular in Europe in the 1990s and since then many European schools have included CLIL programmes in their educational projects as an innovative approach to the teaching of foreign languages as well as a means to achieve the aims stated by the European Commission regarding multilingualism.

According to some researchers (Muñoz, 2007; Pérez-Cañado, 2012) CLIL emerged as an opportunity to overcome the low levels of foreign language proficiency in Europe, which on many occasions stem from an overuse of form-focused instructional approaches. This is why the introduction of CLIL as a new foreign language teaching approach was seen "as an alternative that could overcome the deficiencies in previous language models" (Muñoz, 2007: 17) by many stakeholders (i.e. academic community, educational authorities, families and schools).

Some of the advantages of CLIL over traditional foreign language instruction relate to the enhancement of genuine communication through the use of meaningful input and output (i.e. content subject matter) at the same time as attention to linguistic form and L2 development is provided and supported. Owing to this integration and balance between content and language, CLIL settings are usually seen as optimal instructional contexts for language learning to take place. However, even though the integration of language objectives into content teaching is the main theoretical principle behind CLIL, "evidence of this integration is scant" (Cenoz, Genesee and Gorter, 2014:

10). As a result, voices within CLIL and other communicative classroom settings have arisen calling for the need to give greater attention to language objectives in order to guarantee language learning and development (Lorenzo 2007, Lyster 2007, Muñoz 2007, Pérez-Vidal 2007, García-Mayo, 2011). The lack of attention to linguistic form was one of the main pedagogical flaws identified in immersion and content-based instruction programmes, which were based on the assumption that content learning on its own would automatically trigger language learning. As Swain claimed back in 1988, "not all content teaching is necessarily good language teaching" (p. 68). Therefore, if CLIL is to achieve its potential, efforts need to be made to incorporate the lessons learnt from previous content-based approaches into CLIL settings.

Many research studies within instructed SLA have been conducted to test the efficiency of CLIL programmes in Europe and the vast majority have reported positive effects of CLIL on language learning. Despite that, as described in the following section, CLIL research designs might have had a direct impact on the results obtained in favour of CLIL learners. In addition, comparability across research findings and national contexts has been difficult to achieve due to the wide variety of CLIL models implemented in Europe, which range from programmes that offer between four and six hours of CLIL instruction a week to minimal CLIL exposure which provides learners with one hour of CLIL instruction a week. The focus of study in this dissertation will be to examine the effects of minimal CLIL exposure, a widely implemented CLIL programme in Catalonia, on young learners' oral production skills, a research topic which has generally remained unexplored within CLIL research as most CLIL studies have been conducted in secondary school settings.

As for the development of oral production skills in CLIL settings, most studies (Järvinen, 2005; Mewald, 2007; Juan-Garau, 2010; Bret 2011) report significant

differences in fluency in favour of CLIL learners. Furthermore, speaking skills are believed to be one of the areas which benefits the most from CLIL instruction (Dalton-Puffer, 2011) due to the communicative nature of CLIL lessons. Despite that, some pieces of CLIL research (Stotz and Meuter, 2003; Wannagat, 2007) highlight the lack of teaching practices to foster oral development in CLIL settings as well as the limited use of the L2 on the part of CLIL learners. Thus, in an attempt to further investigate the effects of CLIL on young learners' oral production skills, this study will aim at shedding some light into the development of oral production skills in CLIL+EFL and EFL settings.

1.1 Justification of the study

The imbalance between CLIL programmes being implemented and the scarcity of CLIL research in the mid 1990s led to the publication of the first CLIL research studies in the 2000s. This first wave of empirical research studies has been strongly criticised by authors like Bruton (2011, 2013) on the basis of their methodological flaws and limitations, which in many cases derived to hasty conclusions on the effects of CLIL on foreign language learning all over Europe. Despite this, this first series of CLIL research studies was necessary in order for CLIL research to develop. As Dalton-Puffer and Nikula (2014: 118) acknowledge, "as with any other research area in the process of developing and consolidating, there are areas of CLIL research that need more intensive research". Thus, a new wave of CLIL research is currently being conducted with the aim of filling the gap in the field and overcoming previous methodological limitations.

The study presented in this dissertation, which belongs to a larger longitudinal research project called *CLIL-SLA Project*[1], responds to the calls for new methodological designs within CLIL research, as it aims at addressing some of the challenges found in previous research in an attempt to obtain reliable research findings and contribute to the consolidation of CLIL research as a solid research ground. In the following paragraphs an overview of the most relevant methodological challenges this study attempts to overcome are presented.

According to Dalton-Puffer and Nikula (2014) and Pérez-Cañado (2012), future CLIL research needs to control for moderating variables such as amount of English input (hours of instruction) between CLIL and non-CLIL groups, learners' extracurricular exposure to English or type of school. Among these moderating variables, amount of English input received by CLIL and non-CLIL groups seems to outstand. Up to the present, CLIL research has compared learners who have been exposed to EFL instruction and learners who have received the same amount of EFL instruction in addition to CLIL instruction over a certain period of time, which means that CLIL learners are usually exposed to greater amounts of English input than non-CLIL learners due to their extra CLIL exposure. As a consequence, differences in favour of CLIL learners are hard to be attributed to CLIL exposure only. This is precisely one of the gaps that this dissertation attempts to fill, as amount of English instruction has been kept constant between CLIL and non-CLIL groups throughout the two years of the study. In addition to amount of English instruction between groups, this dissertation also controls for the amount of previous English exposure within the school context, the learners' extracurricular exposure to English and the type of school.

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

Another distinctive feature of this piece of research relates to its longitudinal nature, since this dissertation is based on a two-year longitudinal study in which data from the CLIL and non-CLIL groups has been collected and compared at four times (T0, T1, T2 and T3) throughout two academic years. So far, most European research conducted on CLIL has been cross-sectional. Because of this, Pérez-Cañado (2012) and Dalton-Puffer and Nikula (2014) highlight the importance of carrying out longitudinal research projects which trace language development in CLIL settings.

Regarding participant selection, as Bruton (2011) criticised, CLIL research has tended to conduct studies in which CLIL and non-CLIL groups are not comparable in terms of the learners' English proficiency level and language aptitude, as CLIL groups are usually made up of talented and motivated English language learners who have voluntarily chosen to enrol in CLIL courses. As the author claims "students who will be academically motivated to succeed in the FL" are usually selected to participate in CLIL programmes (p. 524). In this study, a careful participant selection was carried out in order to guarantee the homogeneity between CLIL and non-CLIL groups. In addition, data on CLIL and non-CLIL learners' achievement scores in the measures selected to analyse their oral production skills is also provided before the implementation of CLIL in order to assure the comparability of CLIL and non-CLIL learners at the onset of the research study.

Another relevant aspect that needs to be addressed in terms of research design according to Pérez-Cañado (2012) is the use of mixed research designs in which both quantitative and qualitative data analyses are included. In this dissertation, the statistical analyses are complemented with a more qualitative analysis of the output produced by a selection of participants with different fluency levels over the two years of the study. Furthermore, as Pérez-Cañado suggests, relevant data obtained from classroom

observation sessions, student opinion questionnaires and interviews will be sporadically reffered to in an attempt to further interpret the results obtained.

In addition to the efforts made to overcome methodological limitations found in previous CLIL research, this dissertation also contributes to the field by exploring the effects of CLIL on young language learners in their last two years of primary education. So far, most quantitative CLIL research has been conducted in secondary school settings with teenagers. This is one of the first attempts within European CLIL research to trace the development of young learners' oral production skills in CLIL and non-CLIL settings.

To sum up, this dissertation contributes to previous CLIL research in several ways. Firstly, the methodological design responds to the call for more rigorous CLIL research by controlling moderating variables and guaranteeing the comparability of CLIL and non-CLIL groups. Secondly, the longitudinal nature of the data gathered in this dissertation will provide the field with new research findings on the effects of CLIL on oral production skills and their evolution. Thirdly, the data is analysed using quantitative and qualitative methods in order to obtain a more detailed picture of language learning in CLIL settings. In addition, data triangulation will also be fostered in the interpretation of the results. Finally, this study will contribute to CLIL research and instructed SLA by analysing the evolution of young language learners' oral production skills in CLIL and non-CLIL settings.

1.2 Research questions and hypotheses

As stated above, the main focus of this study is to investigate the L2 English oral output of young learners who have been exposed to CLIL and EFL instruction (CLIL group) and young learners who have only received EFL exposure (non-CLIL group) in terms of complexity, accuracy and fluency in two oral tasks: an interview and a narrative task. Thus, the research questions that this dissertation seeks to answer are the following:

Research question 1 relates to the effects of CLIL+EFL and EFL instruction on complexity, accuracy and fluency at T1, T2 and T3 in the interview and narrative tasks:

RQ 1 Are there statistically significant differences in propositional, syntactic and lexical complexity, accuracy and fluency achievement scores between young learners in the CLIL and non-CLIL groups keeping amount of English input (hours of instruction) constant between the two groups at T1, T2 and T3 in the interview and narrative tasks?

Research question 2 is concerned with the development of syntactic and lexical complexity, accuracy and fluency in CLIL+EFL and EFL settings from T0 to T1, from T1 to T2, from T2 to T3 and from T0 to T3 in the interview and narrative tasks:

RQ 2 Do CLIL and non-CLIL young learners' achievement scores in complexity, accuracy and fluency develop significantly from T0 to T1, from T1 to T2, from T2 to T3 and from T0 to T3 keeping amount of English input (hours of instruction) constant between the two groups in the interview and narrative tasks?

Research question 3 relates to the effects of the learners' initial proficiency level on complexity, accuracy and fluency achievement scores in CLIL+EFL and EFL settings in the interview and narrative tasks:

RQ 3 Are there statistically significant differences in complexity, accuracy and fluency achievement scores between young learners in the CLIL and non-CLIL groups according to their initial proficiency level keeping amount of English input (hours of instruction) constant between the two groups at T1, T2 and T3 in the interview and narrative tasks?

RQ 3.1 Are there statistically significant differences in propositional, syntactic and lexical complexity, accuracy and fluency achievement scores between high achievers in the CLIL and non-CLIL groups at T1, T2 and T3 in the interview and narrative tasks?

RQ 3.2 Are there statistically significant differences in propositional, syntactic and lexical complexity, accuracy and fluency achievement scores between low achievers in the CLIL and non-CLIL groups at T1, T2 and T3 in the interview and narrative tasks?

RQ 3.3 Do CLIL low achievers obtain more statistically significant differences in propositional, syntactic and lexical complexity, accuracy and fluency achievement scores than CLIL high achievers when compared to their respective peers in the non-CLIL group at T1, T2 and T3 in the interview and narrative tasks?

Research question 4 is concerned with the effects of initial proficiency level on the development of complexity, accuracy and fluency in CLIL+EFL and EFL settings in the interview and narrative tasks:

RQ 4 Do CLIL and non-CLIL young learners' achievement scores in complexity, accuracy and fluency develop significantly from T0 to T1, from T1 to T2, from T2 to T3 and from T0 to T3 according to the learners' initial proficiency level keeping amount of English input (hours of instruction) constant between the two groups in the interview and narrative tasks?

RQ 4.1 Do CLIL and non-CLIL high achievers' scores in propositional, syntactic and lexical complexity, accuracy and fluency develop significantly from T0 to T1, from T1 to T2, from T2 to T3 and from T0 to T3 in the interview and narrative tasks?

RQ 4.2 Do CLIL and non-CLIL low achievers' scores in propositional, syntactic and lexical complexity, accuracy and fluency develop significantly from T0 to T1, from T1 to T2, from T2 to T3 and from T0 to T3 in the interview and narrative tasks?

RQ 4.3 Do CLIL low achievers' scores in propositional, syntactic and lexical complexity, accuracy and fluency show greater development from T0 to T3 than those obtained by CLIL high achievers in the interview and narrative tasks?

Research question 5 relates to the relationship between syntactic complexity, accuracy and fluency in CLIL+EFL and EFL settings from T0 to T3 in the interview and narrative tasks:

RQ 5 What is the relationship between syntactic complexity, accuracy and fluency achievement scores obtained by young learners in the CLIL and non-CLIL groups from T0 to T3 keeping amount of English input (hours of instruction) constant between the two groups in the interview and narrative tasks?

RQ 5.1 Is there a correlation between syntactic complexity and accuracy achievement scores in CLIL+EFL and EFL settings from T0 to T3 in the interview and narrative tasks?

RQ 5.2 Is there a correlation between syntactic complexity and fluency achievement scores in CLIL+EFL and EFL settings from T0 to T3 in the interview and narrative tasks?

RQ 5.3 Is there a correlation between accuracy and fluency achievement scores in CLIL+EFL and EFL settings from T0 to T3 in the interview and narrative tasks?

Research question 6 is concerned with the qualitative description in terms of complexity, accuracy, fluency and oral narrative development of the output produced at T0 and T3 by a selection of CLIL and non-CLIL learners with different fluency levels in the interview and narrative tasks:

RQ 6 What are the characteristics, in terms of complexity, accuracy, fluency and oral narrative competence, of the output produced by a selection of young learners with different fluency levels at T0 and T3 in CLIL+EFL and EFL settings in the interview and narrative tasks?

RQ 6.1 What are the characteristics, in terms of lexical and syntactic complexity, accuracy and fluency, of a highly fluent, a fluent and a dysfluent young learner's oral output at T0 and T3 in CLIL+EFL and EFL settings in the interview and narrative tasks?

RQ 6.2 What stages of oral narrative development are identified in the oral output of a highly fluent, a fluent and a dysfluent young learner at T0 and T3 in CLIL+EFL and EFL settings in the narrative task?

According to previous research on the effects of CLIL on the learners' oral production skills (Admiraal, Westhoff and de Bot, 2006; Hüttner and Rieder-Bünemann, 2007 and 2010; Mewald, 2007; Zytadiβ, 2007; Lasagabaster, 2008; Juan-Garau, 2010; Lorenzo, Casal and Moore, 2010; Várkuti, 2010; Bret, 2011), learners who have been exposed to CLIL instruction significantly outperform learners who have only received EFL instruction in fluency, mainly. Additionally, CLIL learners also perform better than non-CLIL learners in areas like syntactic complexity, lexical complexity and accuracy. Thus, in the light of these CLIL findings, the hypotheses posed in this dissertation are the following:

Hypothesis 1 Keeping amount of English input (hours of instruction) constant between CLIL and non-CLIL groups, statistically significant differences will be found in favour

of CLIL learners in fluency in the interview and narrative tasks but not necessarily in propositional, syntactic and lexical complexity or accuracy.

Hypothesis 2 Keeping amount of English input (hours of instruction) constant between CLIL and non-CLIL groups, CLIL learners' achievement scores in propositional, syntactic and lexical complexity, accuracy and fluency will show periods of significant development in the interview and narrative tasks.

With regard to the effects of initial proficiency level on language learning in CLIL settings, very little research has been published up to date. Pladevall-Ballester's study (in press) on the effects of young learners' initial proficiency level on their receptive skills in CLIL settings, for instance, reports that low and mid proficient learners' development in listening skills is greater than that found for high proficient learners, indicating that weaker English learners in primary schools benefit more from CLIL instruction than highly proficient learners in terms of their listening skills. As for reading skills, Pladevall-Ballester's findings show that low achievers surpassed mid achievers at the end of the investigation, providing further evidence of the positive impact CLIL has on weaker English students' language development. Aguilar and Muñoz's (2014) study on the effects of CLIL in tertiary education also yields similar results. In this case, the language development of three groups of postgraduate students with different proficiency levels who enrolled in a CLIL course for a semester was analysed in regard to their listening and grammar skills. The results of the study indicate that, as found in the case of young language learners, less proficient learners obtained higher gains in receptive and grammar skills than more proficient learners.

As for the effects of young learners' proficiency level on the development of their oral productive skills, no research has been conducted yet. Thus, according to previous research on CLIL and on the effects of initial proficiency level on receptive and grammar skills, the following hypotheses are raised:

Hypothesis 3 Keeping amount of English input (hours of instruction) constant between CLIL and non-CLIL groups, more statistically significant differences will be found throughout the study in favour of CLIL low achievers when compared to their respective peers in the non-CLIL group in the interview and narrative tasks. The effects of CLIL exposure are predicted not to be so noticeable among CLIL high achievers.

Hypothesis 4 Keeping amount of English input (hours of instruction) constant between CLIL and non-CLIL groups, the achievement scores obtained by low achievers from the CLIL group will show greater development throughout the study than those obtained by high achievers when compared to the their respective peers in the non-CLIL group,

As for the relationship between syntactic complexity, accuracy and fluency measures in language development (see Chapter 4), research conducted within the framework of the (Extended) Trade-off Hypothesis (Foster and Skehan, 1996; Skehan and Foster, 2012) in the context of task-based learning has generally confirmed the existence of trade-off effects between complexity and accuracy as a result of task characteristics and conditions. These findings suggest that the learners' output may become syntactically more complex and sophisticated at the expense of accuracy in certain learning contexts. Likewise, Ferrari's (2012) longitudinal study on complexity, accuracy and fluency oral development reports the existence of trade-off effects

between complexity and accuracy for certain periods of time along with a gradual growth in syntactic complexity and fluency. Taking these findings into consideration the following hypotheses are raised:

Hypothesis 5 Keeping amount of English input (hours of instruction) constant between CLIL and non-CLIL groups, syntactic complexity, accuracy and fluency are not expected to develop simultaneously in CLIL+EFL and EFL settings from T0 to T3 in the interview and narrative tasks.

H 5.1 An negative correlation is expected between syntactic complexity and accuracy achievement scores in CLIL+EFL and EFL settings from T0 to T3 in the interview and narrative tasks.

H 5.2 A positive correlation is expected between syntactic complexity and fluency achievement scores in CLIL+EFL and EFL settings from T0 to T3 in the interview and narrative tasks.

H 5.3 An negative correlation is expected between accuracy and fluency achievement scores in CLIL+EFL and EFL settings from T0 to T3 in the interview and narrative tasks?

In regard to the hypotheses posed for the qualitative analysis, according to research (Skehan, 1998; Álvarez, 2006; McKay, 2006; Muñoz, Tragant and Torras, 2010; Lyster and Sato, 2013), young language learners rely on the use of formulaic language when communicating in the L2. As a consequence, high accuracy levels are likely to be expected. As for complexity, young learners' lexical and syntactic system develops as they become cognitively more mature and their proficiency level improves.

Concerning the development of young learners' oral narrative competence, Álvarez's (2006) study, which employed the narrative task used in this dissertation, suggests that learners progress through a series of stages which show gradual morphosyntactic development and use of narrative discourse properties. In addition, research conducted by Hüttner and Rieder-Büneman (2007, 2010) on the effects of CLIL on the narrative competence of learners suggests that CLIL learners have a more advanced command over narrative skills than non-CLIL learners. Therefore, taking this into account, the hypotheses proposed for the results obtained from the qualitative analysis are the following:

Hypothesis 6 The oral output produced in the interview and narrative tasks at T0 will be characterized by the use of simple utterances mainly containing formulaic language and content words. The degree of simplicity and amount of formulaic language is expected to vary according to the learners' fluency level. At T3, the output is predicted to be syntactically more complex through the use of coordination and subordination, especially the output produced by CLIL learners. Highly fluent learners are expected to display the highest degree of complexity.

> **H 6.1** In the narrative task, the oral output is predicted to develop through Álvarez's stages (2006) of narrative discourse, from stage 2, which is characterized by the use of isolated L2 content words, to stage 7, when subordination emerges. Highly fluent and fluent learners are predicted to attain stage 7. CLIL learners are expected to progress through stages faster than non-CLIL learners.

1.3 Organisation of the thesis

This section provides a summary of the organisation of the thesis, which includes eight chapters. Chapter 1 introduces the object of the study and presents the research questions and hypotheses this dissertation aims at exploring. The following three chapters describe the theoretical framework used for the interpretation of the results. Finally, the remaining four chapters provide an account of the methodology of the study, the results obtained, the interpretations of the results and the final conclusions.

Chapter 2 presents an account of the differences between SLA and instructed SLA (ISLA) and continues with a description of Housen and Pierrard's (2005) framework for instructed SLA research in which relevant areas for the interpretation of the results obtained in this study are highlighted. Additionally, Ellis' (1990, 1997 and 2005) theories and principles of instructed SLA are described. Finally, the chapter provides a thorough description of three mediating factors which influence the effects of instruction on second language learning, which include type of learner (young language learners), type of instruction (form-focused instruction and meaning-focused instruction) and type of language targeted (BICS and CALP).

Chapter 3 provides a description of the main instructional setting under analysis in this dissertation: Content and Language Integrated Learning. It begins with an account of the emergence of CLIL in Europe, which is complemented with a discussion of the different definitions provided of the term. In addition, the chapter offers a description of the two main influences of CLIL: French immersion programmes and content-based instruction. Finally, the chapter presents a description of micro and macro-level features of successful CLIL programmes and an overview of CLIL research in Europe.

Chapter 4 describes the three areas of L2 oral production which are explored in this dissertation: complexity, accuracy and fluency (CAF) in an attempt to gain insight into L2 oral development. It provides a detailed definition of each of the three elements as well as a description of their subdimensions. Additionally, it offers an account of the factors that influence the manifestations of CAF in L2 oral output by referring to the (Extended) Trade-off Hypothesis (Skehan, 1998 and Skehan and Foster, 2012) and the Cognition Hypothesis (Robinson, 2001). Finally, a selection of CAF research is also reviewed.

Chapter 5 describes the methodology used in this study. It starts with an account of the design of the study and continues with a description of the schools' context and the EFL curriculum in primary schools in Catalonia. After that, the chapter provides a description of the participants, CLIL and non-CLIL learners, the data collection procedures and the data collection instruments. In addition, Chapter 5 explains the measures and unit of analysis used to investigate the degrees of complexity, accuracy and fluency of the oral output young learners in this study produced. Finally, a description of the statistical and qualitative analyses employed to address the research questions of this dissertation is provided.

Chapter 6 presents the results obtained by CLIL and non-CLIL learners in the CAF measures selected to examine their oral performance on the interview and narrative tasks. First, the results obtained from the inter and intragroup comparisons at the four data collection times are presented. Then, the results of the correlation tests carried out to investigate the relationship between CAF measures are also presented for both CLIL and non-CLIL learners. Finally, the qualitative description of the oral output produced by a selection of CLIL and non-CLIL learners in the interview and narrative

tasks is presented with the aim of investigating L2 development in terms of the linguistic features and structures used at the onset and at the end of this study.

Chapter 7 discusses the results obtained by CLIL and non-CLIL learners in the quantitative and qualitative analyses in the interview and narrative tasks and highlights a number of claims derived from the main findings.

Eventually, Chapter 8 provides the final conclusions by answering the research questions posited in this dissertation. In addition, the limitations of the present study as well as future directions of CLIL research are also provided.

CHAPTER 2 Instructed Second Language Acquisition

Chapter 2 presents a description of the characteristics, rationale and scope of instructed Second Language Acquisition as suggested in recent literature. It starts with a brief account of the differences between SLA and Instructed Second Language Acquisition (ISLA) and continues with a description of the theoretical framework provided by Housen and Pierrad (2005) for instructed SLA research. A description of learning contexts in ISLA and the theory of Instructed Second Language Acquisition proposed by Ellis (1990, 1997 and 2005) follow. Finally, this chapter also provides a detailed overview of the main factors which might influence the effects of instruction and have important implications for the present study.

2.1 SLA versus instructed SLA

As Rod Ellis claimed back in 1984, two different types of Second Language Acquisition (SLA) research can be identified according to the context in which a second language is acquired: naturalistic SLA and classroom SLA, more recently re-labelled as instructed SLA. SLA refers to those contexts in which acquisition of the second language takes place naturally through unguided spontaneous interactions in genuine social situations, whereas ISLA makes reference to those settings in which the second language is learned through guided formal instruction. For the purposes of this study, the term *instruction* will be interpreted and used to refer to "any type of systematic attempt to enable or facilitate language learning by manipulating the mechanisms of learning and/or the conditions under which language learning occurs" (Housen and Pierrard, 2005: 2).

The origins of such a clear distinction between SLA and ISLA emerged from Krashen's widespread Acquisition-Learning hypothesis (1981), in which he claimed that the processes involved in language acquisition and learning are essentially different on the basis of the type of knowledge they produce. According to Krashen, acquisition in naturalistic contexts derives to subconscious linguistic knowledge, while learning in instructional contexts leads to conscious metalinguistic knowledge which is used by the learner to construct and modify L2 output. Other researchers like Gass (1989), on the contrary, believe that second language acquisition involves the same processes independently of the context where the language is learnt. Nowadays, the idea of a strong correlation between context and type of L2 knowledge produced seems 'rudimentary' (Housen et al., 2011: 86), as both types of knowledge can be developed in the two contexts.

Regardless of the type of knowledge produced in ISLA contexts and the mechanisms that are activated in SLA and ISLA, the perspective that will be adopted in this dissertation concerning the role of instruction in the acquisition of a second language is the one presented by Housen et al. (ibid), who claim that SLA is considered to be a socio-cognitive process which can be affected by external and contextual factors. As many researchers claim (Housen and Pierrard, 2005; Spada and Tomita, 2010), there seems to be general consensus that instruction does affect the acquisition of second language. However, what still remains to be explored is the way in which different types of instructional practices affect the process of Second Language Acquisition, which is precisely one of the aims of this research study.

2.2 The scope of intructed SLA research

Research in the field of intsructed SLA has focused its attention on attempting to identify the effects of instruction on several areas of SLA in order to inform researchers and teachers about the processes of second language acquisition. As Ellis (2008: 774) points out, "ISLA constitutes an area of enquiry where the concerns of researchers and teachers can be brought together". In other words, ISLA provides not only SLA researchers with relevant data which can be used to address several theoretical issues within SLA but also teachers who can use such data to incorporate new pedagogical practices into their methodologies.

As Housen and Pierrard (2005) thoroughly present, the study of ISLA derives from three main concerns. First of all, there is the need to describe the processes within ISLA, since, as they claim, ISLA might become the main form of SLA in the developed world. Secondly, research findings within ISLA have the potential to inform L2 language pedagogy and educational stakeholders about the instructional practices and interventions which seem to favour second language acquisition the most. Finally, the study of ISLA can provide invaluable information on relevant theoretical issues such as the nature of language learning and processing. The purposes of this investigation are clearly oriented towards the second and third concerns.

2.2.1 A framework for instructed SLA research

According to Housen and Pierrard (2005), the effects of instruction can be varied in the sense that they can affect many different subareas within the field of SLA. Thus, in an attempt to provide a clear and well-defined framework for ISLA research, they put forward the following scheme to establish links between the different areas of study within ISLA research and their effects on SLA (see table 2.1).

Table 2.1
SLA areas which can be affected by instruction (Housen and Pierrard, 2005)

Dimensions of SLA	**Components of SLA**	**Type of language knowledge developed**
Route of acquisition	Exposure	Implicit versus explicit knowledge (analysed and metalinguistic knowledge)
Rate of acquisition	Learners' propensity to learn	
Ultimate level of attainment	Learning processes and mechanisms (knowledge internalisation, knowledge modification and knowledge consolidation).	Declarative versus procedural knowledge

According to Housen and Pierrard's framework (Table 2.1), instruction may affect any of the three dimensions within SLA presented above. To start with, instruction can have effects on the learners' route of acquisition, meaning that second language learners in a classroom context might acquire linguistic features in a different order from learners in naturalistic settings. Secondly, instruction can also affect rate of language learning in the sense that instruction can lead to faster or slower acquisition of the L2. And finally, instruction can also have an influence on the ultimate level of proficiency achieved by second language learners.

In terms of the SLA components that can be affected by instruction, the authors point out three main elements: exposure, learners' propensity to learn and learning processes and mechanisms. First, instruction offers learners input and output opportunities, which are essential in second language acquisition, thus the type, amount and quality of that exposure may have an important impact on the learning of the L2. In addition, instruction can also influence learners' willingness to use and learn the target language by motivating them. Lastly, another of the elements which seems to be heavily influenced by instruction and which has received the greatest amount of attention in SLA and ISLA research is the learning processes that are activated in (I)SLA. L2

learning processes (i.e. knowledge internatisation, modification and consolidation) can be classified into the three dimensions of L2 performance: complexity, accuracy and fluency. In the first place, instruction might promote the internalisation of linguistic features which might lead to the development of syntactic and lexical complexity, resulting in the production of more elaborate output. Besides, instruction can also lead learners to modify their L2 knowledge and performance in order to become more accurate L2 users. And finally, instruction can also benefit L2 learners by promoting the consolidation of L2 knowledge, which may result in the production of more fluent L2 speech.

Regarding the type(s) of knowledge that might be developed by instruction, Housen and Pierrard (2005) present two main distinctions: implicit/explicit knowledge and declarative/procedural knowledge. According to SLA literature, implicit knowledge is abstract knowledge about the language which has been acquired subconsciously and incidentally, whereas explicit knowledge, which can be classified into analysed and metalinguistic knowledge, is knowledge about the language learned intentionally. Analysed knowledge involves a conscious understanding of how structural features work, while metalinguistic knowledge is defined as knowledge about the language (Ellis, 2006). As for the distinction between declarative and procedural knowledge, SLA literature based on skill-acquisition theories, which see language learning as any other type of learning, define declarative knowledge as knowing 'that' or as 'factual knowledge' about the language (DeKeyser, 1998: 48) whereas procedural knowledge or 'automatic knowledge' is the result of proceduralisation or automatisation of declarative knowledge (Ellis, 2008: 480). According to skill-acquisition theories, proceduralisation of declarative knowledge stems from practice, which is understood as attempts to use language meaningfully to communicate (Ellis, 2008).

In brief, Housen and Pierrard's overview (2005) of the SLA areas which can be affected by instruction provide a clear framework in which to categorise the impact of instruction on second language learning. The scope of this dissertation will primarily focus on studying the rate of acquisition, exposure and the learning mechanisms or processes which take place in the learning and development of oral production skills in English in two different instructional contexts (CLIL+EFL and EFL settings).

2.3 Learning contexts in instructed SLA

Learning contexts in SLA and instructed SLA have been often classified into dichotomous terms (e.g. naturalistic vs classroom contexts) which neglect not only variation within contexts but also contextual factors such as learning opportunities, amount of input provided or extra-mural exposure to the foreign or second language which can affect the process of language learning as well (Housen et al., 2011). Taking that into account, this section will attempt to provide a clear and thorough description of the learning contexts using Housen et al.'s descriptive framework of ISLA contexts.

ISLA learning contexts can be researched at three levels: the *learner's individual learning context*, the *curricular context* and the *extra-curricular context*. The learner's individual learning context includes individual factors such as the learner's aptitude, preferences or motivation and also wider social traits like the learner's social networks and interactional practices. The curricular context is made up of the school's pedagogical lines of work and the language policies implemented which in many cases are imposed by educational authorities. At this research level, factors such as curriculum design and pedagogical approaches implemented gain relevance since they will determine the activation of certain foreign language learning processes and mechanisms. In addition, according to the school's traits and stakeholders' policies, two

sub-types of learning contexts can be identified at the curricular context level: *the language-content classroom* and *the language-subject classroom.* The language-content classroom context is seen as a form of bilingual education in which the foreign or second language is used as the medium of instruction in content subjects, while the language-subject classroom context focuses on the foreign/second language as the object of study. Finally, the extra-curricular context level comprises the sociolinguistic, demographic, cultural and institutional conditions inside and outside the school. Contextual factors considered at this level play a crucial role in the interpretation of ISLA research findings as they might provide a feasible explanation of, for instance, why successful pedagogical practices implemented in a multilingual country cannot be expected to have the same outcomes in a monolingual country. At this level, two more sub-type levels can be distinguished: the *school level* and the *community level*. In the context of ISLA, the school level might include opportunities to have informal exposure to the target language by means of native speakers in the staff whereas the community level context is made up of factors such as amount of extra-mural exposure to the target language or the sociolinguistic makeup of the community.

In this dissertation, two curricular contexts are going to be studied and compared: the language-subject classroom (EFL instruction) and the language-content classroom (CLIL instruction) in combination with the language-subject classroom (EFL instruction). This study presents the classroom as a space where factors such as pedagogy, input, tasks and meaningful language use, among others, are manipulated in an attempt to create learning opportunities for learners. In ISLA, the classroom becomes the main, or in many cases the only, environment in which learners are exposed to the target language (TL). However, as stated above, external factors like extra-mural exposure to the target language and individual factors are also crucial for the

understanding of ISLA research findings. Because of this, the interpretation of the results in this study will be framed, when possible, within the three levels of ISLA contexts, especially the curricular level. To do so, Ellis' theories (1990, 1997, 2005) and accounts of ISLA will also be considered in an attempt to create a solid theoretical framework in which to interpret the findings of this study.

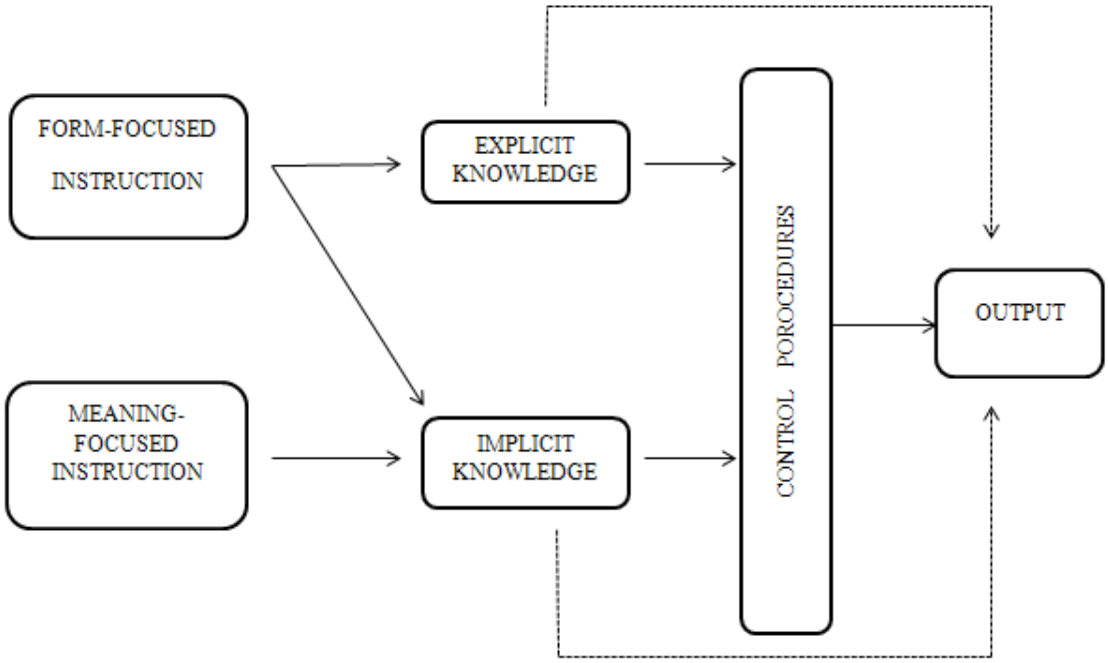
2.4 Ellis' theory of Instructed Second Language Acquisition

Rod Ellis' interest in foreign language learning pedagogy and the impact of pedagogical practices on the process of foreign and second language learning make him one of the most prolific researchers in the field of instructed Second Language Learning/Acquisition. Consequently, this section will focus on the theory of instructed second language acquisition he put forth (1990 and 1997), as well as on the language pedagogy principles he proposed in his 2005 article.

2.4.1 An integrated theory of Instructed Second Language Acquisition (Ellis, 1990)

In 1990 Ellis proposed "An Integrated Theory of Instructed Second Language Acquisition" (see Figure 2.1). As he highlighted, the purpose of the theory was two-fold. Firstly, the theory aimed at providing a set of statements or hypotheses (later re-labelled as 'principles' in his 2005 article) about classroom L2 learning. Secondly, the theory was also an attempt to "make an account of classroom L2 learning that is relevant and accessible to teachers" (Ellis, 1990: 174).

Figure 2.1
A model of instructed L2 acquisition (Ellis,1990).



The theoretical framework in which Ellis placed his integrated theory of instructed second language acquisition in 1990 was strongly based on cognitive learning theories. However, as he acknowledged, cognitive theory on its own was "unable to provide an adequate explanation of how [L2] knowledge is acquired in the first place" (1990: 175). Because of this, Ellis also drew on linguistic theories and approaches to language learning such as Universal Grammar and Accessibility Hierarchy in order to account for his theory of ISLA. As he presented in his 1990 work, his theory recognises that both linguistic and cognitive accounts are needed to explain how learners use instructional input to construct their interlanguages. That is why his initial theory was labelled as 'integrated' theory.

Ellis' (1990) integrated theory of ISLA is based on the principle that " L2 knowledge is differentiated" (p. 184). According to his theory, implicit and explicit L2 knowledge are stored in different parts of the brain and they can be accessed in various

degrees of automaticity. He adopts a non-interface position, meaning that explicit and implicit knowledge are seen as dichotomous rather than continuous, which is in accord with Krashen's Monitor Model (1985) which distinguishes between acquired knowledge (implicit) and learnt knowledge (explicit).

The distinction of explicit and implicit knowledge is crucial for the understanding of Ellis' integrated theory of ISLA, as he posits that different types of instructional input will result in different kinds of knowledge. He claims that explicit knowledge is derived from form-focused instruction, which is based on the teaching of linguistic features, while implicit knowledge derives mainly from meaning-focused instruction which is based on the principle that language needs to be used in context to fulfil a communicative purpose. He also claims that opportunities to practice the L2 in meaning-focused instructional contexts provide learners with the conditions needed to activate control procedures that will help them automatize their L2 knowledge which may result in greater accuracy and fluency in the learner's L2 output.

At this point of the theory and contrary to other theorists who adopted a non-interface position, Ellis defends the position that implicit knowledge can also be acquired in form-focused instruction through the learning of 'lexicalised sentence stems' or formulaic language that will be used later on by the learner to hypothesise about the L2 system. Additionally, Ellis' theory is also characterised by the importance given to explicit knowledge since, as he suggests explicit knowledge functions as an 'acquisition facilitator' of implicit knowledge for it helps the learner notice linguistic features in the input which might be subsequently incorporated into the learner's interlanguage system.

2.4.2 A theory of Instructed Second Language Acquisition (Ellis, 1997)

The first modification Ellis' theory of ISLA went through in 1997 is the omission of the word 'integrated' in the term used to label the theory, even though he did not explicitly mention the reason why he did so in his 1997 work, it is clear by his account of the theory that linguistic factors were not given the same degree of importance as in his 1990 proposal. Besides terminological aspects, the theory posited in 1997 was characterised by adopting a weak interface position regarding the distinction between implicit and explicit knowledge (instead of a non-interface position as presented in 1990) and by elaborating on various areas of language learning in instructed second language acquisition, namely L2 knowledge types, the learning of implicit knowledge and the role of explicit knowledge in instructed SLA.

As for the position Ellis adopts regarding the relationship between implicit and explicit knowledge, his new version of the theory supports, as previously mentioned, the weak interface model on the basis that " explicit knowledge derived from formal instruction may convert into implicit knowledge, if learners have reached a level of development that enables them to accommodate new linguistic material" (1997: 115). Despite this, he rejects the idea that all knowledge starts being declarative, which would be one of the assumptions of a strong interface position, by arguing that "the process of learning a language would become impossible if every rule out of the thousands that comprise the grammar of a language had to be first learnt as explicit knowledge" (1997: 113).

Ellis' (1997) theory proposes that L2 knowledge can be divided into 4 categories (see Figure 2.2) according to the types of knowledge (i.e. implicit or explicit) and the degree of cognitive effort required by the learner to access it (i.e. controlled or automatic processing). This new classification of L2 knowledge stems from the idea of

control posited in his previous work, which suggested that learners could access L2 knowledge in varying degrees of automaticity according to whether acquired L2 knowledge was new or old.

Figure 2.2
Types of L2 knowledge (Ellis, 1997)

Type of Knowledge	Controlled Processing	Automatic Processing
Explicit	A A new explicit rule is used consciously and with deliberate effort.	B An old explicit rule is used consciously but with relative speed.
Implicit	C A new implicit rule is used without awareness but is accessed slowly and inconsistently.	D A fully learnt implicit rule is used without awareness and without effort.

As seen from Figure 2.2, explicit knowledge of the language is accessed slowly if the rule or linguistic item is new and has not been practised enough so as to be accessed rapidly through automatic processing. Likewise, access to implicit knowledge will be relatively slow if the knowledge is new and has not been fully internalised. Only when implicit knowledge is fully internalised, will it be accessed rapidly and without effort by the learners.

Another of the aspects Ellis revisits and elaborates on in his second version of the theory presented in 1997 concerns the acquisition of implicit knowledge. In his first account of the theory, Ellis (1990) claimed that in order for input to become intake, two conditions had to be fulfilled. First, "the learner had to attend subconsciously" (p. 190) to the linguistic feature in the input and, secondly, the learner's cognitive level should match the effort required for the learning of that feature. His 1997 theory, on the other hand, suggests "that input can become implicit knowledge when the learner performs the following operations: noticing, comparing and integrating" (p.119). In other words, linguistic features in the input can become implicit knowledge when learners notice

those linguistic features and compare them to the features they usually produce in their output. Only then will learners be able to incorporate the noticed features into their language system.

Finally, Ellis' 1997 theory of instructed SLA also expands on the role of explicit knowledge in developing implicit knowledge. His theory keeps defending the idea that explicit knowledge acts as an 'acquisition facilitator' due to its role in the operations of 'noticing' and 'comparing' presented above. As Ellis (1997) states, explicit knowledge may help the learner notice not only the linguistic features in the input but also the gap between the target form and their interlanguage system. In so doing, according to Ellis, explicit knowledge may lead the learner to engage in grammatical processing.

To sum up, Ellis' theory of instructed SLA (1990, 1997) constitutes one of the first attempts to account for the process of language acquisition in instructional contexts where input, exposure and output opportunities are quite limited in nature. Despite some limitations, his theory seems to provide a feasible framework for the results of this study to be interpreted as the effects and learning processes of both meaning-focused instruction (CLIL) and form-focused instruction (regular EFL instruction) will be deeply analysed.

2.4.3 Principles of instructed language learning (Ellis, 2005)

As Ellis acknowledged, his 2005 article represents an attempt to draw together the findings from SLA research with the aim of providing set of generalisations of instructed language learning (see Figure 2.3) in which language pedagogy and teaching practices could be framed. By doing so, Ellis seems to revisit and revive the origins and rationale of SLA research when research and the actual practices of language teaching were closely connected (Ellis, 2010).

Figure 2.3
List of Ellis' principles of instructed language learning (2005).

1. Instruction needs to ensure that learners develop both a rich repertoire of formulaic expressions and a rule-based competence.
2. Instruction needs to ensure that learners focus predominantly on meaning.
3. Instruction needs to ensure that learners also focus on form.
4. Instruction needs to be predominantly directed at developing implicit knowledge of the L2 while not neglecting explicit knowledge.
5. Instruction needs to take into account the learner's 'built-in syllabus'.
6. Successful instructed language learning requires extensive L2 input.
7. Successful instructed language learning also requires opportunities for output.
8. The opportunity to interact in the L2 is central to developing L2 proficiency.
9. Instruction needs to take into account individual differences in learners.
10. In assessing learners' L2 proficiency it is important to examine free as well as controlled production.

According to Ellis (2005), the principles are mainly based on the computational model of L2 learning which views acquisition in terms of input, input processing and output, and fail to acknowledge, as the author himself claims, the role of social context and social relationships in the process of learning a language. Despite that, the principles Ellis proposes do present a reliable framework to discuss and interpret research within the field of ISLA.

To start with, Ellis suggests that teaching practices in a foreign language classroom should promote the development of what Skehan's (1998) called the 'exemplar-based system' and the 'rule-based system' (principle 1). Both systems are necessary in the learning of a language as each system has the potential to trigger the development of different areas of the interlanguage system: the exemplar based system promotes the development of fluency while the rule-based system may have an impact on accuracy and complexity (Ellis, 2005). As explained in one of the following sections on the characteristics of young language learners' language (2.5.1.4), this principle gains special relevance when teaching young language learners as, due to their low stages of cognitive development, they mainly rely on prefabricated and memorised chunks of language to produce output. It is not until they become older that explicit

teaching of grammatical rules starts to be effective in the process of foreign language learning.

In addition, Ellis also emphasises the importance of meaning-oriented instructional practices (principle 2) and insists on the idea that it is the pragmatic interpretation of 'focus on meaning' which leads learners "to view the L2 as a tool for communicating and to function as communicators" (p. 212). Teaching practices, then, need to enable learners to engage in genuine communication as active participants in order to create the optimal conditions needed for active learning. However, attention to form is also important in the process of language learning (principle 3). Authors like Schmidt (1990, 2001) or Long (1991), for instance, defend the position that noticing or attention to certain linguistic features within a communicative context may promote their learning. A more detailed account of meaning-focused instruction and form-focused instruction along with their links to the present study will be provided in section 2.5.2 on the effects of instruction types.

As for principle 4, Ellis claims that the development of implicit knowledge "should be the ultimate goal of instruction" (2005: 214) in order to help learners communicate fluently and confidently in the L2. To do so, instruction needs to provide learners with opportunities to participate in communicative tasks, as it is believed that implicit knowledge develops from meaning-focused instruction. Concerning explicit knowledge, Ellis revisits his theory of instructed second language acquisition and suggests that even though "the extent to which explicit knowledge can be converted into implicit knowledge remains controversial" (ibid: 215), there are a number of researchers who claim that explicit knowledge may promote the activation of acquisitional processes such as noticing which facilitate the incorporation of new knowledge into the learners' interlanguage system. On the basis of these accounts, Ellis justifies the role of

instruction aimed at the development of explicit knowledge within the process of language learning, although he stresses the idea that the development of implicit knowledge should be given priority.

Regarding the importance of the learner's 'built-in syllabus' (principle 5), Ellis defends the position, based on research findings, that both instructed and naturalistic learners follow the same order and sequence of acquisition. That might explain why, in some cases, instructed learners do not acquire what they have been taught, as the target feature may have been above their actual stage of development. Because of this, Ellis suggests that the effectiveness in the teaching of grammatical structures depends on whether the way in which they have been taught is compatible with the natural processes and orders of acquisition. In spite of this, Ellis recognises the difficulty for teachers to determine whether a student is ready for a specific target feature as well as the impossibility to provide individualised approaches to all the learners in a classroom and concludes by saying that the teaching of a target feature which is *slightly* above the learners' level might help them advance along their 'built-in syllabus rather than preventing acquisition.

Principle 6 suggests that successful instructed language learning requires extensive L2 input. According to research, amount and quality of L2 input has an impact on language learning and extensive exposure to good-quality input promotes the development of implicit knowledge, and consequently, the development of L2 learners as effective L2 communicators. Despite the importance of input in the learning process, Ellis, as well as many other researchers, defends the idea that output is also crucial in the learning of a language (principle 7). According to Swain's Output hypothesis (1985, 1993), being exposed to input alone is not enough to learn a language. As the author (1993) states, producing output affects not only fluency in the sense that it "permits the

development of automaticity" of students' linguistic resources but also accuracy as "producing language may force the learner to move from a semantic processing to a syntactic processing" (p. 159).

In addition, Ellis (2005) posits that the understanding of input and output on their own is senseless as both co-occur in oral interactions which, according to sociocultural theories, are the "matrix in which acquisition takes place" (p. 219). According to principle 8, then, interactions in the L2 provide the perfect context for learners to not only automatise already learnt linguistic knowledge but also create new knowledge by means of meaning negotiation, corrective feedback and output modification. Moreover, Ellis suggests that the promotion of interactions within instructed settings is more likely to take place through the use of 'tasks'[2] in which learners have to fulfill a communicative objective rather than through exercises.

Principle 9 acknowledges the variability in the rate of learning and ultimate level of achievement and claims that instruction needs to take into account individual differences. According to the author, instruction should be varied and flexible, meaning that it should cover different learning styles and strategies. This way, a wider range of learner-types would be able to benefit from instruction. In addition, Ellis also acknowledges the role of motivation in the process of foreign language learning. That is why he claims that teaching practices which reinforce the learners' intrinsic motivation are very much needed.

Finally, principle 10 states the importance of examining free production over controlled production in the assessment of L2 proficiency, since according to the author it is 'free production' (as opposed to multiple choice exercises, among other types of measurements) which "constitutes the best measure of learners' L2 proficiency" (p.

[2] The term task is interpreted as an "activity in which meaning is primary; there is some communication problem to solve; there is some sort of relationship to comparable real-world activities; task competition has some priority; the assessment of the task is in terms of outcome" (Skehan , 1998: 95)

221). In other words, the author defends the idea that free production (which is best obtained through 'tasks') are more representative of the learners' proficiency level than controlled production, as the kind of language learners find outside the classroom is more likely to resemble free production.

To sum up, this section has presented a brief overview of Ellis' work within the field of ISLA. In so doing, relevant aspects of the process of foreign language learning in minimal input contexts such as the classroom have been tackled and presented in an attempt to provide a deep understanding of the learning processes that underlie ISLA. In the next section, an account of factors that affect the effectiveness of instruction along with relevant literature on their impact will be presented in order to gain insights into the context and object of study of the present study: young language learners' oral development in two instructional settings.

2.5 Mediating factors in instructed SLA

According to Housen and Pierrard (2005) and de Graaff and Housen (2009), the effects of instruction are mediated by three factors: type of learner *(who)*, type of instruction *(how)* and type of language taught *(what)*. In other words, the study of the (un)effectiveness of instruction in second or foreign language acquisition needs to take those three aspects into account in order not to misinterpret the outcomes of instruction. In the next subsections, these three factors and their effects on instruction will be described.

2.5.1 Type of learner

Regarding the first of the three factors, *the learner*, research suggests that the degree of effectiveness of instruction will be determined by aspects such as aptitude, motivation and learning strategies (Ellis, 2008). Age is also an important aspect to consider when examining the effectiveness of instruction since instructional practices may fail to achieve its assumed benefits if the cognitive demands they posit do not match the learners' cognitive maturity. Due to the importance of the learner as a mediating factor in the study of instructed SLA, here follows a detailed description of the population under study in this investigation, *young language learners* (YLL), which includes some literature review on their characteristics as foreign language learners.

Even though this dissertation acknowledges in many cases the importance of individual factors such as aptitude and motivation in the process of language learning, it will not be the focus of this section. The main areas of investigation within this section will be cognitive development and age instead.

2.5.1.1 Young Language Learners (YLL): Definition and characteristics

According to research (McKay, 2006; Drew and Haselgreen, 2008; Nikolov and Djigunović, 2011), young language learners are those pupils who are learning a second/foreign language during their schooling years in primary education. Hence, learners who are categorised as young language learners are those children between the ages of six and twelve who are learning a second/foreign language in a wide variety of educational settings, among which we find immersion programmes, content-based language and CLIL programmes, scheduled foreign language lessons or awareness programmes. As McKay (2006) points out, young language learners can be classified according to the status of the target language within their community. That is to say,

young language learners may be *foreign language learners* learning a language which is rarely spoken in their local community or they could be *second language learners* learning a language which is used in their community. For the purposes of this research and the characteristics of the population studied, the term *young language learners* will be exclusively used to refer to foreign language learners who are learning the foreign language, English in this case, in instructional contexts.

Research conducted on the effects of age on instructed foreign language learning seems to agree on the linguistic areas which are positively affected by an early start (i.e. in primary education). Cameron (2001) and Pinter (2006), among others, highlight that the most salient advantage of an early start concerns the capacity of the younger learner to acquire the phonological system of the foreign language, which, according to Brewster, Ellis and Girard (2004) would be called a low-order process. According to them, higher-order processes such as meaning relations, on the other hand, are better learned by older learners. In addition to that, Johnstone (2002) points out that learning a foreign/second language as a young language learner might, in the long run, trigger a more positive attitude towards language learning and a deeper understanding of language as a system.

2.5.1.2 YLL and cognition

Unlike adults, young (language) learners are in a "stage of constant cognitive, social, emotional and physical growth" (McKay, 2006: 6). Because of this and particularly due to their constant cognitive development, the study of the mental operations underlying the processes of foreign language learning are so complex to explore.

In terms of cognitive development (see Table 2.2 below), young language learners would be classified into the so-called 'pre-operational stage' and 'concrete operational stage'. The former is featured by the lack of logical thinking while the latter is characterised by the use of logical adult-like thinking in immediate situations (Piaget, 1926). According to Piaget's theory of development and learning, then, learners at the 'concrete operational stage' (stage where the participants of this study belong to) are not prepared to generalise logical adult-like thinking in abstract situations until they reach the age of eleven/twelve when they enter the 'formal operational stage'.

Table 2.2
Piaget's stages of development (taken from Pinter 2006)

Stage	Ages	Cognitive development
Sensori-motor stage	From birth to two years of age	Children learn through interaction with the environment around them.
Pre-operational stage	From two years to seven years of age	Children rely on perception and lack logical thinking.
Concrete operational stage	From seven to eleven years of age	Children's thinking begins to resemble logical adult-like thinking. This ability is restricted to the immediate context.
Formal operational stage	From eleven years onward	Children are able to think beyond the immediate context in more abstract terms.

Piaget's stages of development have been strongly criticised in research. As Pinter (2006) claims, the development of 'logical adult-like thinking' at operational stages is, at least to some extent, influenced by the educational practices and beliefs of the schooling system children are immersed in. Likewise, Cameron (2001) highlights the fact that the studies carried out to support Piaget's theory of development were not child-friendly and underestimated children's abilities, which might have affected the results obtained.

In spite of such criticisms, it is worth pointing out that some research within the field of ISLA in CLIL settings (Vallbona, 2009; Bret, 2011) does observe a turning point in the productions of learners learning a foreign language at the age of eleven, when learners would enter the 'formal operational stage'. In terms of foreign language learning, the type of mental operations learners are capable of carrying out at the 'formal operational stage' would be reflected in the emergence of syntactically complex language, through the use of subordination, and greater fluency. Bret's cross-sectional study (2011) on the effects of CLIL on students' oral production skills in grades 5 (10 year olds) and 6 (11 year olds) of primary education, for instance, showed that having been exposed to the same number of hours of CLIL instruction (105 hrs), CLIL seemed to have a much clearer impact on 6th graders than 5th graders. In particular, learners in grade 6 performed significantly better in fluency and syntactic complexity than their EFL peers in the same grade, while no significant results were found in favour of CLIL learners in grade 5. Likewise, Vallbona's study (2009) which also studied the effects of 105 hrs of CLIL on 5th and 6th graders' writing skills, among other aspects, also reported a greater number of significant differences in grade 6 than in grade 5 in favour of CLIL. In her study, Vallbona found that CLIL learners in grade 6 were significantly better in syntactic complexity, accuracy and fluency, whereas CLIL learners in grade 5 only obtained significant results in fluency.

2.5.1.3 The Critical Period Hypothesis (CPH) and YLL in ISLA

As Nikolov (2009: 2) points out "young learners are widely perceived to acquire languages in a qualitatively different way from adolescents and adults". As previously stated, young language learners seem to have some advantages over older learners in foreign language learning. However, it is also true that, as some researchers suggest,

older learners can compensate for an early start through the use of strategies and analytical learning methods which younger learners lack (Cameron, 2001; Lightbown and Spada, 2006).

Such discussions on the effects of age on foreign language learning are inevitably linked to the Critical Period Hypothesis (CPH) proposed by Lenneberg (1967), which claims that natural language learning can only take place before puberty and which has been understood and interpreted by many as 'the younger the better'. In spite of that widely accepted interpretation, as DeKeyser and Larson-Hall (2005) claim the 'younger the better' only applies to certain types of language learning (i.e. implicit learning), which do not typically take place in instructional settings. In other words, the authors' understanding and accounts of the CPH are based on the idea that natural (foreign) language learning (i.e. from exposure only) is limited in older teenagers and adults. Foreign language learning in instructional settings seldom relies on mere exposure, that is why the authors emphasise that "the implication of critical period research seems to be that instruction should be adapted to the age of the learner, not that learners should be necessarily taught at a younger age" (p. 88).

As Nikolov and Mihaljevic-Djigunović's (2006) literature review of the interpretations of the CPH also shows, there seems to be little consensus on the effects of age on foreign language learning in certain academic communities. Long (1991), for instance, supports the idea that the CPH affects the whole process of language learning, whereas researchers such as Scovel (1988) restricted the effects of the CPH to pronunciation only. DeKeyser (2000) and Paradis (2004) suggest that the CPH affects learners' implicit learning skills preventing them from experiencing natural language learning at an older age. Singleton (2003), on the other hand, suggests that research findings which support the CPH, or a sensitive period for language learning, could also

be explained in terms of cognitive changes which affect all kinds of learning not only language learning. All in all, due to the lack of consensus on the extent to which age and maturational constraints affect foreign language learning at an early age, Singleton (2005: 280) concludes that the CPH "cannot plausibly be regarded as a scientific hypothesis […] of something which can be falsified […] or something that can be clearly confirmed or supported".

Even though the CPH has received a great amount of attention in SLA literature, "one wonders how these scholarly discussions are relevant to early foreign language programmes" (Nikolov, 2009: 5) or instructed foreign language acquisition. As some literature suggests (Cenoz, 2003; Muñoz, 2003; Brewster et al., 2004), most research on CPH has been carried out in second language contexts, where the L2 is spoken by the local community, not in educational contexts, where amount of exposure to the foreign language, proficiency level of the teacher and type of instruction, among others, might have a stronger influence on foreign language learners' ultimate level of attainment than maturational and age constraints

Research conducted by the BAF (Barcelona Age Factor) Project (Muñoz, 2006) is, among others, an exception to that in the sense that their study focused on the effects of age on foreign language learning in instructional settings in Catalonia. Their research examines the effects of different initial ages of learning English as a foreign language (8, 11, 14 and +18 years old) on the learners' proficiency level measured by a dictation test, cloze test, listening comprehension, written production and oral test. The results show that late starters outperformed early starters in most of the measures. In addition, the results also point out that while older starters benefited from their explicit learning mechanisms, early starters may have been prevented from their 'potential advantage' (p. 33) due to the limited amount of input received, which leads Muñoz (2006) to conclude

that initial age of learning may favour skills which are typically learnt implicitly and which are favoured by great amounts of input, whereas age might have an effect on rate of acquisition.

Larson-Hall's study (2008) on the effects of age on phonetics and morhosysntax in 'minimal input situations' (i.e. instructional settings), on the other hand, seems to yield dissimilar results. In this study, the author examines the performance on phonemic discrimination and grammaticality judgements tasks of two groups of Japanese students learning English as a foreign language. One of these groups was made up of 61 learners who started learning English between the ages of three and twelve, whereas the other was composed of 139 Japanese EFL learners whose learning age onset was between twelve and thirteen. The results reported in the study show that younger starters performed slightly better in the phonemic discrimination task and basic morphosyntactic abilities. Despite such surprising findings, the author calls for a cautious interpretation of the results, as the amount of input received by the learners is crucial in the understanding of the results. That is to say, early starters had received a greater amount of input due to their early start, which might explain why they out performed late starters. This view is partly supported by researchers like Nikolov and Mihaljevic-Djigunovic (2006), who acknowledge that length of exposure may favour foreign language acquisition, although longer periods of exposure to the target language do not guarantee higher levels of L2 attainment.

In brief, what can be concluded from research on the effects of age on foreign language acquisition is that previous interpretations and research findings of the CPH have to be applied with caution in 'minimal input situations', as most research has been conducted in naturalistic contexts. However, voices seem to arise defending the position that an early start will inevitably lead to a greater amount of input in the foreign

language, which should benefit foreign language learning, especially in instructional settings. All in all, consensus seems to be met on the fact that young learners learn a language differently from older learners. Taking that into account and in an attempt to describe the main characteristics of language learning among YLL, a description of Skehan's dual-model system (1998) will be provided and supported with relevant research studies in the next section.

2.5.1.4 Characteristics of YLLs' language system

According to Skehan's (1998) model of language learning and development *'the dual-mode system'*, language learners draw on two different but complementary systems when learning and using a language: the *exemplar-based system* and the *rule-based system*. The exemplar-based system is built upon unanalysed language chunks which are stored by the learners as units wholes (i.e. formulaic language) with a communicative function. The rule-based system, on the other hand, consists of grammatical rules and knowledge which learners use to produce creative output.

Due to the analytical nature of the rule-based system, the cognitive effort required by learners to access grammatical knowledge to construct utterances is considerable for learners who have not still developed a metalinguistic capacity. Because of this, young language learners draw on their exemplar-based systems when communicating (Lyster and Sato, 2013). According to McKay (2006) the use of formulaic language is crucial for young language learners to develop their interlanguage, as it helps them gain implicit knowledge of the language that can be gradually used to restructure already stored language chunks to create new ones.

As a study conducted by Muñoz et al. (2010) on the English oral output produced by 209 second-grade young learners from six European contexts (Croatia,

Italy, Poland, Spain, Sweden, and The Netherlands) showed, the amount of formulaic language used by the learners is dependent on their proficiency level. In other words, learners from contexts with a higher level of English (i.e. Sweden and the Netherlands) used fewer unanalysed chunks of language to complete the tasks than learners from countries with a lower level of English like Poland or Spain. Moreover, the study also points out that some learners, those with a higher level of English probably, were capable of restructuring already learnt chunks of language and combine elements from different sequences to create new and longer utterances that fulfilled their communicative purposes.

These findings seem to support the idea that young language learners' language will be mainly made up of unanalysed units of language. It is not until learners develop their analytical and rule-based capacities that they become capable of analysing those chunks and develop a rule-based system. As Wood (2002) puts forth:

> There is a certain amount of evidence of formulaic sequences being used as a learning strategy in children. It appears that first- and second-language acquisition in children is largely a function of attending to formulaic sequences in language input, adopting them for use, and later segmenting and analyzing them. The analysis may take place later partly as a result of neurological development and a resultant increase in analytic cognitive skills (p.4).

As Skehan suggests (1998: 89), "neither the rule-based nor the exemplar system is ideal separately", as a balance between the two systems is needed for successful language learning at a certain stage of cognitive development. In the case of ISLA, Skehan (ibid) suggests a combination of meaning and form-focused instruction for the development of the two systems. Type of instruction is indeed relevant in Skehan's dual system mode as an over development of one system only, as a result of certain pedagogical practices, might lead to deficient language use and knowledge.

To sum up, YLLs comprise one of the most complex population groups to study within ISLA owing to their unique characteristics as learners and their constant

evolution. That is why pedagogical practices need to be shaped accordingly. This section, then, represents an attempt to gain insights into the dimension of the learner in order to be able to understand the ways in which YLL's characteristics may mediate the effects of ISLA.

2.5.2 Type of instruction

Type of instruction constitutes another of the mediating factors which may account for the learning outcomes obtained in ISLA. Up to date, the most well-known classification of instruction types is that which categorises instruction into *form-focused instruction* (FFI) or *meaning-focused instruction* (MFI) according to whether the driving force in instruction is linguistic form or meaning and communication.

The more traditional definitions of these two types of instruction present FFI and MFI as two dichotomous terms. Form-focused instruction refers to those instructional models which view the TL as the object of study and consequently, tend to see language learning as an accumulation of grammatical rules and lexical items which the learner is expected to produce accurately after having practiced them. Meaning-focused instruction, on the other hand, emphasises meaning over form and provides learners with great amounts of (comprehensible) input which the learner is expected to process and assimilate in order to communicate in the TL.

The effects of FFI in language learning have been widely studied in ISLA research. In Ellis' (2008) review of FFI studies, he concludes that "grammar instruction can be effective in enabling learners to progress along the natural order more rapidly" (p. 863). In addition, Norris and Ortega (2000) also show that explicit forms of instruction are more effective that implicit types of instruction. However, in spite of the positive findings in favour of FFI obtained in these pieces of research, Ellis (2008),

points out that the effects of FFI "may only be temporary" (p. 866) if learners are not exposed and asked to participate in interactions in which the target form is used meaningfully.

As for the impact of MFI on language learning outcomes, it is worth mentioning some of the conclusions drawn from the vast amount of research conducted in Canada on immersion programmes. One of the lessons to be learnt from these programmes is that learners "seldom achieve the same high levels of competence in speaking and writing as they achieve in comprehension" due to the "lack of opportunities to engage in extended discourse" (Genesee, 1994: 4-5). Likewise, Cummins (2000) refers to the poor development of productive skills in immersion programmes as a consequence of the "paucity of classroom opportunities". Moreover, authors like Lyster (1998) and Ellis (2008) point out that many pieces of research show that learners that have been in immersion or content-based programmes over long periods of time show problems in producing accurate language. That is why some researchers (Swain, 1998; de Graaff et al. 2007; Lyster, 2007; Muñoz, 2007; Pérez-Vidal 2007, García-Mayo, 2011, Basterrechea, García-Mayo and Leeser, 2014) defend the use of an approach which integrates both FFI and MFI with the aim of promoting a better learning of the TL.

Both forms of instruction, as stated above, have been largely criticised in research. Strong forms of FFI, referred to as Focus on Forms by Long and Robinson (1998), are based on the assumption that learners will learn the language through the accumulation and controlled practice of grammatical rules and lexical items. Likewise, strong forms of MFI have been criticised for assuming that language learning in the classroom is purely incidental or implicit (ibid). In an attempt to conceal both positions in instruction, Long and Robinson (1998) propose the adoption of Focus on Form (FonF), which consists in allocating the learners' attentional resources on linguistic

form when a language feature interferes with communication. FonF acknowledges the importance of both forms of instruction in order to enhance language learning in ISLA. Concerning the effects of FonF on foreign language learning, Harley (1998), for instance, reports positive results of the systematic use of focus on form activities targeted at the teaching of grammatical gender in French to 2nd graders in six French immersion classes. The findings show that learners that had received focus on form instruction were more successful in 'item learning', and consequently more accurate, than learners that had not received the treatment. However, as the author also points out, the experimental groups failed to generalise and apply that learning to unfamiliar words. Similarly, Lightbown and Spada's (2000) study, in which they studied the English oral output of 100 French young learners of English after an intensive ESL course which adopted a communicative approach to the learning of English, reports positive effects of FonF practices such as corrective feedback on the accurate use of the progressive –ing and adjective noun order. According to the authors, the findings of their study suggest that:

> Accuracy, fluency, and overall communicative skills are probably best developed through instruction that is primarily meaning-based but in which guidance is provided through timely form-focus activities and correction in context (p. 443).

Basterrechea et al. (2014) also conducted a study on the effects of focus on form tasks on the noticing of linguistic features in CLIL settings. In particular, they analysed the impact of dictoglosses on the noticing of present and past tenses in a CLIL History lesson. The participants in this study were 16 teenagers aged 15-17 who had already received around 1520 hours of English instruction. Eight students reconstructed the text collaboratily while the remaning eight did it individually. First, the learners listened to a text related to the history topic being dealt with in class twice to take notes. Next, they reconstructed the text (either individually or in pairs). After, they listened to the text

again while taking further notes to complement their fisrt reconstruction of the text. Finally, they reconstructed the text again and answered a retrospective questionnaire to explore their metalinguistic reflections during the task. The results showed an increase in the accurate use of verb tenses between the first and the second reconstructions of the text. Moreover, the authors provided evidence to support Swain's output hypothesis (1985, 1993) regarding the effects of pushed output on noticing the gap between the target language and their interlanguage (Swain's Output hypothesis will be further described in the next chapter). As for the effects of collaborative work, no evidence was found to defend the position that collaborative work in this task is more beneficial than individual work.

As proved in this section, a thorough understanding of these three types of instruction, two if FonFi s seen as a sub-type of FFI, is crucial in order not to misinterpret the results obtained in ISLA research. This dissertation will attempt to contribute to the field of ISLA by analysing and comparing two groups of YLL who have been exposed to two types of instruction over a period of two years: FFI vs MFI + FFI combined. The characteristics of MFI, CLIL in particularly, will be developed in Chapter 3.

2.5.3 Type of language targeted

The effectiveness of instruction can also be affected by the language feature being taught. As Housen and Pierrard (2005: 10) point out, some language features are easier or 'more amenable' to teach than others. Thus, seeing the importance of type of language targeted in instruction for the study of ISLA, this section will provide a description of the two types of language targeted in each of the two instructional

settings under investigation: BICS (Basic Interpersonal Communication Skills) in FFI and CALP (Cognitive Academic Language Proficiency) in MFI (CLIL).

The terms BICS and CALP were first coined by Cummins (1979, 1981) in an attempt to account for the academic failure of fluent EAL (English as an additional language) learners. BICS refers to "conversational fluency in a language" (Cummins, 2008: 71) which is acquired and used in everyday social interactions. This is the type of language targeted in EFL instructional contexts with the aim of providing learners with the basic L2 knowledge to be able to communicate successfully in various informal social contexts. CALP, on the other hand, refers to the language needed by the learner to understand and express, both in written and oral modes, content knowledge within the school context (Cummins, 2008).

According to Mohan (1986), first and second language learners find it difficult to understand and express theoretical knowledge (i.e. academic knowledge) due to the type of language used to express that knowledge, namely CALP. Because of this, careful planning of content and language objectives in content-based approaches or immersion programmes needs to be done in order to help learners understand and verbalise content and academic objectives adequately. According to Snow, Met and Genesee (1989: 202), language gains a crucial role in content-based approaches, as "it provides access to content".

According to Cummins (1999), learners take round two years to acquire BICS, whereas they need between five to ten years to be fully competent in CALP, which might explain why learners show difficulties in understanding the language used in content subjects. However, that does not mean that the sequential order of acquisition of BICS and CALP applies to all contexts in the same way. As the author acknowledges (1999),

> The sequential nature of BICS/CALP acquisition was suggested as typical in the specific situation of immigrant children learning as second language. It was not suggested as an absolute order that applies in every, or even the majority of situations. Thus attainment of high levels of L2 CALP can precede attainment of fluent L2 BICS in certain situations (p.3).

In CLIL instructional contexts, the focus of attention is on CALP rather than BICS, since the language learners are exposed to and are expected to use is more related to theoretical knowledge (i.e. maths, history or science) than to everyday interactions. Moreover, CLIL includes the three aspects which, according to Cummins (1999), promote the development of CALP: cognition, academic content and language awareness. First of all, instruction in CLIL settings generally provides learners with challenging learning experiences which enable them to use their high-order thinking skills. In addition, instruction in CLIL integrates both content knowledge and language. And finally, CLIL instruction enhances the development of critical language awareness. Taking the characteristics of CALP and CLIL into account, there seems to be little doubt that learners exposed to CLIL instruction will achieve higher levels in CALP than learners exposed to regular EFL instruction. However, as a study conducted by Várkuti (2010) in Hungary shows, CLIL learners in secondary school performed significantly better not only in CALP but also in BICS tests which measured both language perception and production than learners that had only received EFL instruction. According to the author, CLIL students were more successful than non-CLIL students due to the 'meaningful nature of communication' in CLIL contexts (p. 76).

In brief, this chapter has constituted an attempt to provide a detailed account of the ISLA theoretical framework used in this dissertation for the interpretation of the results. In addition, this chapter has also highlighted how factors such as population group, type of instruction and type of language targeted in instruction can affect the

effectiveness of instruction. In so doing, relevant issues concerning the research questions and hypotheses posited in this investigation, which will be further developed throughout the subsequent chapters, have been tackled.

CHAPTER 3 Content and Language Integrated Learning

Chapter 3 aims at providing a detailed description of the main object of study in this dissertation: Content and Language Integrated Learning (CLIL). The chapter starts with an account of the reasons that gave rise to the emergence of CLIL within the European context and continues with a discussion of the definition(s) of the term. A description of immersion programmes in Canada and content-based instruction in the US, two of the main influences that shaped CLIL, are also presented. Furthermore, a summary of the features that successful CLIL programmes share along with a selection of relevant CLIL research in Europe is provided. Finally, this chapter concludes by summarising the linguistic outcomes of CLIL at receptive and productive levels of language use.

3.1 Emergence and evolution of CLIL in Europe

Back in the mid1990s the European Commission started to launch a number of recommendations to improve the quality of foreign language teaching and learning in the European Union. Those recommendations were introduced into the national educational authorities in the form of pieces of legislations or official documents such as the *Council resolution of 31 March 1995 on diversifying language learning and teaching within the education systems of the European Union* or the White Paper on education and training *Teaching and Learning-Towards the learning society* published in 1995.

In the 1995 Council resolution, for instance, the council emphasises the need to "promote, by appropriate measures, qualitative improvement in knowledge of the languages of the European Union within the Union's education systems" in an attempt

to materialise the concept of a multilingual and mobile European citizen. To do so, the 1995 Council resolution continues with the provision of the following measures: "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". Likewise, the White Paper on Education and Training (1995) focuses on the need to have a good command of at least two foreign languages in order to "develop everyone's employability and capacity of economic life" (p.13) within the EU. Moreover, it proposes that "secondary school pupils should study certain subjects in the first foreign language learned, as is the case in the European schools. Upon completing initial training everyone should be proficient in two Community foreign languages" (p.47).

Seeing the importance the European authorities were giving to the mastery of foreign languages and to the quality and promotion of innovative foreign language teaching methods, it did not come as a surprise to find that a decade after the publication and dissemination of official documents such as the one presented above, the implementation of programmes in which a foreign language was used to teach content subjects "was part of mainstream school education system in the great majority of countries at primary and secondary levels" in the EU (Eurydice European Unit, 2006: 21). Those programmes were and are referred to as *Content and Language Integrated Learning,* CLIL (first coined by Marsh in 1994), and their rapid spread and implementation throughout the EU created a completely new scenario for the teaching and learning of foreign languages which affected both academic and educational communities. According to Pérez-Cañado (2012), the origin and widespread use of CLIL can be explained in terms of *reactive reasons* (the need to seek for a solution to the deficient foreign language competence in some parts of Europe), *proactive*

responses (the enhancement of Europe's level of multilingualism) and the advancements and outcomes in SLA and teaching research in the last decades.

Regarding the term CLIL, Marsh (2002: 58) defines it as a "generic umbrella term which would encompass any activity in which a foreign language is used as a tool in the learning of a non-language subject in which both language and the subject have a joint curricular role". According to this definition, both language and content objectives have to be met in CLIL settings, however, the way in which this 'joint curricular role' is operationalized remains unclear as there are no precise guidelines which help CLIL practitioners put CLIL into practice. According to Nikula and Marsh (1998), that is precisely one of the strengths of CLIL as "the term is broad enough for the specific blend of content and language objectives to be made according to the specific objectives of the school in which the method is used" (p.16). However, even though the 'flexible' nature of CLIL might be seen as an advantage by some, other researchers warn that the capacity CLIL has to model itself according to the context is also a potential weakness (Coyle, 2007). According to the author, its main weakness lies in the wide variety of interpretations of that flexibility, which might derive in the implementation of teaching practices which move away from the 'curricular joint role' of language and content. In addition to that, it is also worth highlighting that Content and Language Integrated Learning (CLIL) as an educational approach was developed in Europe and is, therefore, European-oriented (Wolff, 2007), which means that CLIL has the capacity to fit not only into various school contexts with a specific set of goals regarding foreign language learning and teaching, but also into different countries, each of them with its own linguistic reality and objectives.

Nowadays, after two decades of CLIL implementation in Europe and as a consequence of that 'flexibility', authors like Cenoz et al. (2014) still call for a

clarification of the definition of CLIL. According to them, the definition of CLIL as an 'umbrella term' includes a wide range of educational practices (i.e. language showers, summer camps, different forms of content-based instruction, language across the curriculum projects, etc.) which makes it difficult "to identify the specific characteristics which they all share" (p. 4).

Taking all this into account and in order to avoid misinterpretations of the term CLIL in this study, CLIL will be defined and understood as a form of content-based instruction in which language objectives (linguistic functions, language structures, language skills and vocabulary) are clearly defined according to the linguistic and cognitive demands of the content subject chosen. This definition of CLIL differs from other definitions of immersion programmes or content-based instruction inasmuch as language is not only seen as a mere medium of instruction but also as a part of the content taught. Even though the differences between these types of instructional models and CLIL might not be clear-cut in many cases, attempts have been undertaken by many researchers to distinguish them. In the next section, an overview of the main sources of inspiration of CLIL, namely French immersion programmes in Canada and content-based instruction in North America, as well as an analysis of their similarities and differences will be provided.

3.2 Precedents of CLIL

3.2.1 French immersion programmes in Canada

As Georgiu points out (2012: 496), "CLIL is claimed to be a fusion of a number of theories and approaches" and one of the main influences of CLIL is, without any doubt, French immersion programmes in Canada, which started as an experimental kindergarten immersion class in St. Lambert (Quebec) in September 1965 in response to

the overwhelming failure of traditional approaches to the teaching of French as a second language. This experimental immersion class evolved to a sound programme which is now implemented all over Canada and other countries to improve the learners' proficiency levels in a second language. According to Genesee (1984: 41), the original objectives of St. Lambert French immersion programme were: (1) "To provide learners with a functional competence in both written and spoken aspects of French." (2) "To promote normal levels of English language development." (3) "To achieve academic success." (4) "To promote respect for French-Canadians, their language and culture."

Research on the numerous types of French immersion programmes has confirmed the success of these programmes. According to Perez-Cañado's (2012) overview of research findings in immersion programmes in Canada, learners were found to develop native-like receptive skills in French as well as good attitudes towards L2 learning. In addition, immersion programmes have also been proved not to have a negative impact on the learning of subject matter. However, as Genesee (1994) pointed out, one of the shortcomings of immersion programmes is that learners "seldom achieve the same levels of competence in speaking and writing as they achieve in comprehension" (p.4). In addition to that, research in immersion settings (see Cummins, 2000) also reports that the levels of grammatical competence are far from being native-like, due to the fact that, as Lyster (2007) highlights, content-based approaches like immersion programmes prime comprehension over production, preventing learners from developing their morphosyntactic system in the TL.

As for the relationship between CLIL and immersion programmes, there are important discrepancies among CLIL researchers. Lasagabaster and Sierra (2010), for instance, claim that the main similarity between CLIL and immersion models is that non-language subjects are taught in an additional language. However, as the authors

mention, they clearly differ in the status of the additional language within the local community, which in the case of immersion programmes in Canada was French, one of the two official languages in the country. In the case of CLIL programmes, on the other hand, the additional language is, in many contexts, rarely spoken within the community, which decreases considerably the amount of extra-mural exposure to the language chosen to teach CLIL. Because of this, Lasagabaster and Sierra (2010: 2) point out that "the reader might expect the word *foreign*" in the term *Content and Language Integrated Learning*. Somers and Surmont (2012), on the contrary, claim that there is no reason why CLIL could not be implemented through an official minority or heritage language, as the Eurydice report unit (2006) and Cenoz et al. (2014) suggest in their definitions and accounts of CLIL.

Theoretical discrepancies aside, another of the most relevant differences between CLIL and immersion Lasagabaster and Sierra (2010) put forward has to do with the profile of teachers. In immersion programmes, the teachers are usually native speakers of the language used in instruction, whereas in CLIL programmes the teachers rarely show a proficient or even advanced mastery of the foreign language, which undoubtedly has important effects on the linguistic model CLIL teachers provide their learners with.

Even though CLIL and immersion programmes are quite differentiated by a fraction of CLIL researchers (Coyle 2007; Lasagabaster and Sierra 2010; Pérez-Cañado, 2012), authors like Somers and Surmont (2012) defend the potential of CLIL to converge with other models. Similarly, Cenoz et al. (2014) are also against the trend of distinguishing CLIL from immersion and suggest the use of CLIL as "an umbrella construct" (p. 13) in which to place the wide variety of L2 or FL content-based teaching models, among which we find the different types of immersion and other forms of

content-based teaching. All in all, as Dalton-Puffer (2011: 183) puts forth, "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". Additionally, CLIL and immersion programmes are very often differentiated on the basis of the total amount of exposure to the TL they provide learners with. That is to say, programmes in which the foreign language is used to teach one or two subjects a week are usually referred to as CLIL programmes, while models in which half of the curriculum (at least) is taught in a second/foreign language are called immersion programmes.

Despite the theoretical and terminological discussions regarding the definitions and characteristics of CLIL, the influence of immersion programmes on CLIL cannot be denied, as both L2 teaching models stand on similar foundations. Similarly, content-based instruction (CBI) programmes in the US also helped to model CLIL into a teaching approach on its own. In the next section, the main features of content-based instruction will be presented along with an account of the similarities and differences between CLIL and CBI.

3.2.2 Content-based instruction in the US

Content-based instruction (CBI) also influenced the emergence of CLIL in Europe. These programmes were developed during the 80s in the US as a result of the success of immersion programmes in Canada (Lightbown, 2000) and are defined as "an approach to second language instruction that involves the use of a second language to learn or practise content" (Met, 1998: 35). The original aim of these programmes was to provide content instruction while developing the proficiency of the language in order to mainstream immigrant students eventually to classes where content instruction was

provided in the national language (Short, 1993). Thus, as Dalton-Puffer (2011) points out, content-based instruction was perceived as a way of offering language instruction to large numbers of immigrants whose level of the official language was quite limited. CBI is generally offered in primary schools, not in secondary schools where all learners receive regular content and language instruction regardless of their proficiency level (Short, 1993; Langman, 2003).

The choice of content-based instruction programmes in the US as opposed to other L2 models to enhance the level of English of immigrant students was deeply rooted in the principles of communicative language teaching. In content-based instruction, language was no longer the object of study but rather the medium of instruction as it was believed that a contextualised and meaningful use of the second language would enhance its learning. As Met (1998: 36) claims, "learning through content provides students with opportunities to use language as it functions in the real world: to communicate authentic meanings, for authentic purposes, and to accomplish authentic tasks". Research on CBI suggests that content-based programmes are a means for learners to continue their academic development while fostering language proficiency (Stoller, 2004). However, CBI has also been criticised for its "purely comprehension-based approach to language teaching" (Lightbown, 2000: 437). As Swain (1988: 68) claimed, content teaching does not necessarily enhance good language learning. According to the author, the language some learners are exposed to and use in CBI is far from being functionally varied (i.e. *functionally-restricted language*) which prevents learners from developing their interlanguage.

Langman's (2003) case study, for instance, examined the effects of ESL trained content teachers in a secondary school in Texas on the English development of limited English proficiency (LEP) students over a school year. LEP students at this school did

not receive language instruction at secondary levels, as it was believed that the content instruction offered by teachers who had received 15 hours of ESL training would be enough to enhance the learners' English level. The analyses of the classroom observation sessions and the instances of language produced by these students, however, revealed that after a course of a school year "the majority of learners showed little or no gains in their English language development" (ibid: 20). According to the author, these findings support the idea that focused language instruction is also necessary to enhance the proficiency English levels of LEP students, since, as shown in her study, content-based instruction on its own does not have a significant impact on language learning.

CLIL seems to have overcome that by emphasising the two areas of learning: content and language. In that sense, CLIL could be viewed as an improved version of CBI which highlights the importance of language teaching and learning (Ruiz de Zarobe, 2008; Järvinen, 2005; Lorenzo, 2007).

As shown in this section, CLIL seems to have learnt the lessons from immersion programmes and content-based instruction. However, there are many other aspects which determine the success of CLIL as an effective foreign language teaching model. In the next section, a description of the macro and micro-level features of successful CLIL programmes will be provided and discussed in an attempt to identify the keys to successful language learning in CLIL settings.

3.3 Successful CLIL programmes

Most CLIL research conducted in the European context highlights a series of characteristics and recommendations CLIL programmes, CLIL practitioners and CLIL stakeholders should follow in order for CLIL to display its full potential. Most of these

recommendations and features can be summarised in de Graaff et al. (2007) and Navés' (2009) pieces of work, in which the authors describe the characteristics of successful CLIL programmes as well as instances of effective CLIL performance. In this section, the main characteristics of successful and effective CLIL programmes will be presented and discussed.

The features the authors mentioned above put forward can be classified according to the level at which they are implemented or take place. Thus, the label 'macro-level features' will be used to describe those characteristics of CLIL programmes which are decided outside the CLIL classroom, while the label 'micro-level features will be employed to classify those aspects of CLIL programmes that have to do with the actual teaching practices that are visible within the classroom setting. See table 3.1 for a summary of the main features of effective CLIL programmes according to de Graaff et al. (2007) and Navés (2009).

Table 3.1
Macro and micro-level features of effective CLIL programmes.

Macro-level features	Micro-level features
Respect and support for learners' L1 Teacher's profile and training Long-term stable teaching staff Joint effort of all parties High expectations Optionality	Appropriate teaching materials A rich repertoire of active teaching behaviours Consistent integration of cognitively demanding academic content and tasks Exposure to input (i+1) Meaning and form-focussed processing Output production The use of learning strategies

3.3.1 Macro-level features of successful CLIL programmes

One of the key features Navés (2009)[3] presents as crucial for CLIL to be successful is the acceptance that L1 development contributes to successful foreign language learning. According to her and other researchers, L1 learning and development

[3] The author uses the term CLIL and Bilingual Education (BE) interchangeably.

has a positive impact on the learners' second or foreign language as "literacy in the primary language transfers to the second language" (p. 28). That is why she suggests that school contexts where CLIL programmes are implemented should also promote the development and growth of the L1 system and literacies so as to guarantee the success of CLIL.

Another important aspect which was not mentioned by Navés (2009) and which is worth highlighting at this point is the development of L2 literacies in CLIL and BE programmes. According to Pladevall-Ballester (in press), the fact that primary learners in CLIL programmes are not explicitly trained in L2 literacies may be problematic and may affect the learning outcomes of CLIL instruction. As Piske (2010) puts forth, primary learners exposed to BE should receive not only L1 but also L2 literacy instruction in order to guarantee L2 literacy development and the success of these programmes.

Secondly, Navés claims that the profile of CLIL teachers is a major determinant of the success of CLIL. According to her, CLIL teachers need to be bilingual in order for them to recognise the needs and problems learners experience when learning a second or foreign language as well as to provide learners with a good model of the TL. In addition, according to the author, the effectiveness of CLIL programmes will depend on the training CLIL teachers receive. That is to say, CLIL teachers need to receive quality teacher training courses in pedagogical and theoretical aspects of language acquisition in order for CLIL to succeed. The stability of the teaching staff involved in the implementation of CLIL as well as the continuity of the CLIL programme is also a key factor in analysing the effectiveness of CLIL. As the author points out, learners need a minimum of seven years in order to function adequately in a second language.

Concerning the role stakeholders play in the success of CLIL programmes, Navés (ibid: 31) defends the idea that effective CLIL programmes "require the joint effort of all parties involved: educational authorities, parents and teachers at both district and school level". Furthermore, according to the author, successful BE in the US all share the figure of a leader who proactively involves the teachers, the community and the private sector in the process of programme design. In this respect, parents have a lot to do since, in many cases, it is their interest in innovation and enriching teaching models which promotes the implementation of new pedagogical approaches.

In addition, another of the macro-level features which determines the success of CLIL programmes is the high expectations posited by the teachers on the learners' potential in learning, regardless of their individual differences or socio-economic backgrounds. Finally, the author also states that effective CLIL programmes should be optional, not imposed, as in the case of immersion programmes in Canada, where parents saw them as a right not as an imposition.

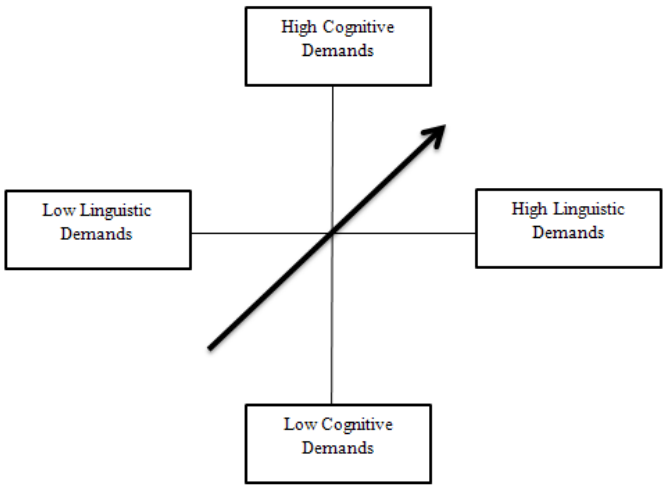
3.3.2 Micro-level features of successful CLIL programmes

Regarding the micro-level features that determine whether a CLIL programme is effective or not, Navés (2009) points out that CLIL learners need to be exposed to and use good CLIL materials which integrate both content and language objectives. As she claims, however, there is a lack of good quality CLIL teaching materials nowadays which may have a direct influence on the learning outcomes of CLIL programmes. Moreover, the author also highlights that successful CLIL programmes are implemented by teachers who display a wide range of good quality teaching behaviours such as describing tasks accurately, giving instructions clearly, keeping an adequate teaching pace or linking new information to learners' previous knowledge. Effective CLIL

teachers, as good teachers in general, have to be good communicators. As for the integration of content and language, the author says that CLIL programmes need to follow a careful planning regarding the cognitive and linguistic demands posited by the tasks used in the classroom in order to make content and language accessible to the learners. According to Coyle (2006), CLIL teachers are encouraged to use the adapted version of Cummins' matrix (1984) in order to analyse the type of tasks they are using to examine whether both content and linguistic development are fostered (see Figure 3.1).

As Navés (2009) states, CLIL programmes need to consistently integrate language and content. However, this has to be done in a way that respects the learners' learning stage in terms of cognition and language. To do so, Coyle (2006) suggests that the planning of CLIL tasks needs to follow an upward direction. In other words, the evolution of a CLIL programme which starts to be implemented for the first time in a school needs to move from tasks which are linguistically and cognitively low to tasks which little by little become more demanding in terms of language and cognitive effort (see Figure 3.1).

Figure 3.1
The CLIL Matrix adapted from Cummins 1984 (Coyle, 2006).



Regarding the actual teaching practices that take place in a CLIL classroom, de Graaff et al. (2007) identified a number of indicators for effective teaching performances in CLIL settings. According to the authors, effective CLIL teaching performances should include exposure to input (i+1), meaning and form-focussed processing, output production and the development of learning strategies.

As regards input exposure, one of the most relevant theories within SLA is Krashen's Input Hypothesis (1985) in which he claimed that learners need to be exposed to comprehensible input which is slightly above their current level (i+1) in order to acquire a language. Nowadays, nearly thirty years after the publication of this work, SLA researchers have provided substantial evidence which shows that exposure to input is not enough to function as a second/foreign language speaker. However, it is crucial for the development of the learners' interlanguage. According to Muñoz (2007), for instance, comprehensible and abundant input is not enough for acquisition. It has to fulfil other conditions. It has to be authentic, varied and it has to fulfil a communicative purpose. In CLIL settings, as de Graaff et al. (2007) point out, the input is very likely to accomplish these conditions, as the driving force in CLIL settings is meaningful language communication.

In addition to comprehensible input exposure, the authors (ibid: 608) claim that "input is only effective if the input is processed for meaning". Meaning-processing is crucial in CLIL settings, thus, the effectiveness of CLIL programmes will very much depend on the extent to which CLIL teachers promote meaning processing through the use of tasks and strategies which help learners understand input and generate new knowledge. Despite the importance of meaning-focussed processing in CLIL settings, second/foreign language learning also requires attention to linguistic form. According to de Graaff et al. (2007), CLIL teachers have to draw their students' attention to specific

language features in order to guarantee language learning. This is in line with previous research in other contexts. Doughty and Williams (1998), for instance, suggest that an approach based on *form* (not FormS) may be needed to move from a "communicatively effective language toward targetlike second language ability" (p. 2). As previously stated, CLIL programmes focus on both content and language, because of this, attention to form is essential for CLIL to be successful. In CLIL settings, however, attention to form is not supposed to rule the lesson; rather it needs to be dealt with in a way which is purposeful and easily included in a meaning-focused instruction setting. An example of how attention to form could be integrated into CLIL lessons is provided by Schmidt's Noticing Hypothesis (1990) in which he suggests that only noticed input may become intake and, consequently, part of the interlanguage system. The author points out that learners may notice forms in the input which are frequent, salient and highlighted by instructional practices. The fact that the Noticing hypothesis seems to be non-intrusive in meaning-focused instructional settings makes it extremely relevant for addressing linguistic form in CLIL settings.

In addition to the importance of input, meaning and form, output is also viewed as a crucial element for second language acquisition. As seen in Chapter 2, producing language may assist language acquisition, according to Swain's Output Hypothesis (1985, 1993). First of all, the author claims that by producing language learners are forced to use and apply their language knowledge, which might derive in the automaticity of the learners' linguistic resources. Secondly, language production may force learners' involvement in some kind of syntactic processing while attempting to articulate a message, which may result in grammatical development. Thirdly, according to Swain, producing language provides learners with the opportunity to test their hypotheses about the language and realize the gaps in their interlanguage system.

Finally, language production may also generate responses (i.e. feedback) from interlocutors which learners can take in and use to modify their output in the future. There seems to be little doubt that in settings such as CLIL where meaning is the main focus of attention, output and interactive opportunities should be abundant. However, as will be seen in the following sections, greater efforts to foster L2 ouput in CLIL should be made.

Finally, de Graaff et al. (2007) also highlight that the effectiveness of a CLIL programme might be dependent on the role of the teacher in assisting learners in developing a rich repertoire of learning strategies to help them handle content or language difficulties. According to the authors, strategies such as inferring the meaning of a word by the context, using prior knowledge, meaning negotiation or paraphrasing may play a crucial role in helping learners function in a CLIL classroom successfully.

Altogether, as seen in this section, the success of CLIL will depend on a wide variety of aspects ranging from stakeholders support and adequacy of teaching materials to the use of appropriate teaching practices which enhance and promote language learning and development. In the next part of this chapter, a selection of the most relevant research on CLIL in the European context will be provided and discussed.

3.4 Overview of CLIL research

The rapid spread of CLIL across the EU in the last twenty years can be attributed to "claims of the success of CLIL without substantial empirical evidence" (Cenoz et al., 2014: 14). So far, CLIL research findings seem to back up the supremacy of CLIL over regular EFL instruction. However, as Bruton (2011, 2013) and Pérez-Cañado (2012) point out, those findings need to be interpreted with caution as most studies show a lack of control over moderating variables (i.e. amount of exposure, learners' profile or

optionality of CLIL courses) which may have affected the results obtained. In this section, a detailed review of the most important studies conducted on CLIL in Europe at primary and secondary school levels will be discussed in an attempt to provide a comprehensive picture of the situation of CLIL research in the EU.

3.4.1 CLIL research in Europe

In the European context, Finland was one of the first countries where CLIL was introduced into state mainstream education successfully. As for research studies, the main areas of interests have been cognitive development, the effects of CLIL on L2 syntactic development, language choice in CLIL classrooms and affective factors such as motivation. Generally, findings seem to indicate that CLIL learners outperform their peers in L1 instruction in almost all the aspects mentioned above.

Jäpinnen (2005), for example, analysed the cognitive development of learners (N=335) attending CLIL science and mathematics lessons in English, French or Swedish and learners (N=334) who received instruction in these two areas in Finnish. The tests used to measure the learners' cognitive development included problem solving, classifications and conceptual similarities among others. The tests were adapted to what the learners had studied in the lessons and to their ages, which ranged from 7 to 15. The results of this research showed that CLIL and non-CLIL learners achieved similar test scores in terms of cognitive development. Despite these positive results, the author points out that younger learners (aged 7-9) had some difficulties with certain abstract concepts taught in the CLIL lessons, which suggests that "in the beginning in CLIL environments, teachers have to consider very carefully the contents taught through a foreign language to young learners" (p. 163). The second group of students, on the other hand, aged 10-14, showed a faster cognitive development than their peers

in the L1 context. Finally, very few differences in cognitive development were found between CLIL and non-CLIL learners in the third age group (aged 15). All in all, Jäppinen's research showed that CLIL, in general terms, did not have a negative impact on content learning. Furthermore, in some cases, learners' cognitive development seemed to benefit from CLIL exposure.

The impact of CLIL instruction on content learning is nowadays one of the main concerns of CLIL implementation in Europe. However, according to research in Finland CLIL does not affect content results negatively. Jäpinnen's (2005) study along with other pieces of research conducted by Merisuo-Storm (2007) and Seikkula-Leino (2007) seem to indicate that CLIL and non-CLIL learners achieve similar levels of content learning and that the development of L1 literacy is not affected by CLIL instruction.

Regarding syntactic development in CLIL settings, Järvinen (2005) analysed and compared the syntactic development of relativization in English of 90 CLIL students aged 7 to 11 (from grade 1 to grade 5) and 47 learners aged 9 to 11 (from grade 3 to grade 5) who did not receive CLIL instruction. The syntactic development of the learners was measured using elicited imitations of a number of sentences containing relativization. Before presenting the results, it is worth highlighting that, as the author acknowledges in her discussion of the results, CLIL learners in this study were "an exceptional group […] in terms of motivation and aptitude" (p. 451), since learners who enrolled in the CLIL courses were selected students. The control group, on the other hand, was made up of average school children. The results of the study, then, showed that CLIL learners in grades 3, 4 and 5 produced significantly longer, more complex and more accurate sentences in the elicited imitations than their counterparts in the non-CLIL groups. As Järvinen suggests, the results can be accounted for on the basis of the quantity and quality of input CLIL learners received. That is to say, CLIL learners were

exposed to a greater amount of input than non-CLIL learners, as CLIL learners from grade 3 onwards received both EFL (2 hrs/week) and CLIL instruction (4-5 hrs/week) while the control group was only exposed to 2 hours of EFL instruction.

CLIL settings have also been studied from a pragmatic perspective, Nikula's (2005) study, for instance, analysed the educational situations learners in EFL and CLIL classrooms are involved in and how these situations affect the learners' language choice. This study was based on six EFL lessons and seven CLIL lessons in secondary school. CLIL students in this study received all the subjects in English with the exception of Finnish as a mother tongue, whereas EFL students' exposure to English came from regular EFL instruction only, which means that CLIL learners received a significantly greater amount of English input than their peers in the non-CLIL group. In addition, CLIL was voluntary in the schools where this research was conducted. In other words, "students participating in it were more likely to be more interested in English and they were also likely to have better skills in English at the outset" (p.31). The observations and analyses of CLIL and EFL lessons showed that the learners' use of English in CLIL and EFL contexts differed in the degree of personal involvement displayed when using the language. In CLIL settings, learners were actively involved as English users in the classroom discourse in the sense that they displayed more initiatives and used their personal experiences. In the EFL settings, on the other hand, learners used English to talk about imaginary situations or material-related tasks rather than their actual personal experiences or interests. The fact that EFL learners were more personally detached from English than CLIL learners "reinforced the role of English as something to practice rather than to communicate" (p. 54) in EFL lessons. According to the author, EFL contexts should shift their attention from classroom materials to learners as active communicators capable of using English meaningfully.

As for the effects of CLIL on affective factors such as motivation in Finnish schools, Merisuo-Storm (2007) attempted to find out whether there was a difference between the learners' attitudes towards foreign language learning in CLIL and monolingual settings. The CLIL group was formed by 70 primary learners aged 10-11 who started learning English in grade 1 and the control group was made up of 75 learners of the same ages who had been exposed to EFL instruction for two years. The results of this research show that the answers provided by the learners in CLIL settings to questions on attitudes to foreign language learning were significantly higher than the learners' scores in monolingual settings. The author, then, concludes that in CLIL classroom the learners' attitudes towards foreign language learning are more positive than in EFL settings. Likewise, Seikkula-Leino's (2007) study, in which 116 CLIL learners in grades 5 and 6 of primary education and 101 non-CLIL learners in the same grades were asked to evaluate themselves as foreign language learners, showed that CLIL pupils demonstrated a stronger motivation to learn foreign languages. However, they also felt that their knowledge of the foreign language was worse than that of non-CLIL learners, probably due to the fact that CLIL settings are more demanding in terms of both content and language learning.

Sweden, on the other hand, despite its geographical proximity to Finland does not share the success of CLIL. According to Sylvén (2013), CLIL research in Sweden is "surprisingly scarce" (p. 305) and "findings from CLIL research do not match those obtained elsewhere" (p. 301). One of very few pieces of research conducted on CLIL in Sweden is Sylvén (2004) which analysed vocabulary acquisition in CLIL and non-CLIL settings in secondary school. The results of this study showed that CLIL learners outperformed non-CLIL learners in vocabulary proficiency. However, the study also reported that CLIL learners already obtained higher vocabulary scores at the onset of

the investigation, making it difficult to claim that CLIL had a positive impact on vocabulary acquisition. In addition to that, the author also found evidence to claim that CLIL learners had a greater amount of extra-mural exposure to English than non-CLIL learners, which might have favoured their results in the vocabulary tests administered.

Another example of CLIL research in Sweden is Lim Falk (2008), in which the author investigated, among other aspects, students' interactions in CLIL lessons in English and Swedish content lessons in secondary schools. Lim Falk based her analyses on eight recorded lessons of CLIL and six recorded content lessons in the L1. Her results indicate that there is less student interaction in CLIL lessons than in L1 content lessons. Furthermore, the participation in CLIL lessons was mainly visible when the CLIL teacher switched to Swedish.

The (very few) results of CLIL in Sweden seem discouraging as regards the implementation of CLIL as a solid and effective foreign language teaching model. That might be explained as Sylvén (2013) suggests in terms of contextual factors. First of all, the author puts forth that one of the reasons why CLIL seems to fail in the Swedish context is its non-recognition as a teaching model in the national curriculum. In her view, that lack of recognition of CLIL makes it difficult for schools to implement it as "there is no guidance" (p. 305). In addition, another of the explanations that may account for the failure of CLIL in Sweden is the scarcity of research on its effects which seems to be an obstacle for CLIL to adapt to the Swedish context. Thirdly, the author also points out that CLIL teachers are not required to credit their mastery of the foreign language or foreign language teaching methodology. On the whole, as seen from this brief description of the CLIL situation in Sweden, one can suggest that, as stated by Navés (2009) the success of CLIL depends largely on stakeholders' support and involvement.

The picture in central Europe differs from that described in Sweden. To start with, Austria is one of the countries in this part of Europe with the longest tradition in CLIL projects and research. Mewald (2007) was one of the first attempts to study oral production skills in CLIL and non-CLIL settings. In this study, the author used six oral tasks (a prepared monologue, an interview, a summary, a question-formation task and an unprepared monologue) to analyse the oral output of 36 learners in secondary school who received CLIL instruction and 36 learners of the same ages in mainstream schools with no exposure to CLIL. The analyses of the oral data were carried out using a combination of holistic and analytical measures. The fluency measurement employed showed that CLIL learners obtained higher results in fluency, however, the results were not significant. As for the learners' ability to produce continuous speech, the study reports that, again, CLIL learners outperformed their non-CLIL peers in terms of average scores. Regarding accuracy, results indicate that CLIL learners made fewer grammatical and lexical mistakes than their counterparts in the non-CLIL groups. Likewise, CLIL learners made use of a wider range of syntactic structures and grammatical categories. In general terms, then, CLIL learners performed better than non-CLIL learners in the six communicative tests. Despite this, the author emphasises on several occasions that " the complexity of the output (of CLIL learners) did not exceed the language one could have expected from EFL lessons" (p. 168).

Hüttner and Rieder-Bünemann (2007, 2010) also analysed the effects of CLIL on the levels of narrative competence in Austrian secondary schools. In their study, 44 pupils (22 learners in each group: CLIL and non-CLIL), were examined using macro-level and micro-level features of narrative competence. The results reported in the study indicate that CLIL learners outperformed non-CLIL learners in macro-level features of narrative competence such as realising all the plot elements: identification of the

problem, unfolding of the solution and resolution of the problem. In addition, CLIL learners also did better in micro-level features of narrative competence, namely anchor tense consistency and verb errors. On the whole, CLIL learners seemed to have a more advanced command over narrative skills than non-CLIL learners. However, as pointed out in previous research studies in this section, it is not clear that the advantage of CLIL learners stemmed from CLIL instruction only. As the authors mention (2010: 77), "the children in the CLIL groups were not randomly selected". In other words, CLIL learners enrolled voluntarily in the CLIL programmes which means that CLIL learners' motivation or language aptitude, as the authors acknowledge, might have had an effect on the results obtained. In addition, CLIL learners were exposed to greater amounts of English input than non-CLIL learners, which may have also had an impact on the results of the study in favour of CLIL learners.

Several CLIL initiatives have also been implemented in Switzerland. The Zurich Project 21 comprises one of the most remarkable initiatives in the country. The Zurich Project 21 consists of a collection of educational initiatives to enhance foreign language learning at primary school levels. One of the most popular initiatives was to teach 90 minutes a week of content subject in English in primary school (20 minutes every day approximately). Stotz and Meuter (2003) carried out a two-year evaluation to assess the outcomes of this initiative with regard to classroom interaction and productive skills. Classroom observation sessions revealed that there were some inconsistencies in the implementation of CLIL methodology. Moreover, the study also revealed that those teachers who had an advanced level of the English language were better at using English as a vehicle of communication, while those who had an intermediate level relied on explicit vocabulary teaching and linguistic explanations to a greater extent. As for the interaction patterns found, the authors pointed out that the interactional patterns

resembled those found in typical frontal EFL lessons. The learners used English to answer the teachers' questions, but the utterances were made up of one or two words. Spontaneous interactions, on the other hand, were always in German. In addition, little evidence for meaning negotiation was found. The authors conclude that those learners who had been exposed to successful models of content and language integration obtained better results in the listening and speaking tests than those who were instructed in deficient models of CLIL.

Várkuti (2010) also attempted to analyse the effects of CLIL in Hungary. As the author claims, "it is expected that CLIL will create a linguistically more challenging environment resulting in improved language learning" (p. 67). Thus, in order to see the presumed effects of CLIL, the author analysed the English language competence of CLIL (N=816) and non-CLIL (N=631) learners in secondary school in terms of their conversational and academic language use. CLIL learners received 5.30 periods of EFL instruction plus three content subjects in English a week, while the non-CLIL group was exposed to 5.30 periods of EFL instruction a week. The English competence was measured by means of two tests. Test I covered several aspects of conversational language use in addition to vocabulary and grammar needed in informal interactions. Test II included more cognitively demanding exercises which measured complex vocabulary and grammar. The findings of this study indicate that the communicative linguistic competence of CLIL learners was significantly better than that of non-CLIL learners. To start with, CLIL learners obtained higher mean scores in the test that measured conversational skills than their counterparts. Regarding the learners' academic linguistic competence, the CLIL group also performed statistically better than the non-CLIL group. Thus, the author concludes that CLIL environments are more effective in terms of foreign language learning. Despite the encouraging results, it is

worth highlighting that CLIL learners, as stated in previous studies, received a higher number of contact hours to English than non-CLIL learners within school which had undoubtedly an impact on the results obtained.

In Germany, Wannagat (2007) observed the interaction patterns found in a CLIL history class in English with secondary school learners. The results reported in the study were quite discouraging in terms of students' output opportunities, as some of the findings pointed out that 75% of the students' turns in a CLIL lessons were aimed at responding to teachers' questions, whereas only 18.8% of student talk initiated an interaction. These results fall far short from the expected outcomes of CLIL. Once again, output production in CLIL lessons seems to be one of the main shortcomings of CLIL. Zytadiß (2007), on the other hand, carried out a four-year longitudinal study to analyse the English linguistic competence and knowledge of CLIL learners and learners in mainstream education along with the effects of CLIL on L1 content levels. The results of this study revealed that the English proficiency levels of CLIL learners were higher than those in mainstream classes. In addition, no differences were found in content learning levels between CLIL and L1 learners.

In the Netherlands, Admiraal et al. (2006) comprises one the most solid pieces of research as regards the effects of CLIL or Bilingual Education (BE) on overall English proficiency. In this study the authors analysed the English proficiency levels of secondary school learners in terms of receptive vocabulary knowledge, reading comprehension and oral proficiency. They analysed the results obtained by two groups of learners: the BE group (N=584) and the mainstream group (N=721). The BE group had 50% of the total number of lessons in English whereas the mainstream group was only exposed to regular EFL instruction. The findings reported in the study showed that the BE group performed better than the mainstream group in receptive word knowledge.

However, the results also showed that the BE group had an advantage at the beginning of the study, which means that the two groups did not have similar levels in this measure at the outset of the study. As for reading comprehension levels, the BE group obtained a significantly higher mean score than the mainstream group. The data obtained from the oral test, which included tasks in which learners had to use language in real-life situations, were analysed holistically using measures for general oral proficiency and pronunciation. There results showed significant differences for the BE group in the two areas analysed. According to these results, BE seems to positively affect the English proficiency levels of secondary learners. However, as pointed out in previous studies, the amount of English instruction offered to learners in the BE group was much greater than in the mainstream group, which makes it difficult to claim that BE, rather than extensive exposure to the TL, has a positive impact on English proficiency levels.

Regarding the southern part of Europe, Spain in particular, it is worth mentioning that the amount of research and CLIL initiatives that have been carried out in the last decade surpasses that of many EU countries. As Coyle (2010: viii) says, "Spain is rapidly becoming one of the European leaders in CLIL practice and research". The autonomous regions in Spain that have produced the biggest amounts of CLIL research are Andalusia, the Basque Country, Catalonia and Madrid. In the next few pages, a detailed description of the educational policies as well as research projects on the implementation of CLIL and CLIL outcomes will be provided. The case of Catalonia, which is the geographical context of this dissertation, will be dealt with separately in the next section.

To start with, the wide variety of CLIL models found in Spain stems from the specific educational policies in each autonomous region. However, all the CLIL

initiatives that have been implemented in Spain for the last decade share the common aims of improving the communicative competence in English and fostering multilingualism.

In the autonomous region of Andalusia, for instance, the most important language policy in regard to foreign language teaching was the 2005 *Plan de Fomento de Plurilingüalismo* (Plan to promote plurilingualism). According to Lorenzo (2010), the plan put forth a great number of policies in order to enhance English proficiency levels (the creation of 400 bilingual primary and secondary schools, the use of native speakers, the provision of teacher training centres and in-service training in bilingual education, student mobility programmes, among others). The Bilingual schools that had been set up from this Plan started to implement CLIL in 40% of the school curriculum. Lorenzo et al. (2010) was one of the first studies to analyse the effects of CLIL in Andalusia. Their study compared the linguistic competence in English, French and German of learners in primary and secondary schools who had been exposed to CLIL in these TLs (N=754) and learners of the same age groups in mainstream classrooms who only received instruction in the TL in regular foreign language lessons (N=448). The overall results showed that CLIL learners clearly outperformed non-CLIL learners in reading, listening, writing and speaking skills. Once again, the results reported in this study point to the direction that abundant exposure to the target language in the curriculum (40% of the school curriculum was taught in the foreign language), rather than the actual methodological practices of CLIL, might account for the results obtained.

In the Basque Country, the autonomous government decided to start a *Plurilingual Experience* in 2003 which put forward, among other measures, a minimum of 7 hours of CLIL a week in the four years of compulsory secondary education (Ruiz

de Zarobe and Lasagabaster, 2010). Lasagabaster (2008) examined the implementation of CLIL in four different schools in the Basque Country. The total sample was 198 secondary school learners who had started to learn English at the age of 8. The learners were distributed into three groups: group 1 was made up of 28 learners who had only received EFL instruction (3 hours a week), group 2 consisted of 113 learners who had been involved in CLIL programmes for two years (4 hours of CLIL a week) and also received regular EFL instruction and group 3 was formed by 57 learners who had been exposed to 4 hours of CLIL a week over a period of one year and had also received the 3 mandatory hours of EFL instruction. The instruments used to analyse the learners' level in English were the Oxford Grammar Test and a writing and oral task. The findings reported indicate that the CLIL groups outperformed significantly their non-CLIL peers in all the tests. The author concludes that these results confirm the effectiveness of CLIL in a context where the foreign language is seldom used outside the school context. Nevertheless, he warns that CLIL in these schools was voluntary, so "students who chose the demanding CLIL programmes may have been more academically gifted and more motivated than their non-CLIL counterparts" (p. 38). In addition to the voluntary nature of CLIL in this study, it is also important not to overlook the fact that CLIL learners received more hours of English input than non-CLIL learners, which makes it difficult to claim that CLIL, on its own, is the only factor that accounts for the positive results obtained.

Lázaro and García-Mayo (2012) also examined the effects of CLIL on L1 use and morphosyntactic development in the oral output of 15 teenagers in a narrative task. In this study, the authors analysed the role of L1 use on discourse markers and repair sequences and several morphological and syntactic aspects (i.e. the use of correctly inflected verbs, the use of pronouns and the use of subordination) at two data collection

times. The results reported indicate a significant decrease in the use of L1 words to refer to discourse markers and to generate repair sequences. In regard to morhosyntactic elements, the analyses yielded significant improvement in all the aspected studied from T1 to T2. As for the comparison between CLIL and non-CLIL groups, the results showed a clear advantage of CLIL learners. However, as the authors acknowledged, the CLIL group was exposed to four more years of English instruction than the non-CLIL group. In addition, CLIL was optional, which means that CLIL learners might have been more taleneted language learners than non-CLIL learners.

In Madrid, two plans to enhance foreign language learning have been implemented since the 1990s (Llinares and Dafouz, 2010). The first one was the *MEC/British Council project*, in which the teaching through authentic materials across subjects and exposure to natural language, among other measures, were implemented in infant and primary schools. The second plan was the *Bilingual Project* which in the year 2009-10 was implemented in 206 schools in the region of Madrid. The schools under this plan are required to teach 30% of the syllabus in English, which means that learners are exposed to 5 hours of EFL lessons plus 3 hours of CLIL a week. In terms of research, the UAM-CLIL Project developed a research project to investigate the effects of CLIL on learners in secondary school who had received CLIL instruction under the MEC/British Council Project. Llinares and Whittaker (2007), for example, examined classroom productions (oral interviews and written texts) of 10 CLIL learners. Their findings suggest that CLIL learners achieved the aims stated in the syllabus in regard to their participation in discussions and their ability to respond to the information presented by the CLIL teacher. As for their writing skills, the results reported indicate that CLIL learners moved away from the use of memorised chunks of the language they

had studied in class. Moreover, the author observed a rise in fluency in the form of long answers in the interview tasks and longer written productions.

As seen from this literature review, the outcomes of CLIL in Spain as regards the learners' linguistic competence in English are impressively positive. However, despite the positive results obtained, critical voices have arisen questioning the rigour of Spanish CLIL research. One of the main flaws found in CLIL research within Spain is that amount of exposure to the foreign language is not kept constant between CLIL and non-CLIL groups, which means that CLIL learners usually receive a greater amount of exposure than non-CLIL groups. In addition, as stated in this section, CLIL is very often voluntary. That implies that learners who choose to take CLIL lessons are good language learners with a high motivation to learn foreign languages.

All in all, CLIL research is nowadays abundant in the EU. Spain, however, stands out from the rest Europe for its wealth in CLIL projects and research as well as for its self-criticism on the rigour of research. In the next section, a review of CLIL projects and research in Catalonia will be provided along with a brief description of the aims, scope and preliminary findings of the CLIL-SLA Project.

3.4.2 CLIL research in Catalonia

CLIL was recognised as a sound foreign language teaching model by the Catalan authorities in 2005 in the *Pla Experimental de Llengües Estrangeres* (Foreign Language Experimental Plan), which included, among other measures, the implementation of CLIL projects in Catalan schools. Up to 2011, the public calls to participate in this plan allowed a total of 1339 infant, primary and secondary state schools to start the training and implementation of CLIL. Seeing the growing interest in CLIL through this plan, the Catalan government decided to launch a new plan in 2011, *Pla Integrat de Llengües*

*Estrangeres (*Foreign Language Integrated Plan), which focused exclusively on CLIL initiatives. The first public call of this new Plan gathered 38 schools. In addition to these figures, it is worth highlighting that many other CLIL initiatives without funding have been undertaken in semi-private and private schools in Catalonia in the form of CLIL projects or semi/fully bilingual education (Navés and Victori, 2010).

Due to the rapid spread of CLIL within educational settings, the Catalan academic community started to carry out research studies to examine the effects of CLIL on foreign language learning. It is a fact that CLIL research in the Catalan context did not inform such CLIL initiatives, as the first comparative studies on CLIL and EFL instruction within the Catalan context did not appear until 2009. As Navés and Victori (2010: 34) point out in their review of CLIL literature in Catalonia, "research in our context is still quite scarce". Two of the first studies on the effects of CLIL on the English competence of learners are Vallbona (2009) and Vallbona (2011). In these studies, the author carried out a cross-sectional study to examine the effects of Science CLIL on the English competence of 5th and 6th graders. The data collected from the CLIL group (N= 51) was compared to the data collected from the non-CLIL group (N=51). The CLIL group had received 1 hour of CLIL a week over a period of two years, besides the 3 mandatory hours of EFL instruction a week, while the non-CLIL group had only been exposed to EFL instruction. The results of these pieces of research showed that CLIL seemed to have a greater statistical impact on 6th graders than 5th graders. The statistical analysis showed that CLIL learners in grade 6 significantly outperformed their counterparts in the non-CLIL group in reading, listening, lexical complexity, fluency and accuracy.

Subsequently, a subsample of the participants from these studies was analysed by Bret (2011) in regard to oral production skills. The results of this study reported

significant differences in favour of CLIL learners in grade 6 in syntactic complexity, accuracy and fluency. Likewise, Juan-Garau (2010) analysed the oral fluency of EFL learners (N=11) and CLIL learners (N=16) in secondary school. The results reported that CLIL learners speak more fluently than non-CLIL learners, as significant differences were found in the number of words per minute that CLIL and non-CLIL learners produced.

Miret (2009) also analysed the effects of CLIL on secondary school learners' English proficiency level and writing skills. The CLIL group (N=64) performed significantly better than the non-CLIL group (N=51) in some writing measures (content organisation, topic development, vocabulary and language use). As for the proficiency tests, CLIL learners also outperformed their non-CLIL counterparts significantly in the dictation and multiple-choice task aimed at assessing grammar knowledge. Similarly, Navés and Victori (2010) reported overall significant differences in favour of CLIL learners. In their study, the authors analysed the effects of CLIL on general language proficiency among primary and secondary learners (N=114) and compared the results to a group of EFL learners in the same grades (N=284). The results showed that CLIL learners were significantly better than their peers in the non-CLIL group in the proficiency tests. As for their study on writing skills (ibid), they also reported statistically significant results in favour of CLIL learners in fluency, syntactic and lexical complexity and accuracy. In addition to these results, the authors also report that 7th grade learners in the CLIL group obtained similar results to those obtained by non-CLIL learners one or two years older in the dictation, reading comprehension, grammar, listening skills and writing measures.

Nowadays, more recent CLIL research derived from the CLIL-SLA Project coordinated by Dr Pladevall-Ballester at Universitat Autònoma de Barcelona is being

published. This two-year longitudinal research project focuses on the effects of CLIL and EFL instruction on the learners' competence in English in grades 5 and 6 of primary school and attempts to overcome some of the methodological limitations found in previous CLIL research, namely the lack of control over amount of English exposure between CLIL and non-CLIL groups and the lack of data regarding the initial proficiency level in English of CLIL and non-CLIL groups. Up to the present, this research project stands out as the first attempt within the Spanish context to conduct a longitudinal piece of research which keeps amount of English input constant between CLIL and non-CLIL groups.

Vallbona's dissertation (2014), which is part of the CLIL-SLA Project, focuses on the analysis of the overall linguistic competence and writing skills of CLIL learners attending Science or Arts & Crafts lessons in English (N=97) and non-CLIL learners (N=107) throughout a period of two years in primary school. Her findings show that, keeping amount of English input constant between the two groups, CLIL learners in Science lessons outperformed their non-CLIL counterparts in listening and in the total number of adjectives used in their written productions. On the other hand, CLIL learners who received the subject of Arts & Crafts in English obtained statistically significant differences in writing in terms of complexity, accuracy and fluency.

In a subsequent study undertaken within the same research project by Pladevall-Ballester (in press) on the effects of CLIL subject selection on the listening and reading skills of primary school learners, it is shown that CLIL learners (N=96) who attended Science CLIL lessons obtained higher results than those (N=42) who received Arts & Crafts CLIL instruction. Findings from this study reveal that the Science CLIL group was significantly better than the Arts & Crafts group in the listening test after having received 2 years of CLIL instruction. As for reading skills, no significant differences

were found between the groups. Despite that, the Science CLIL group showed greater improvement than the Arts & Crafts group. Another important finding derived from this study relates to the impact of CLIL instruction on learners who have foreign language learning difficulties or problems (low achievers), average foreign language learners (mid achievers) and those who are talented foreign language learners (high achievers). According to the author, low and mid achievers' progress in listening throughout the two years of the investigation is higher than the progress displayed by high achievers in the same time period. Regarding the reading results, the data reported indicate that low achievers from the two groups of CLIL overcome their mid achievers counterparts, confirming that CLIL does not have a negative effect on learners who face difficulties learning a foreign language. According to the author, CLIL favours weaker language learners' L2 development, especially in listening.

Regarding the effects of CLIL on affective factors, Pladevall-Ballester's (2014) analysis of students', teachers' and parents' perceptions after one year of CLIL implementation in five primary schools showed that around 80 % of the students interviewed liked CLIL as they saw a purpose in learning and using the foreign language. Those who did not like attending CLIL lessons, on the other hand, said that they did not enjoy CLIL lessons because they had problems to understand some concepts and follow the teachers' explanations. The parents' answers resembled those given by their children inasmuch as most of them believed their children liked CLIL. Regarding teachers' perceptions, they all agreed on the fact that learners seemed to be more motivated to learn English since the implementation of CLIL. However, they also pointed out that, when dealing with challenging content, learners tended to feel frustrated due to their limited knowledge of the language. In addition, teachers also expressed that the lack of time to prepare the lessons and the lack of peer and

institutional support were their main concerns as CLIL teachers. As for the perception of the general learning outcomes of CLIL, interestingly, none of the three groups surveyed agreed on their answers. Learners reported that in CLIL lessons they were learning content while parents thought they were learning language. Only teachers said that learners were learning both content and language.

To sum up, this section has provided a detailed picture of the situation of CLIL in Catalonia. As seen from the literature review presented, the rapid spread of CLIL all over the territory and the increasing amount of research conducted reveal that, even though CLIL seems to have a lot to offer to ISLA, much needs to be done to improve the implementation of CLIL principles and the quality of CLIL research. In the next section, a summary of the learning outcomes of CLIL at receptive and productive levels of language use will be offered in an attempt to embrace the main findings obtained by CLIL research in Europe.

3.5 CLIL and language outcomes

This final section of the chapter attempts to provide a general and updated overview of the linguistic outcomes of CLIL at both receptive and productive levels of language use based on Dalton-Puffer (2011) and Ruiz de Zarobe's (2011) pieces of work. According to Ruiz de Zarobe (ibid: 146), "contrary to research on other educational approaches, notably Canadian immersion programmes, the dichotomy between receptive and productive skills shows contradictory results in CLIL". As shown in this chapter, a comparison of research findings across Europe exemplifies some of these contradictory results.

According to the authors, the receptive language areas that are mostly affected by CLIL exposure are vocabulary and reading skills. Regarding listening skills, Ruiz de

Zarobe (ibid) is c autious in c laiming th at th ey a re p ositively a ffected b y C LIL instruction on the basis that more research is needed. One of the pieces of research that could c ontribute t o t he c onfirmation t hat l istening s kills a re f avoured b y C LIL instruction i s V allbona (2014) w hich, i n s pite of t he f act t hat a mount of E nglish exposure b etween the groups w as ke pt c onstant, r eveals t hat t he C LIL group outperformed statistically better than their non-CLIL counterparts in the listening tests.

As for productive aspects of language learning and use, several areas of speaking and writing skills are clearly favoured by CLIL instruction according to the authors. To start with, both authors claim that spontaneous L2 speaking skills are enhanced in CLIL settings, e specially f luency. Furthermore, D alton-Puffer (2011: 187) s teps f urther a nd says t hat "overall t he evidence i s r obust e nough t o w arrant t he v erdict t hat C LIL definitely fosters spontaneous L2 speaking skills". Even though CLIL research has very often pr ovided evidence t o s upport s uch a claim, r esearchers t hemselves ha ve acknowledged on many occasions that CLIL learners who participated in their studies started with a linguistic advantage over their counterparts and had more contact hours to English t han t heir non -CLIL p eers. Furthermore, C LIL research w ithin E urope ha s highlighted that there is little room for learners to participate in oral interactions in the CLIL c lassroom, which ma kes it d ifficult to th ink th at C LIL o n its o wn f osters th e development of learners' speaking skills. Altogether, one has the feeling that little has been acco mplished or clarified i n t erms o f t he e ffects o f C LIL o n s tudents' o ral production s kills i n t he last years. A s V an de Craen et a l. (2007: 71) stated " CLIL education leads to erratic results as far as speaking is concerned".

As r egards writing s kills, t he a uthors poi nt out t hat aspects s uch a s morphosyntactic an d l exical co mplexity as w ell as f luency are en hanced i n C LIL settings. A s f or a ccuracy, c ontradictory da ta ha s be en f ound. O n t he one ha nd, a s

Dalton-Puffer (2011) states, accuracy seems to improve through CLIL exposure, in spite of the focus on meaning in CLIL classrooms. On the other hand, however, according to Ruiz de Zarobe (2011: 146) "accuracy does not benefit much from the approach". Again, the data seems to be inconclusive on this aspect.

To summarise, this chapter aimed at presenting a thorough picture of the situation of CLIL practice and research in Europe. As suggested, CLIL seems to have overcome some of the limitations of previous L2 teaching approaches like immersion and content-based instruction and it is perceived as a powerful means to the enhancement of foreign language learning in the current school system in the EU. As for CLIL research, findings seem to confirm the supremacy of CLIL programmes over regular foreign language instruction. However, CLIL research is currently being revisited in an attempt to step further and find more reliable methodological designs which help researchers clarify what the real effects of CLIL are.

CHAPTER 4 The study of L2 English oral output: complexity, accuracy and fluency

This chapter aims at providing a detailed description of the three dimensions of L2 oral production which are analysed in this study: complexity, accuracy and fluency (CAF). It begins with a brief account of the emergence of the CAF construct and continues with the examination of the main challenges of CAF research in an attempt to identify some of the theoretical and methodological flaws from previous CAF literature. In addition, Skehan's (1998) *Trade-off Hypothesis*, Skehan and Foster's (2012) *Extended Trade-off Hypothesis* and Robinson's (2001) *Cognition Hypothesis* are presented and described with the aim of gaining understanding of the manifestations of CAF during L2 oral production. Finally, a description of the most relevant CAF findings with regard to the *(Extended) Trade-off Hypothesis* and the *Cognition Hypothesis* will be provided.

4.1 CAF: origins and evolution as research variables

The study of L2 performance and proficiency within instructed SLA was strongly linked to the notions of accuracy and fluency in the 1980s, when research on L2 language pedagogy explored the effects of learning and teaching conditions (i.e. the Audio-lingual method and the Communicative approach) on accurate versus fluent speech production (Housen and Kuiken, 2009; Housen, Kuiken and Vedder, 2012). Complexity, on the other hand, was added to the CAF triad a decade later by Skehan (1996), when he suggested that learning goals in task-based learning (TBL) should be categorised into three main areas: accuracy, fluency and complexity. As Housen and

Kuiken (2009) highlight, this was the first time the three elements of CAF were seen as the main dimensions of L2 proficiency. Since then, a great deal of research on the effects of age or teaching practices on L2 learning has used complexity, accuracy and fluency measures as dependent variables to analyse L2 language gains and development.

Nowadays, CAF has emerged as "the primary epiphenomena of the psycholinguistic processes and mechanisms underlying the acquisition, representation and processing of L2 systems" (Housen et al., 2012: 2). According to research, and as it was presented in Chapter 2, the three dimensions of CAF are strongly connected to cognitive processes which lead to transformations and adjustments within the L2 interlanguage system. Complexity, for instance, implies the *internalisation* of new L2 elements. Accuracy is associated with the process of *restructuring* or *modification* of L2 knowledge in order to comply with L2 norms. Finally, fluency is seen as a form of *automatisation* or *proceduralisation* of L2 knowledge.

Despite the interest that researchers have shown in the study of CAF, CAF research still faces challenging issues. Housen et al. (2012) identify some key aspects within CAF research which need to be tackled in order to establish CAF as a solid and uncontroversial research ground. These aspects, or challenges as the authors name them, are the definition and measurement of CAF, the identification of the processes and mechanisms underlying CAF, the connections and interdependency of CAF elements and the factors that affect the manifestation and development of CAF in L2 use and learning. In the following sections these aspects will be explored and discussed.

4.2 Definition and measurement of CAF

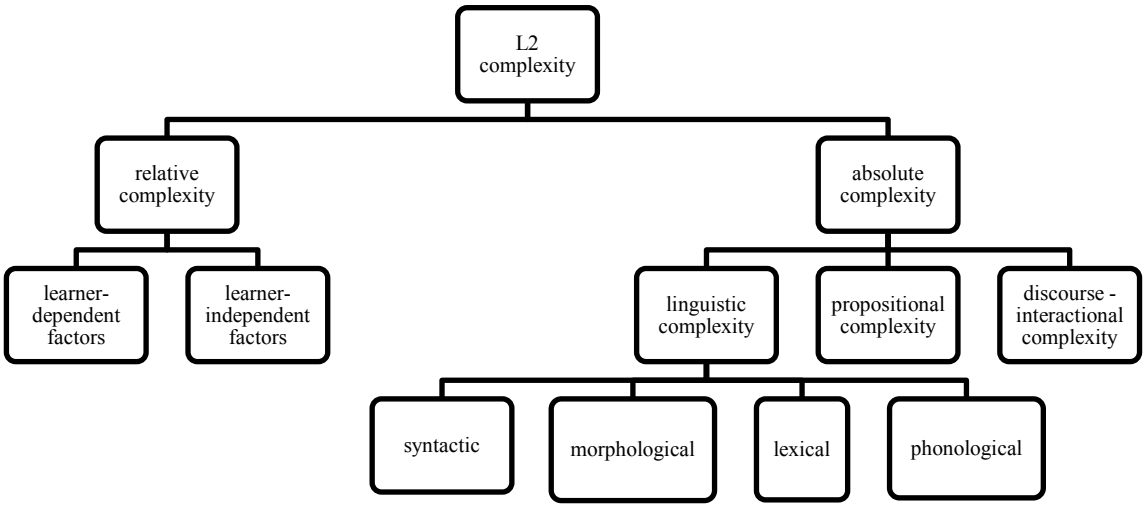
According to Housen et al. (2012), one of the challenges of CAF research is to provide findings which allow for the comparability of results across studies. Nowadays, CAF research is far from providing results which can be compared across research studies and its key findings are on many occasions inconsistent when compared to previous literature. One of the possible explanations that accounts for such inconsistencies is the failure to present clear definitions and explanations of the terms *complexity, accuracy* and *fluency.* Instead, researchers provide vague definitions of the terms or mere accounts of how these areas are measured. Thus, in an attempt to step ahead within CAF research, Housen et al. (2012) suggest the provision of clear and concrete definitions of CAF. In the next subsections, the three elements of CAF will be presented and explained in order to provide a clear account of the three areas that are used in this study to analyse L2 English learners' output. Additionally, this section will also present a description of the different subdomains that compose each of the three dimensions of CAF and the measures that can be used to analyse them.

4.2.1 Complexity

According to Bulté and Housen (2012), *L2 complexity* can be approached from a relative and an absolute perspective (see Figure 4.1). According to the relative approach, the notion of complexity is seen as difficulty (i.e. mental effort) in learning an L2 target structure, which may stem from either individual factors such as language aptitude or memory capacity (learner-dependent factors) or from factors like perceptual saliency or frequency of the L2 target structure (learner-independent factors). The absolute approach, on the other hand, views *complexity* as the number of elements that

make up a language feature or a language system. In this dissertation, complexity will be analysed from the absolute perspective.

Figure 4.1
A taxonomy of complexity constructs (Bulté and Housen, 2012).



As seen from Figure 4.1, absolute complexity consists of three components: linguistic complexity, propositional complexity and discourse complexity (see Table 4.1 for a definition of the components that form L2 complexity). In turn, linguistic complexity can be classified into syntactic, morphological, lexical and phonological complexity. This study focuses primarily on the two major components of linguistic complexity (grammatical complexity, syntactic basically, and lexical complexity) and propositional complexity. However, discourse-interactional complexity will also be dealt with in reference to the classroom observation sessions gathered in this study in order to further interpret the results obtained from the statistical analyses.

Table 4.1
Definition of the L2 complexity components according to Bulté and Housen (2012).

L2 COMPLEXITY COMPONENTS	
Linguistic complexity	Grammatical complexity: Elaboration, size, range, variation and breath of L2 grammar.
	Lexical complexity: Elaboration, size, range and breath of repertoire of L2 lexical items and collocations.
Propositional complexity	Number of information or idea units which a speaker/writer encodes in a given language task to convey a given message content.
Discourse-interactional complexity	Number and type of turn changes that learners initiate and the interactional moves and patterns they engage in.

In addition to the identification and account of complexity types, Bulté and Housen (2012) move further in their attempt to categorise the construct of complexity by providing the levels on which grammatical and lexical complexity can be examined. According to the authors, grammatical complexity can be studied from an abstract level, that is to say, research can look into the features of a system or structure in terms of the number of components that compose them or the relationship between the features of that system or structure. The observational level focuses on the language behaviour used for the manifestation of grammatical complexity. Finally, the operational level or statistical constructs refers to the use of analytical measurements that have been designed to account for the degree of complexity in a given language sample. In this study, propositional and linguistic complexity will be examined at the observational level.

Moving on into the dimension of syntactic complexity (the CAF area which has received the greatest amount of attention in research), Norris and Ortega (2009) identify five subdomains which can be measured: complexity via subordination, clausal complexification via coordination, overall complexity, subclausal complexity via phrasal elaboration and the sophistication and acquisitional timing of linguistic forms.

According to the authors, complexity via subordination is analysed by dividing the total number of subordinate clauses by the total number of units in a sample. The choice of the unit will be given by the type of data (written versus oral) and the proficiency level of the participants. The study of subordination is more appropriate with L2 intermediate level learners while the analysis of coordination, which is measured by dividing the total number of coordinate clauses by the number of clauses, is more adequate for the study of beginner L2 learners. According to the authors, this measure is uncommon in literature as most SLA research aims at studying intermediate or advanced level learners. Overall complexity, on the other hand, relates to the number of elements that compose a multi-clausal unit such as the T-unit, AS-unit or utterance and can be measured by using mean length measures of any type of unit. Subclausal complexity is another of the dimensions within syntactic complexity and it refers to the composition of the clause. This type of syntactic complexity can be measured using mean length clause measures. Finally, the sophistication and acquisitional timing of linguistic forms has been studied by a number of SLA researchers (Ellis and Yuan, 2005; Robinson, 2007) by means of raw frequencies of certain forms that are believed to be more sophisticated or later acquired (i.e. modal verbs, passives or infinitival phrases) in relation to the effects of planning and task complexity.

Due to the nature of the participants in this study and the type of oral data collected, the two main areas under investigation in this dissertation regarding syntactic complexity are coordination and subordination. As previously mentioned, the production of coordination is more frequent than subordination at beginner levels, which is why coordination was employed as the main measure for syntactic complexity. In addition to that, it was observed that some participants started producing subordination and that the amount of subordination increased over time. Because of this,

subordination was also included as a syntactic complexity measure. In regard to the actual measures and unit of analysis, Chapter 5 provides a more detailed description and justification of the measures employed.

As for lexical complexity, Bulté and Housen (2012) identify three main subdimensions: lexical diversity, lexical density and lexical sophistication. Lexical diversity refers to the number of different words in the output. The most common measures to examine lexical diversity are the number of word types, the type-token ratio and the Guiraud Index (which is a variation of the type-token ratio which controls for text length). The second subdimension of lexical complexity is lexical density and is defined as the proportion of lexical words in relation to the total number of words used. The most widely employed measure for the analysis of lexical density is total number of lexical words divided by the number of words. Finally, lexical sophistication refers to the use of infrequent or more advanced words and is measured using word frequency lists on which words are ranked according to their frequency or sophistication. The focus of this study within the area of lexical complexity will be the analysis of lexical diversity. To so do, different word types (i.e. nouns, verbs and adjectives) have been computed in relation to the total number of words in English used.

4.2.2 Accuracy

Accuracy refers to "the extent to which an L2 learner's performance deviates from the norm" (Housen et al., 2012: 4). From this definition, accuracy would seem to be the most straightforward dimension within CAF (Palloti, 2009), however, challenges may be faced when considering which type of norm to use as target norm. As Housen et al. (2012) suggest, on some occasions and depending on the participants, errors or 'deviations' "should be tuned to non-standard and even non-native usages" (p.4) fully

accepted within a particular community. These are aspects that need to be addressed when analysing accuracy. As Iwashita et al. (2008) indicate, coders sometimes disagree on what they consider to be errors. Taking this into account, some researchers suggest that the 'A' in CAF should be not only understood as accuracy but also as appropriateness and acceptability (Housen et al., 2012). Another challenge in regard to accuracy relates to the interpretation of the different accuracy levels learners display. According to Palloti (2009: 592), "accuracy per se is not a direct indicator of interlanguage development", particularly at beginner levels when errors are desirable and allow learners to develop their interlanguage by restructuring their L2 knowledge. U-shaped learning behaviours during the acquisition of some linguistic features (i.e. the acquisition of the English past tense), for instance, have proved that the presence of errors might be a more reliable indicator of language development than their absence at certain levels of proficiency.

As for the analysis of accuracy, Iwashita et al. (2008) distinguish two methods: global accuracy and specific accuracy. In the global accuracy approach all errors are considered, while in the specific accuracy approach only specific error types are computed. The most common measures to analyse global accuracy are the number of errors per 100 words, the total number of errors per unit or the number of error-free units. With regard to the specific approach, different ratios can be computed in which the correct use of specific linguistic features (third person singular, verb tense, plural markers or prepositions) are divided by the total number of words in which the feature is produced. According to the authors (ibid), the global accuracy approach is the hardest to conduct due to the difficulty in establishing coding consistency. In this dissertation both approaches have been employed in an attempt to provide a broader picture of the participants' accuracy levels. Global accuracy was analysed using the percentage of

error-free units whereas specific accuracy was measured by the percentage of correct verb forms.

4.2.3 Fluency

Fluency is defined as "the capacity to produce speech at normal rate and without interruption" (Skehan, 2009: 510) and it consists of three subareas: breakdown (dys)fluency, repair (dys)fluency and speed fluency. Breakdown (dys)fluency refers to the number, length and location of pauses. Repair (dys)fluency relates to the number of false starts, self-corrections and repetitions produced in speech. Finally, speed fluency is associated with the rate of linguistic units produced. In this investigation, the focus is placed on speed fluency due to the proficiency level of the participants (i.e. beginner level) and the characteristics of their L2 oral output, which is generally formed by long and abundant pauses and a great number of repetitions and false starts. Speed fluency is usually measured by the number of words or syllables per minute. The measure selected to examine speed fluency was speech rate in words, which is obtained by dividing the total number of words by the total task time in minutes. Additionally, an L1 word ratio was also used as an indicator of breakdown disfluency.

Even though the definition and scope of fluency seems uncontroversial, recent research has called into question whether some fluency measures, particularly length-based measures, are actually measuring fluency. According to Wolfe-Quintero, Inagaki and Kim (1998), all metrics involving mean length (i.e. mean length of utterance, mean length of T-unit, mean length of clause or mean length of phrase among others) should be understood as fluency measures, not as complexity measures as most literature presents them. Norris and Ortega (2009), on the other hand, defend the traditional position that length-based measures measure complexity. To back up their position, the

authors cite Oh (2006). In this study, the author examined the fluency levels of 78 essays in English written by advanced learners of English. The fluency measures employed were the total number of words, the total number of T-units, the total number of clauses, the total number of words per minutes spent writing and two length-based measures which were mean length of T-unit and mean length of clause. Oh submitted the results to a factor analysis and the results showed that the first four measures loaded very highly on one factor, while the two length-based measures loaded very highly on the second factor. According to Norris and Ortega's (2009) interpretation of these results, Oh's (2006) findings suggest that length-based measures such as number of words per T-unit or number of words per clause do not analyse the same dimension of L2 production as number of words, number of clauses, number of T-units and words per minute. According to Norris and Ortega (ibid), the results obtained by Oh (2006) support the position that length-based measures are complexity measures, not fluency measures.

All in all, as seen from this section and as has been acknowledged by many researchers, CAF research needs to clearly present and define the areas of L2 production that are being analysed in order to allow for the comparability of findings across research. This section has attempted to contribute to research by providing clear definitions of the CAF construct and their scope within this study. In the next section, two opposing views regarding the cognitive processes and mechanisms underlying CAF as well as the relationship between CAF elements and their manifestation in L2 output will be discussed.

4.3 Cognitive processes and CAF elements

Complexity, accuracy and fluency levels of L2 production have been widely examined in relation to attention and its allocation during L2 output production on task performance. Within this research area, two competing models, namely Skehan's (1998) *Trade-off Hypothesis* and Robinson's (2001) *Cognition Hypothesis,* have triggered the vast bulk of research within CAF in an attempt to gain deeper insights into the conditions or factors that affect task performance and the manifestations of CAF.

4.3.1 The Trade-off Hypothesis versus the Cognition Hypothesis

The *Trade-off Hypothesis* was first proposed by Skehan in 1998 and further developed in subsequent years resulting in the *Extended Trade-off Hypothesis* (Skehan and Foster, 2012). Skehan's (1998) and Skehan and Foster's (2012) proposal is based on the assumption that our attentional capacity is limited, consequently, the allocation of the learners' attentional resources during L2 performance on complexity, accuracy and/or fluency will depend on the characteristics of a task or communicative context. See Figure 4.2 for a summary of task features and their effects on L2 performance.

Figure 4.2
Task characteristics and their effects on CAF (Skehan, 1998).

Accuracy effects	Complexity effects	Fluency effects
Structured tasks Familiar information	Tasks requiring transformations, interpretations or divergent views	Structured tasks Familiar information

According to Skehan (1998), complexity, accuracy and fluency are likely to be enhanced as a result of task characteristics. Therefore, depending on the area which wants to be analysed or assessed, different types of tasks are advisable. For example, if accuracy is seen as the goal of assessment, structured tasks which have a clear time line

or macro-structure and deal with familiar information are likely to be preferred. Similarly, if complex language is expected, tasks in which more elaborate language is required seem to be more adequate. Likewise, interactive tasks tend to lead to greater accuracy and complexity, while monologic tasks may produce more fluency (Skehan, 2003). Additionally, Skehan (1998) also suggests that giving time to learners to plan the task will also have a positive effect on complexity and fluency, mainly. Yet, under certain conditions, accuracy also seems to improve as a result of planning.

Skehan (1998), then, defends the position that performance in the three areas (i.e. complexity, accuracy and fluency) "entails competition for attentional resources" (p. 168), which means that attention to one area may be raised at the expense of the others. According to Skehan (2009) and Skehan and Foster (2012), the natural tension is placed between complexity and accuracy. That is to say, if complexity and accuracy are competing for attentional resources, it is *likely* that only one will be raised in performance. Despite that, complexity and accuracy may be simultaneously increased *under certain conditions* (i.e. as a result of a combination of task characteristics and effective use of planning time), which Skehan and Foster relabelled as the *Extended Trade-off Hypothesis* (2012).

Robinson's (2001) Cognition Hypothesis, in contrast, defends the position that there are no capacity constraints on attention, which implies that attentional resources do not compete for being allocated on a specific dimension of L2 output production. Furthermore, the author suggests that raising task complexity will result in higher levels of linguistic complexity and accuracy (see Figure 4.3).

Figure 4.3
Proposed effects of task complexity on accuracy, fluency and complexity (Robinson, 2001).

Monologic tasks	
Simple	**Complex**
+fluency -complexity -accuracy	-fluency +complexity +accuracy
Interactive tasks	
Simple	**Complex**
+ fluency -accuracy +comprehension checks/ clarification requests	-fluency +accuracy +comprehension checks/ clarification requests

Robinson's (2001) predictions on the effects of task complexity on monologic tasks are based on the assumption that complexity of oral output will take place as a result of increasing complex functional demands of the communicative situation, in this case, tasks. In addition, greater attention to speech, derived from the high degree of complexity of the communicative task demands, usually leads to higher levels of accuracy. As for fluency, the author states that learners' fluency levels will decrease as task complexity increases. Regarding interactive tasks, Robinson claims that increasing task complexity leads to greater amounts of meaning negotiation in the form of comprehension checks and clarification requests, which in turn has a negative effect on the length and complexity of speech. In other words, interaction may prevent participants from producing complex output. In terms of fluency and accuracy, the assumptions on which Robinson based his predictions also hold for interactive tasks.

According to Robinson (2001), tasks should be sequenced according to their complexity in order to enhance language learning. To do so, he proposes a set of task features (resource-directing or resource depleting dimensions) which can be manipulated in order to increase or decrease task complexity according to the target level of instruction (see Figure 4.4).

Figure 4.4
Cognitive dimensions and their effects on output (Robinson, 2001).

Resource-directing dimensions	Resource-depleting dimensions
+/-few elements +/- here and now +/- reasoning demands	+/- planning +/-single task +/-prior knowledge

Robinson's proposal suggests that the presence (+) or absence (-) of certain task dimensions affects task complexity and consequently linguistic complexity. According to the author, resource-directing dimensions can increase or decrease task complexity by directing the participants' resources to aspects of language. For instance, tasks where few elements are used (+ few elements) will likely lead the learner to use singular forms rather than plural forms. Similarly, tasks which refer to the past (- here and now) will trigger the use of past tenses. As for tasks in which learners have to justify their statements (+ reasoning), more instances of subordination to give reasons are expected, for example. Resource-depleting dimensions, on the other hand, cannot be manifested through any type of linguistic feature but affect task complexity and in turn linguistic complexity. For example, tasks in which learners are not given any planning time (- planning), tasks in which the learner cannot use his prior knowledge (- prior knowledge) and tasks which have a dual objective (-single task) are claimed to increase task complexity.

To sum up, it is worth highlighting that even though Skehan and Robinson's respective hypotheses and models have opposing departure points, both models acknowledge the effects of task types and conditions on the manifestations of CAF elements and on their interrelationship. In the next section, a selection of relevant CAF research will be provided with the aim of shedding some light on the factors that

influence the manifestation of CAF and their development within Skehan or Robisnson's framework.

4.3.2 CAF research

Robinson (1995), for instance, explored the effects of task complexity on the performance on oral narrative discourse of 12 adult intermediate learners of English by manipulating the +/- here and now dimension of the task. Two procedures were used to analyse the effects of task complexity. First, the participants were asked to tell a story while looking at the pictures of a narrative task using the present tense (+ here and now). After that, the learners were given a different story, told to examine the pictures carefully and asked to tell the story in the past tense (- here and now) without looking at the pictures. The measures employed to analyse task performance were categorised into three groups: grammatically defined measures, intonationally and pausally defined measures and lexically and pausally defined measures. First, grammatically defined measures consisted of target-like use of articles and S-nodes per T-unit. Amount of silence (over a 2-second pause) and propositional content per unit were used as intonationally and pausally defined units. Finally, number of lexical words per unit and lexical words used in the task were the lexically and pausally defined units. The results reported in this study did not confirm any of the hypotheses posited by Robinson within the framework of the Cognition Hypothesis, according to which, performances under the - here and now dimension would lead to greater complexity, accuracy and dysfluency, since no significant differences were found between task complexity and any of the measures presented above, except one, number of lexical words per unit. Despite the lack of significant results in accuracy and syntactic complexity, Robinson concludes that even though no significant differences were found, mean scores and

nearly significant results do show a trend towards higher accuracy and complexity scores when task complexity increased. In addition, the author also highlights that sample size, task procedures, narrative choices and the measures used might have affected the results obtained.

The data gathered and analysed in Robinson (1995) was revisited in Robinson, Cadierno and Shirai (2009). In this second analysis, the authors used specific measures of tense-aspect marking, as they suggested that specific measures were more adequate to capture the effects of complex tasks on task performance. The hypotheses posited by the authors were that complex – *here and now* tasks would lead to a greater use of non-prototypical past tenses and progressive forms than + *here and now* tasks. The results reported showed that the mean scores of non-prototypical use of past tenses and progressive forms were higher or slightly higher on complex tasks than on simple tasks, which in some cases reached a marginally significant difference, which seemed to confirm the assumptions of the Cognition Hypothesis.

As for findings in line with the Trade-off Hypothesis, Foster and Skehan (1996) examined the effects of task choice and planning conditions (i.e. implementation conditions) on the L2 English oral performance of 32 pre-intermediate level adult learners. The students were divided into a control group and two experimental groups and were asked to carry out three tasks: a personal information exchange task, a narrative task and a decision-making task. According to the authors, the personal task was the easiest one as it required the least cognitive effort. The narrative and the decision-making task were seen as the most demanding ones, as the cognitive load required to perform the task was higher. As a consequence of the cognitive effort needed by the learners to do the tasks, the authors hypothesised that less attention would be given to linguistic form in these two tasks. In addition to task choice, the authors also

attempted to study the effects of planning conditions on task performance. Thus, the control group had no planning time before the task. Instead, they were just given a brief introduction to the objective of the task. The experimental groups, on the other hand, were given the same brief introduction as the control group followed by 10 minutes of individual planning. One of the experimental groups received instructions for detailed planning, which means that the participants were told to consider syntax, lexis, content and organisation. The other experimental group received no guidance on to what to consider for carrying out the task (i.e. undetailed planning).

The results reported did not show a significant effect of task familiarity on complexity measures probably due to the small sample used (ibid: 1996: 311). However, the analyses do indicate strong effects of planning conditions on fluency measures in the three tasks. In particular, the study reports that planning conditions have a positive impact on the total number of pauses and the amount of total silence. Similarly, significant results were found between planning conditions and the learners' performance in the complexity measure used (clauses per c-unit) in the three tasks. The effect of planning condition type provided inconclusive results across tasks and areas of performance. On the one hand, complexity results increased as the planning condition became more detailed, on the other, less detailed planning led to more accurate performances (fewer error-free clauses). The results on the effects of task familiarity on task performance were not conclusive either. Additionally, the results obtained across the different tasks seemed to report tradeoff effects between complexity and accuracy as a result of task type and planning conditions. In the personal task, for instance, the results for complexity were higher than for accuracy. The narrative task produced high scores in complexity while accuracy results decreased. Finally, the decision-making task produced high levels of accuracy and complexity, this last dimension being positively

affected by planning conditions. All in all, the authors conclude that "the combination of task type and task condition provides scope for complexity to operate as a viable goal without compromising accuracy" (ibid: 318).

Larsen-Freeman's (2006) time-series design offered a more qualitative perspective to the study of CAF. In this study, the author examined the written and oral productions of five adult Chinese learners of English who received EFL instruction in the verb tense-aspect system and article usage. The participants were asked to write a story of a past event and tell it orally. Both tasks were repeated (the same story each time) every six weeks over a period of six months. The measures used to analyse the written data were average number of words per t-unit for fluency, average number of clauses per t-unit for grammatical complexity, error-free T-units for accuracy and a type-token ratio for vocabulary complexity. In addition, both the written and oral data were analysed into idea units. The results of this investigation reported no significant differences from one time of data collection to the other, however, the analyses did reveal that the five participants made progress. A deeper analysis of the data, however, indicated that each of the five participants was following a different route to development in complexity, accuracy and fluency, probably due to "the way individuals have chosen to allocate their limited resources" (p. 601).

Following Robinson's hypothesis, Levkina and Gilabert (2012) investigated the effects of task complexity on the oral output of forty-two intermediate learners of English by manipulating pre-task planning time and +/- elements task characteristics. The measures selected to analyse the participants' oral output were speech rate in syllables for fluency, the Guiraud Index of lexical richness for lexical complexity, mean number of clauses per AS-unit for syntactic complexity and mean number of error-free clauses per AS-unit for accuracy. The results showed that fluency decreased without

planning time, even though the impact did not reach significance. In addition, lexical complexity was also negatively affected by the lack of planning time while syntactic complexity seemed to be unaffected by the presence of planning time. Finally, accuracy was not significantly affected when learners were provided with planning time. Regarding the effects of +/- elements, the results suggest that when the task was complex (i.e. more elements in a task) there was a positive significant impact on lexical complexity, which seems to support Robinson's (2001) Cognition Hypothesis. As for syntactic complexity and accuracy, no differences were found between complex and simple tasks (+/- elements), which does not seem to back up the Cognition Hypothesis, which claims that accuracy and complexity may increase as a result of task complexity. The authors justify the lack of significant results in accuracy and syntactic complexity by referring to the measures selected.

In addition, Ferrari's (2012) longitudinal study relates to some of the hypotheses posited in this research project. In particular, Ferrari's findings suggest that the study of L2 development and performance should be examined in relation to task type, as different tasks lead to different types of output in terms of CAF measures. In this three-year longitudinal study, the author analysed the oral production of four L2 learners of Italian and compared their output to that of two Italian native speakers in four tasks: a film retelling task, a narrative task, an informal interview and a telephone call. The measures employed were average number of subordinate clauses per AS-unit and average number of words per clause for syntactic complexity. The percentage of error-free AS-units was used to analyse accuracy whereas the average number of silent pauses as well as the average number of hesitation phenomena were employed as fluency measures. The results obtained from this longitudinal study indicate that accuracy improves in the long run, although there were periods in which a trade-off effect was

found between an increase in complexity and a decrease in accuracy. In addition, fluency also showed improvement over the three years of the study. Another of the findings from Ferrari's (2012) piece of research related to the effects of task type on CAF. According to the author, CAF results were sensitive to task characteristics. For instance, interactive tasks seemed to produce less complex language (which resembled native speakers' performance on the same task) and more accurate language. Monologic tasks, on the other hand, contained more complex language and led to lower scores in fluency. Even though the size of the sample was quite limited, a number of implications for the study of L2 development are addressed by the author. Firstly, the findings from this study suggest that L2 use is sensitive to task, consequently, L2 development should be studied using data from a variety of tasks. Additionally, the data examined also seems to indicate that interpretations of development in CAF should be dependent on task type, as a decrease in one of these dimensions (i.e. syntactic complexity) might be an indicator of development rather than incompetence (p. 292).

Finally, Kuiken and Vedder (2012) examined the oral and written productions of L2 Italian and L2 French learners using complexity, accuracy and fluency measures with regard to the learners' proficiency level and task complexity. To do so, three studies were conducted. This section, however, will only refer to the study which analysed L2 oral performance. In this study the authors analysed the oral performance of 44 Dutch students of Italian in a monologic task consisting in leaving a phone message on an answering machine in which the dimension + /- elements was manipulated. The measures used were total number of errors per AS-units, number of 1st, 2nd or 3rd degree errors per AS-unit for accuracy, number of dependent clauses per clause for complexity and the total number of word types per square root of two times the total number of word tokens for lexical variation. The results indicated a significant

influence of proficiency level on accuracy. As for syntactic complexity, no differences were found between high and low proficient learners. Regarding the effect of task complexity, the analyses showed a significant effect of task complexity on accuracy (i.e. decrease of errors) and on syntactic complexity inasmuch as fewer dependent clauses were found in the more complex version of the task, which, according to the authors, contradicts the Cognition Hypothesis. The authors conclude by saying that even though certain task characteristics may lead to the production of particular language features, "there is no general effect of task complexity on syntactic complexity" (p.165).

The CAF research presented in this section has attempted to gain understanding of the CAF construct in terms of the factors that trigger its manifestation, the relationship between the three dimensions and their development. As seen from the literature review provided, no consensus seems to have been reached on the reasons that lead to the production of accurate and complex speech. However, agreements are encountered when claiming that task types and conditions affect output production. In Chapters 6 and 7 the results of the analyses conducted to investigate the effects of CLIL+EFL instruction and EFL instruction on CAF measures in two tasks will be presented and discussed. Additionally, this study will also focus on the development of CAF measures in the two instructional settings in an attempt to examine the relationship between CAF elements and the processes of L2 language development in instructed SLA.

CHAPTER 5 Method

This chapter presents the procedures, methodology and measures used to analyse the oral data collected in this dissertation. As previously mentioned, this is a two-year longitudinal study which aims at exploring the effects of CLIL on L2 English young learners' oral production skills. To do so, an investigation based on a time-series design was conducted in which the oral output of two groups of learners from two Catalan schools was studied and compared. One of these groups, the non-CLIL group (also referred to as the control group), was formed by learners who had been exposed to EFL instruction, whereas the other group, the CLIL group, was composed by learners who had received CLIL and EFL instruction

5.1 Design of the study

As most research in the field of Applied Linguistics, this study has a quasi-experimental design. The independent variable was type of instruction received by the learners, namely CLIL plus EFL instruction or EFL instruction alone, whereas the dependent variable was the learners' oral performance analysed by means of complexity, accuracy and fluency measures. As for the moderator variables controlled in the study, there was school and learners' proficiency (i.e. high and low achievers). Finally, amount of English instruction was used as a control variable, since the number of hours learners had of English instruction (EFL alone or EFL plus CLIL) was held constant between the two groups of learners throughout the two years of the study.

As seen in Table 5.1, the oral data were collected at four different times (T0, T1, T2 and T3) in order to explore the processes of language learning and study the progress learners underwent in CLIL and non-CLIL settings after 122.5, 164.5 and 244 hours of

English instruction. As the table shows, data collection sessions from the non-CLIL group started in the school year 2010-2011, whereas the data from the CLIL group were gathered in the subsequent academic year 2011-2012. For the purposes of the study, the data collection sessions could not be simultaneously conducted in the two groups (see section 5.4.1 for further details on data collection procedures).

Table 5.1
Data collection times and amount and type of English instruction received by the groups according to data collection times.

	Groups	**Data collection date**	**Hours of English instruction**	**Type and amount of English instruction**
Time 0	Non-CLIL	September 2010	-	-
	CLIL	September 2011	-	-
Time 1	Non-CLIL	June 2011	122.5 hours	3.50 hrs/week EFL instruction
	CLIL	April 2012	122.5 hours	1 hr/week CLIL instruction (27 hours in total)
				3.50 hrs/week EFL instruction (95.5 hours in total)
Time 2	Non-CLIL	December 2011	164.5 hours	3.50 hrs/week EFL instruction
	CLIL	September 2012	164.5 hours	1 hr/week CLIL instruction (37 hours in total)
				3.50 hrs/week EFL instruction (127.5 hours in total)
Time 3	Non-CLIL	June 2012	244 hours	3.50 hrs/week EFL instruction
	CLIL	March 2013	244 hours	1 hr/week CLIL instruction (54 hours in total)
				3.50 hrs/week EFL instruction (191 hours in total)

5.2 Schools' context

This study was carried out in two semi-private Catalan schools which started the implementation of CLIL in grade 5 of primary education in September 2011. This research project started in September 2010, though, in order to be able to gather data

from the non-CLIL group in grade 5, from the same schools. One of the schools is located in a city near the metropolitan area of Barcelona, from now on referred to as School I, whereas the other one is situated in a town in central Catalonia, referred to as School II in this study. Despite being situated in two very different geographical areas, both schools share similar socio-economic data as both schools are composed by families with a mid-high socioeconomic status. Both schools offer schooling for infant school, primary education and compulsory secondary education. School I also offers non-compulsory secondary education, which prepares students for Spanish university entrance exams. In addition, both School I and II have more than one class for each of the grades from infant school to secondary school.

Neither of the schools had had previous experience in implementing CLIL courses before the start of this study in September 2010. However, various initiatives had been adopted in the two schools to introduce English as a foreign language in kindergarten in the form of short sessions (from 30' to 60' a week) in which young learners were exposed to the foreign language through story-telling and game sessions.

In September 2011 both schools started to implement CLIL in grade 5. The content subject selected to teach CLIL was *Science*. Thus, since September 2011 up to June 2013, learners had 3.50 hours of EFL instruction plus 1 hour of *Science* a week in which learners studied content from the science curriculum in primary school that had not been introduced or dealt with in the curricular subject of Science in Catalan, the school's vehicular language.

The teachers in charge of delivering the CLIL subject, *Science*, were primary school EFL teachers whose level of English was higher than B2, according to the Common European Framework of Reference for Languages. The teachers had not taught CLIL before but had attended a training course in CLIL before the

implementation of CLIL in their respective schools. See Table 5.2 for a summary of the most important characteristics of the schools selected for this study.

Table 5.2
Schools' main features.

	SCHOOL I	SCHOOL II
Location	Near Barcelona metropolitan area	Central Catalonia
EFL Experiences	30 ' sessions in kindergarten	30'-60' sessions in kindergarten
Previous CLIL experience	No	No
CLIL Subject	Science	Science
CLIL training before CLIL implementation	Yes	Yes
CLIL Teacher's profile and level of English	School English teacher - higher than B2	School English teacher – Higher than B2

5.2.1 EFL curriculum in primary schools in Catalonia

As Nikolov (2009: 23) points out, the construct used to measure the linguistic outcomes of EFL instruction among young language learners "has to be in line with the curriculum" in order to assure that neither linguistic nor cognitive task demands affect the results obtained. Because of this, it is crucial to know what learners at early stages of cognitive and linguistic development can and cannot do.

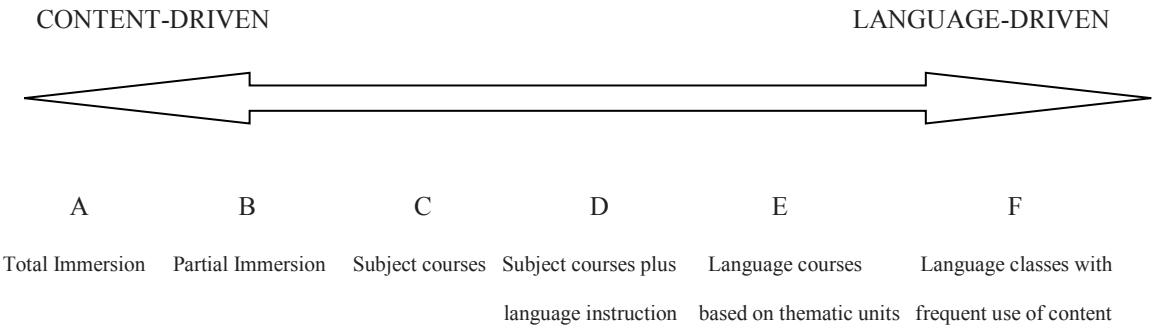
According to the Catalan EFL curriculum[4], learners at the end of their 4th year of primary education, around 9-10 years old, should be able to participate in short oral interactions which are familiar to them such as dialogues about their daily lives and routines. In addition, they should also be able to understand and produce everyday formulaic language expressions for greetings, expressing emotions, describing the weather and providing personal information.

As presented in the EFL curriculum for primary school, EFL lessons should focus primarily on communicative competences (comprehension and production) and

[4] http://www.xtec.cat/web/curriculum/primaria/curriculum

meaning instead of linguistic form by means of thematic units in which topics like food, clothes, schools and homes serve as a means to foster English language learning. According to the various models for the integration of language and content teaching presented by Met (1998) (see Figure 5.1), the types of EFL instruction implemented in primary schools in Catalonia would belong to groups E and F and would be categorised as language-driven programmes.

Figure 5.1
Language-content continuum (Met, 1998).



The foreign language teaching programme that the two schools selected for this study started to implement in September 2011 (CLIL+EFL instruction), on the other hand, would be further away from language-driven programmes than EFL instruction on its own and would belong to group D 'Subject courses plus language instruction'. The learners in this study that received this type of instruction, the CLIL group, were exposed to language lessons plus subject science lessons in English (i.e. CLIL) that emphasised both content and language learning. The topics dealt with in the CLIL lessons in the schools selected in this study were Science topics such as the cell, plants and vertebrate and invertebrate animals among others. As for the curriculum in CLIL programmes, it needs to be said that the Catalan Department of Education has not explicitly stated the competences, methodologies and assessment criteria schools need

to implement when teaching CLIL. However, Catalan stakeholders do suggest Coyle's (2006) 4Cs framework for CLIL implementation in Catalan schools:

> The 4Cs Framework suggests that it is through progression in knowledge, skills and understanding of the content, engagement in associated cognitive processes, interaction in communicative context, developing appropriate language knowledge and skills as well as acquiring deepening intercultural awareness through the positioning of self and otherness, that effective CLIL takes place (p.15).

The 4Cs Framework identifies the 4 main areas which need to be addressed in CLIL lessons in order for CLIL to be effective and successful, which are Content, Cognition, Communication and Culture. As Coyle states (ibid), CLIL instruction focuses on content, cognitive and linguistic development and moves away from purely language-driven programmes in which language form is the driving force.

5.3 Participants

The sample of participants used in this study belongs to the *CLIL-SLA Project,* which studies the effects of CLIL on young learners' linguistic competence in English and motivation. The subjects were 52 Catalan/Spanish bilingual primary students learning English as a foreign language from two of the five Catalan schools used in the CLIL-SLA Project.

The subjects that participated in this study were randomly selected from the two schools in a way that heterogeneity within each of the two groups was assured. To do so, the selection of young learners was carried out taking into account their grades in the three language subjects (Catalan, Spanish and English) in an attempt to have a mixture of talented and untalented language learners in each group. As Table 5.3 shows, the non-CLIL group was made up of 20 learners (10 from each school) who received regular EFL instruction (3.50 hours a week) during the study, whereas the CLIL group was composed by 32 learners (18 from School I and 14 from School II) who were

exposed to EFL and CLIL instruction (3.50 hours a week of EFL plus 1 hour a week of CLIL). Unlike other schools implementing CLIL, the two schools used in this study offered CLIL to all students, which means that learners in the CLIL group were not necessarily more motivated or 'talented' in the learning of foreign languages than their peers in the non-CLIL group.

Table 5.3
Description of participants.

Groups	Participants	Age at study onset	English exposure during the study	Previous English exposure at school	% learners attending extracurricular English lessons
Non-CLIL	20	9-10 years old	3.50 hrs/week EFL instruction	420-437 hrs	26,90%
CLIL	32	9-10 years old	1 hr/week CLIL instruction 3.50 hrs/week EFL instruction	420-437 hrs	25,00%

As can be seen from Table 5.3, at the onset of data collection sessions in both groups, the participants were 9-10 years old and were starting their 5th year of primary education. Regarding the amount of English exposure received in school up to the beginning of the study in September 2010, all the participants had received EFL instruction and the number of hours ranged from 420 to 437 in both schools, which made the schools similar and hence comparable. Likewise, the percentages of learners from the CLIL and non-CLIL groups that attended extracurricular English lessons outside school when this research project started in September 2010 were similar, which diminishes the effect of extracurricular English exposure on the learners' performance in the two oral tasks selected for the study.

5.4 Data collection

5.4.1 Procedures

Data from the two groups were collected and recorded at four different times (Time 0, Time 1, Time 2 and Time 3). Data collection sessions in the non-CLIL group took place from September 2010 up to June 2012, while data from the CLIL group was gathered from September 2011 to March 2013 (see Table 5.4).

Table 5.4
Data collection times for CLIL and non-CLIL groups.

	Time 0	**Time 1**	**Time 2**	**Time 3**
Non-CLIL	September 2010	June 2011	December 2011	June 2012
CLIL	September 2011	April 2012	September 2012	March 2013

As previously mentioned, the onset of the data collection sessions could not take place within the same school year in both groups owing to the fact that CLIL started to be implemented in September 2011. Thus, the data collection sessions from the non-CLIL group started in September 2010 (T0), while in the CLIL group the data started to be collected in September 2011 (T0). As for the rest of the data collection times, calculations of the total amount of English exposure learners received (3.50 hours a week of EFL instruction in the non-CLIL group and 3.50 hours a week of EFL instruction + 1 hour a week of CLIL instruction in the CLIL group) were carried out in order to keep amount of English instruction constant between the two groups. Hence, the data collection sessions in the CLIL group were carried out earlier in the year than in the non-CLIL group due to the fact that CLIL learners received more hours of English instruction a week than non-CLIL learners.

In the first data collection session, Time 0, a bio-data questionnaire, an English proficiency test and two oral tasks in English, an interview and a picture-elicited

narrative, were administered to CLIL and non-CLIL learners. The bio-data questionnaire included questions on the families' socio-economic background, the languages spoken at home, previous English exposure and the contexts in which they used English. The English proficiency test administered to the participants was based on a combination of tasks taken from two of the tests which compose the widely known Cambridge Young Learners English Tests: Movers and Flyers. The tests were designed, piloted and validated prior to their administration to the participants in this study. The scores obtained from the proficiency test were used to establish the learners' initial proficiency level at T0 and classify learners into low and high achievers in an attempt to examine the effects of CLIL on learners with different proficiency levels (see Table 5.5 for the classification of low and high achievers).

Table 5.5
Classification of CLIL and non-CLIL learners according to proficiency level.

	Mark in the proficiency test at Time 0	Number of Non-CLIL learners	Number of CLIL learners
Low achievers	<3.9	10	16
High achievers	≥3.9	10	16

The oral tasks were administered individually to all the participants by a researcher from the CLIL-SLA Project. The first task was the interview in which learners had to answer questions about themselves. After the interview, the participants were shown the pictures of the narrative task. Learners were given a couple of minutes to familiarise with the story and understand the plot. After that, learners were asked to tell the story (see section 5.4.2 for further details on the tasks). The same procedure for the administration of the oral tasks was followed in the three subsequent data collection sessions for Time 1, Time 2 and Time 3.

5.4.2 Data collection instruments

In order to analyse the learners' oral production skills in English two speaking tasks were used: an oral interview (an interactive task) and a picture-elicited narrative (a monologic task). Both tasks had been previously used by the *BAF Project* (Muñoz, 2006) with learners of similar ages from schools in Barcelona and met the learners' linguistic and cognitive developmental stage according to the EFL curriculum presented by the Catalan Department of Education.

5.4.2.1 Oral interview

The interview consisted of seven questions related to the learners' families, personal lives and routines (see Appendix A for the interview). Many of the questions asked in the interview dealt with topics that learners at that age (9-12 years old) have studied and used in their EFL lessons. However, even though learners were familiar with the vocabulary used in the questions, they did not master some of the grammatical structures needed to answer the questions accurately (i.e. the use of future or past tenses) due to their level of proficiency. Despite this, learners did understand the meaning of the questions asked in most cases, as structures such as the ones mentioned above are implicitly dealt with in regular EFL instruction in primary school. If learners did not understand a question, the researcher who was administering the tests reformulated it several times. Rarely did the researcher have to use the learners' L1. Regarding total task time (TTT), Table 5.6 shows the average amount of time in minutes learners needed to fulfil the task at the four data collection times.

Table 5.6
Total task time means in minutes obtained in the interview for Times 0, 1, 2 and 3.

	Time 0	**Time 1**	**Time 2**	**Time 3**
Non-CLIL	3.07	5.22	3.90	3.76
CLIL	4.77	4.80	4.11	3.57

The use of oral interviews to assess and study oral production skills has been strongly criticised on the basis of the unequal power relationships which are established between the interviewer and the interviewee (Luoma, 2004) and the "limited repertoire of functions" young learners show in this type of tests (McKay, 2006: 84). However, the decision to include the interview in this study as a tool of data collection was taken inasmuch as the structure of the task as well as the content of the questions asked are very similar to the type of speaking tasks young language learners in this study performed in their regular EFL lessons. Thus, in an attempt to use tasks which are familiar to the learners, the interview was used as one of the oral tasks in this study.

5.4.2.2 Picture-elicited narrative

Right after the interview, learners were shown a story based on six pictures (see Appendix B for the story). Picture (1) showed a mother and her two children preparing some food to go on a picnic. Picture (2) depicted the mother showing the children on a map the place to go to have the picnic in the background and a dog trying to get into the picnic basket full of food in the foreground. Picture (3) showed the children walking on the street and waving goodbye to their mother. In Picture (4) the learners could see the two children from the story in the countryside. Picture (5) depicted the two children seeing the dog coming out of the basket. And finally, Picture (6) showed the children surprised at seeing the picnic basket almost empty.

The participants were given some minutes to look at the pictures and understand the plot of the story. Once they felt comfortable with the story, they were asked to narrate it. When learners found it hard to tell the story, the researcher asked them some questions to help them produce some language. The questions had been previously selected and were the same for all the learners. Tables 5.7 shows the average amount of time learners used to carry out the task at the four data collection times.

Table 5.7
Total task time means in minutes obtained in the picture-elicited narrative for Times 0, 1, 2 and 3.

	Time 0	**Time 1**	**Time 2**	**Time 3**
Non-CLIL	2.62	2.32	1.83	1.49
CLIL	2.28	1.84	2.15	1.81

As Luoma (2004) claimed, the use of narrative tasks to assess EFL learners' oral production skills is quite common. However, it is important to select good picture sequences which "generate enough talk and provide opportunities for the examinee to show what they know" (ibid: 144). According to the author, good picture-elicited narratives should enable learners "to show their control of basic narrative features such as setting the scene, identifying the characters and the main events and telling them in a coherent sequence" (p. 144). In order to study the appropriateness of the picture-elicited narrative chosen for this study, a pilot study was carried out (Bret, 2011) in which learners of the same ages were administered the picture-elicited narrative selected for this dissertation. The results of the study showed that the task was appropriate for the learners' as they felt comfortable with the story and were able to verbalise most of the 'narrative features' presented by Luoma (2004).

5.4.2.3 Qualitative instruments

Despite the clear focus on second language learning processes of this dissertation, this study will also attempt to address several pedagogical aspects derived from the implementation of CLIL in primary schools. To do so, brief references to the data obtained from several questionnaires and oral interviews administered at the beginning and at the end of the first and second year of CLIL implementation will be made. The questionnaires addressed questions related to the opinion learners, parents and teachers had on CLIL and its effects on English foreign language learning. In addition, teachers were also asked about the implementation process (i.e. difficulty in delivering CLIL lessons, degree of institutional support received, weaknesses and strengths of their CLIL lessons, etc.).

Besides the use of questionnaires to explore the participants' views on the implementation of CLIL, data was also collected from several classroom observation sessions carried out in the two schools within the CLIL classroom. Even though the results derived from the instruments and procedures stated above will be partly interpreted in the discussion section, a thorough interpretation of the the data obtained from these instruments is beyond the scope of this dissertation. See Pladevall-Ballester (2014) for a more detailed study of the qualitative data gathered in the CLIL-SLA Project.

5.5 Data analyses

The oral data collected for this study was analysed using quantitative and qualitative methods. This section provides a description of how the data was treated in the processes of transcription, codification and analysis along with a justification of the measures used.

5.5.1 Task transcription

The recordings of the tasks were transcribed and analysed using CLAN (Computerized Language Analysis) available at the CHILDES website.[5] Transcripts were coded for syntactic and lexical complexity, accuracy and fluency using CHAT transcription symbols. The total number of transcripts used in this research study was 416 (208 of the interview task and 208 of the narrative task).

5.5.2 Measures and unit of analysis

As mentioned earlier, this study used CAF measures to study L2 English young learners' oral production skills and their development. Thus, the selection of the measures was carefully made in order to capture progress among L2 English young learners' speech and gain insight into the development of each of the areas within CAF.

5.5.2.1 Complexity

As explained in Chapter 4, the study of L2 complexity in this research was based on two of the three main components of complexity presented by Bulté and Housen (2012): propositional and linguistic complexity. In order to investigate

[5] http://childes.psy.cmu.edu

propositional complexity, the *total number of units* (TNU) participants produced in each of the tasks at the four data collection times (T0, T1, T1 and T3) was computed.

The unit of analysis used in this study, from now on referred to as *unit*, was specifically designed for this study to analyse young language learners' oral output, a research area which has received little attention, at least from a quantitative point of view. The unit was defined as a context-dependent meaningful utterance, grammatical or ungrammatical, which conveys one piece of information or idea. As seen from the definition, the main criterion to identify and segment utterances into units was semantic. For example:

(1) INVESTIGATOR: How are you?
SUBJECT: I'm fine. [unit]

(2) INVESTIGATOR: How are you?
SUBJECT: My name is …. [NOT a unit]

(3) SUBJECT: The brother and sister preparing a picnic. [unit]

In examples (1) and (3), the subjects' answers were counted as units because they were meaningful in the context of the interaction or communicative situation. In other words, the utterances produced answered the question asked (example 1) or referred to the story being described (example 3). In example (2), on the other hand, the utterance produced by the participant was not computed as a unit since it did not fulfil the requirement of being context-dependent within the context of the interaction.

Besides being context-dependent utterances which conveyed one piece of information or idea, the implicit or explicit presence of a predicate (i.e. lexical verb) within the utterance was also employed as a criterion for the identification of *units*. See examples below:

(4) INVESTIGATOR: Have you got any brothers or sisters?
SUBJECT: It one sister. [unit]

(Intended form: I have one sister)

(5) INVESTIGATOR: What do you do at the weekends?
SUBJECT: Saturday football. [unit]

(Intended form: On Saturdays I play football)

(6) SUBJECT: The sandwich and the dog in the basket. [unit]

(Intended form: the sandwich and the dog are in the basket)

(7) SUBJECT: The boy and the girl on the mountain. [unit]

(Intended form: the boy and the girl are in the mountain)

All the examples provided above lack a predicate, probably due to the learners' low stage of language development. However, the answers given by the participants in examples (4) and (5) show that the learners were able to provide a meaningful answer. Likewise, in examples (6) and (7), the participants failed to produce a verb in the sentences. Despite that, the learners showed traces of SVO order in addition to providing a meaningful utterance. These examples indicate that the learners were trying to verbalise complex grammatical relationships between the phrases in their oral productions other than mere enumerations of elements, which were discarded as units. Hence, in order not to ignore that type of progress among the sample population of this study, the explicit use of a predicate was not necessary to identify units. Additionally, another of the reasons why the implicit or explicit use of a predicate was also employed as a criterion to identify units relates to the difficulty encountered in some cases to identify the extension of ideas on the basis of semantic criteria only (Foster, Tonkyn and Wigglesworth, 2000). This way, the implicit or explicit use of a verb would served to identify the end of the unit. See examples below:

(8) INVESTIGATOR: Have you got any brothers or sisters?
SUBJECT: I have one brother [unit] he name Marc [unit]
(Intended form: I have one brother. His name is Marc)

(9) SUBJECT: Mountain, the girl and the boy the dog. [NOT a unit]

(10) INVESTIGATOR: Tell me about your family.
SUBJECT: my brother, my mother, my father. [unit]

Example (8) illustrates the use of the implicit or explicit presence of a predicate to identify different units in the utterance provided by the learner. Examples (9) and (10), in contrast, display enumerations. In the case of the narrative task (example 9), enumerations were not counted as units, while in the interview task and as a response to certain questions, enumerations could be counted as units, as the answer in example (10).

Due to the nature of oral interactions and the use of the interview, utterances made of phrases or words (even one-word utterances) which contained elliptical material were quite common. Thus, utterances which contained elliptical material and fulfilled the requirements presented earlier were also computed as units. For example:

(11) INVESTIGATOR: Any brothers or sisters?
SUBJECT: Three. [one-word unit]

(12) INVESTIGATOR: What is your brother's name?
SUBJECT: Jordi. [one-word unit]

One-word units were not included within the measure used to analyse propositional complexity (TNU). However, when the TNU was used to calculate other measures like coordination, subordination or error-free units, one-word units were also included in the TNU.

The main reason that led to the design of a new unit for the analysis of the oral data in this study was the inexistence of adequate units to investigate the interlanguage produced by young language learners (see Table 5.8 for a selection of the most widely used units in Second Language Acquisition research).

Table 5.8
Definitions of the most widely used units in Second Language Acquisition research (taken from Foster et al., 2000).

Unit	Definition
T-Unit	'One main clause with all subordinate clauses attached to it' (Hunt, 1965:20). 'One main clause plus whatever subordinate clauses that happen to be attached to or embedded within it' (Hunt, 1966:735). 'A main clause plus all subordinate clauses and non-clausal structures attached to or embedded in it' (Hunt, 1970:4).
C-Unit	'Utterances, for example, words, phrases and sentences, grammatical or ungrammatical which provide referential or pragmatic meaning' (Pica et al., 1989:72).
AS-Unit	'A single speaker's utterance consisting of an independent clause or subclausal unit, together with any subordinate clause(s) associated with either' (Foster et al., 2000:365).
Clause	'Either a simple independent finite clause or a dependent finite or non-finite clause (Foster and Skehan, 1996:310).

The C-unit, for instance may seem the most adequate unit to analyse the type of speech produced by young language learners. However, as Foster et al. (2000) indicate, semantic units such as the C-unit, are difficult to identify since one never knows with certainty where one idea ends and a new one starts. Because of this, semantic units are very often complemented by syntactic units (Foster et al., 2000). Therefore, the main reason why semantic units which focused exclusively on meaning were discarded was because they did not capture progress and development in the learners' L2 morphosyntactic system. See examples below for a comparative analysis of the oral data collected using the C-unit and the *unit* designed in this study.

(13) SUBJECT: One boy, one girl, sandwich, dog.

(14) SUBJECT: One boy one girl look the map.
(Intended form: The boy and girl are looking at the map)

These two utterances, taken from the narrative task administered to the learners, exemplify why the C-unit was not used in this study. According to the definition of the C-unit, utterance (13) would be made up of either one C-unit (the whole utterance) or four C-units (one boy // one girl// sandwich // dog). Utterance (14) would consist of one C-unit (the whole utterance) or four C-units, one unit for each content word used (boy// girl// look// map). Provided that the C-unit was interpreted as an utterance, the examples presented above would be seen as two C-units, regardless of the morphosyntactic features in the learners' output. Likewise, if the criterion to identify C-units was the number of content words used (ideas), both examples would have the same number of C-units, independently of the fact that one of the utterances is an enumeration of elements and the other contains the SVO structure.

As for the use of intonational units such as the utterance, tone-unit and idea unit, it is worth pointing out, as Foster et al. (ibid) claim, that their use with second language learners seems to be inadequate as the pauses in the learners' speech may indicate lexical search rather than the end of message formulation. Because of this, it was decided not to contemplate intonational units as an option to be used with young language learners.

Syntactic units, on the other hand, seemed the most effective for the purposes of this study. However, syntactic units such as the T-unit, one of the most widely used units in Second Language Acquisition research, does not overcome the challenges presented in the analyses of oral data such as the treatment of repetitions, hesitations,

the use of elliptical material and other residual fragments found in speech. Additionally, it seemed inadequate for the analysis of young language learners' interlanguage which is characterised by the use of fragmented utterances and lack of predicates (at least at low stages of language development). The utterances below show why the use of the T-unit was discarded in this study:

(15) INVESTIGATOR: What did you do last weekend?
SUBJECT: My friends dinner in my house.
(Intended form: My friends had dinner in my house)

(16) INVESTIGATOR: What do you usually do at the weekend?
SUBJECT: Bike.
(Intended form: I rode my bike).

According to the definitions of the T-units, neither of the two examples provided above would be seen as T-units due to the lack of a predicate. Even though there is a big difference between the two examples in terms of morphosyntactic development (attempt to verbalise the canonical SVO word order and the use of a prepositional phrase in utterance 15 as opposed to a one-word utterance in example 16), neither of the utterances would be considered a T-unit. However, in this study, utterance (15) would be seen as a *unit* as the explicit presence of the predicate is not compulsory. As for example (16), on the other hand, it was not counted as a unit due to the lack of morphosyntactic information and failure to reconstruct the utterance by means of elliptical material.

In addition to the T-unit, Foster et al. (ibid) propose the use of the AS-unit, which was exclusively created for the analysis of oral data, unlike the T-unit. A priori, the AS-unit seemed a reliable measure to analyse the oral data gathered in this study.

However, the presence of a predicate to identify AS-units was one of the requirements presented by the authors.

Altogether, the units available from previous research did not seem to be designed in accordance with young language learners' cognitive and linguistic characteristics. Because of this, the AS-unit as well as other units such as the clause, which also requires the use of finite or non-finite versb, were discarded in this study in favour of a *unit* which did not need the explicit presence of a verb.

Regarding linguistic complexity, this study focused on the analysis of syntactic and lexical complexity. Syntactic complexity was examined by means of two measures, namely the percentage of coordinate units (% CU) and the percentage of subordinate units (% SU). The coordination measure was obtained by dividing the total number of coordinate units (TNCU) by the total number of units (TNU) including one-word units, in the case of the interview task. Utterances in which learners did not link the clauses with a coordinating conjunction were not counted as coordinate clauses, but as two different units. See examples below for a sample of coordinate units produced by the learners and its codification:

(17) SUBJECT: The mum of the boys looking the map [unit] [coord. unit] and the dog looking the food. [coord. unit] [unit]

(Intended form: The mum of the children is looking at the map and the dog is looking at the food)

(18) SUBJECT: Brother and sister going to the park [coord. unit] [unit] and dog eat sandwich. [coord. unit] [unit]

(Intended form: The brother and sister are going to the park and the dog is eating the sandwich)

(19) INVESTIGATOR: What do you usually do at the weekends?
SUBJECT: I go to with my grandparents [coord. unit] [unit] and I do my homework. [coord. unit] [unit]

(Intended form: I go to my grandparents' house and I do my homework)

(20) INVESTIGATOR: Tell me about your family.
SUBJECT: My mother is working in the school [coord. unit] [unit] and my father works in Barcelona. [coord. unit] [unit]

(Intended form: My mother works in a school and my father works in Barcelona)

All the examples provided above are made up of units which contain predicates. However, as previously stated, context-dependent utterances which did not contain the explicit presence of a verb were also counted as units. Taking this into account utterances which did not include the presence of a verb but were linked by means of a coordinating conjunction were counted as coordinate units. See examples below:

(21) INVESTIGATOR: What are your parents' jobs?
SUBJECT: My mother secretary [coord. unit] [unit] and my father cooking. [coord. unit] [unit]

(Intended form: My mother is a secretary and my father is a cook).

(22) SUBJECT: The boys it is sad [coord. unit] [unit] and the basket no food [coord. unit] [unit].

(Intended form: The children are sad and there is no food in the basket).

Additionally, units which contained predicates but the verb complement was in the L1 were also counted as coordinate units provided that the coordinating conjunction was used. See examples below:

(23) SUBJECT: The bo ys l ook i n t he *cistell* (basket) [coord. unit] [unit] and they aren't the sandwich. [coord. unit] [unit]

(Intended form: The boys look at the basket and there isn't the sandwich).

Regarding the analysis of subordination, it was decided to use the percentage of subordinate units (% SU) pr oduced by the learners, which was calculated by dividing the total num ber of s ubordinate uni ts (TNSU) by t he t otal num ber of units (TNU), including one-word units. See examples below for instances of subordination produced by the learners in this study:

(24) INVESTIGATOR: What's your father's job?
SUBJECT: My father goes in a office [to do maps of house].

(Intended form: M y f ather g oes t o a n of fice t o do hous e plans)

(25) INVESTIGATOR: What do you usually do at the weekends?
SUBJECT: I go to classes [for drive a bike].

(Intended form: I g o to l essons t o learn how t o dr ive a bike)

(26) INVESTIGATOR: What do you usually do at the weekends?
SUBJECT: I go [to play tennis].

(27) INVESTIGATOR: What do you usually do at the weekends?
SUBJECT: In Sunday I think [that we go to Barcelona].

(Intended form: O n S unday I t hink t hat w e a re going t o Barcelona)

(28) SUBJECT: She looks [what the dog eat all the picnic].

(Intended form: S he s ees t hat t he do g ha s eaten a ll t he picnic)

(29) SUBJECT: They look [that in the nest there are a dog]

(Intended form: They see that the dog is in the basket)

Even though some of these utterances are not accurate in the use of subordinate conjunctions, attempts to produce subordination as the ones just presented were considered as subordinate units with the aim of exploring the emergence of subordination among young language learners.

At this point, the definition of *unit* provided earlier needs to be complemented. So far, units have been mainly defined as meaningful utterances that can provide referential meaning on their own. Such a definition seems to work well when measuring simple sentences or sentences with coordination, as coordinate units have meaning on their own in the same way as simple sentences do, however, that does not seem to be the case with subordinate units, since the meaning of one unit is embedded into the other. In spite of that, in order not to penalise learners by diminishing the number of units in the cases in which subordination was used, it was decided to count as units both subordinate and main units within a sentence. Hence, each of the examples provided above would consist of two units one of them being a subordinate unit.

As for lexical complexity, this investigation focused on the degree of lexical diversity unfolded by the learners in their oral productions. To do so, the learners' use of lexical categories was studied by means of three ratios presented as percentages: noun, lexical verb and adjective ratios (% NR, % VR and % ADJR respectively). The calculation of these ratios consisted in dividing the total number of nouns, lexical verbs or adjectives (TNN, TNV and TNADJ respectively) by the total number of words in English (TNWE) learners produced.

5.5.2.2 Grammatical accuracy

The approach adopted in this study to investigate the degree of grammatical accuracy L2 English young language learners unfolded in their oral productions was the one suggested by Iwashita et al. (2008) in which both global accuracy and accuracy of specific grammatical features are considered. The measure used to analyse global accuracy was percentage of error-free units (% EFU), whereas percentage of correct lexical verb forms (% CV) was chosen to study the degree of accuracy in the use of verbs.

Error-free units were considered as such when they did not contain any type of morphological, syntactic or lexical error. Phonological errors, on the other hand, were overlooked as the study of pronunciation was beyond the scope of this dissertation. See examples below for typical EFU and non-EFU produced by the learners in this study:

(30) INVESTIGATOR: What is your name?
SUBJECT: My name is Jordi. [EFU]

(31) INVESTIGATOR: How old are you?
SUBJECT: Eleven. [EFU]

(32) INVESTIGATOR: What do you usually do at the weekends?
SUBJECT: I go to play football. [EFU]

(33) INVESTIGATOR: What are you going to do next weekend?
SUBJECT: Play the football [NON-EFU]

(34) INVESTIGATOR: Ask me a few questions.
SUBJECT: I am sports? [NON-EFU]

(Intended form: Do you like/play any sport?)

(35) SUBJECT: The girl and boy go to the park. [EFU]

(36) SUBJECT: [...] because the dog likes eating. [EFU]

(37) SUBJECT: The boy and girl walk in the forest. [EFU]

(38) SUBJECT: The two brothers comes mountain. [NON-EFU]

(39) SUBJECT: The boy and girl say bye bye in her mum. [NON-EFU]

Regarding the other accuracy measure used in this study, % CV, it is worth pointing out that not only lexical but also morphological and syntactic aspects were considered to categorise verbs as 'correct verb forms'. In other words, morphosyntactically correct yet lexically incorrect verbs were not counted as correct verb forms. The percentage of CV was obtained by dividing the total number of correct verb forms (TNCV) by the total number of verbs (TNV). See examples below for a sample of correct and incorrect verb forms produced by the participants of the study:

(40) SUBJECT: The girl and boy eats [verb] a breakfast. [Not counted as a CV]

(41) SUBJECT: Two boys broke up [verb] the bread. [Not counted as a CV]

(42) INVESTIGATOR: Now it's your turn. Ask me a few questions.
SUBJECT: I like [verb] pasta? [Not counted as a CV]

(43) INVESTIGATOR: What do you usually do at the weekend?
SUBJECT: I have [verb] football. [Not counted as a CV]

(44) INVESTIGATOR: What do you usually do at the weekend?
SUBJECT: Play [verb] [CV] football.

(45) SUBJECT: The dog is [verb] [CV] in it.

(46) SUBJECT: The dog is looking at [verb] [CV] the map.

As previously presented, examples (41) and (43) are not computed as correct verb forms due to lexical reasons. Even though morphological requirements (i.e. verb inflection) within a sentence are met, the verbs used are not the correct ones in terms of meaning. Example (40), on the other hand, infringed morphological rules of subject verb agreement, which is why it was not counted as a correct verb form. Finally,

example (42) shows a lexically acceptable verb within the context of the sentence and interaction but lacks the auxiliary ‘ do’ to mark the interrogative modality of the sentence.

5.5.2.3 Fluency

In regard to fluency, the main focus of this study is on speed fluency, thus, the temporal variable used in this investigation to study the progress young language learners underwent in this area was speech rate in words per minute including pause time (SPRW), which has been proved to be a reliable indicator of oral fluency among non-native speakers (Kormos and Denés 2004; Mora, 2006, Bret 2011). SPRW was calculated by dividing the total number of words in English (TNWE) by the total task time (TTT) in minutes.

Besides looking into speed fluency, this investigation also used an L1 word ratio (L1WR) presented as a percentage to study breakdown (dys)fluency. As Mora (2006) points out, the use of L1 words at low stages of acquisition can be used as a “strong indicator” of non-fluent speech. This measure was obtained by dividing the number of L1 words by the total number of words (TNW) produced.

To sum up, the measures selected for this study to trace the development of complexity, accuracy and fluency over a period of two years (see Table 5.9) were chosen on the basis of their applicability to the data collected and previous research. However, the units of analysis employed in previous research did not seem to be suitable for the analysis of young language learners’ output, as most of them are exclusively used in the analysis of intermediate and advanced language learners. Consequently, in an attempt to capture progress at beginner levels of language development a new *unit* was designed.

Table 5.9
Measures selected for the analysis of young learners' oral output.

Dimension of L2 oral performance		Measures
Propositional complexity		Total number of units (excluding one-word units)
Linguistic complexity	**Syntactic complexity**	% Coordinate units % Subordinate units
	Lexical complexity (Lexical diversity)	Noun ratio Verb ratio Adjective ratio
Grammatical Accuracy	**Global accuracy**	% Error-free units
	Specific accuracy	% Correct verb forms
Fluency	**Speed Fluency**	Speech rate in words
	Breakdown dysfluency	L1 word ratio

5.5.3 Statistical analyses

In order to study the differences between the scores CLIL and non-CLIL learners achieved in the four data collection times in each of the measures and the improvement learners unfolded, generalized linear mixed models were used with SAS v 9.2. The covariates included in the models were data collection time (T0, T1, T2 and T3), group (CLIL/non-CLIL) and the learners' initial proficiency level in English (high and low achievers). Adjusted means with 95% confidence intervals were estimated for each group at the four times and the significance level was set at $p = 0.05$. In addition to these analyses, Spearman's correlation tests were also conducted in order to study the relationship between complexity, accuracy and fluency as a construct.

5.5.4 Qualitative analyses

In an attempt to further analyse the output produced by young language learners in the interview and narrative task at the four data collection times, a detailed description of the linguistic features and structures employed by a selection of six learners (three learners from the CLIL group and three learners from the non-CLIL group) with different fluency levels was carried out to examine the language features produced and their evolution. Additionally, CLIL and non-CLIL learners' output during the narrative task was also analysed using the stages of narrative discourse development proposed by Álvarez (2006). In the next chapter the results obtained from both the quantitative and qualitative data analyses will be presented.

CHAPTER 6 Results

This chapter presents the results of the statistical and qualitative analyses conducted to answer the research questions posited in the dissertation. As previously mentioned, the statistical analyses of the scores obtained by the learners were performed using longitudinal logistic regression models. Adjusted means with 95% confidence intervals were estimated for each group at different times and time periods. In regard to the qualitative analyses, this chapter will present a description of the evolution of the scores obtained in the CAF measures by a selection of CLIL and non-CLIL learners at an individual level. Additionally, a detailed examination of the learners' oral output in terms of linguistic features and structures associated with lexical and syntactic complexity, accuracy and fluency will be provided.

This chapter starts with the presentation of the results obtained from the statistical analyses in each of the CAF measures used to analyse the L2 English oral output of learners who have been exposed to CLIL and EFL instruction (CLIL group) and learners who have only received EFL instruction (control group). The results will be provided for each of the four data collection times (T0, T1, T2 and T3) and task. The results of the interview task will be reported first followed by the results obtained from the narrative task. Intergroup comparisons of these results will be presented and complemented with intragroup comparisons with the aim of analysing the evolution of the two groups of learners separately between different time periods (T0-T1, T1-T2, T2-T3 and T0-T3) in the two tasks.

In addition, interactions between type of instruction received and proficiency level of the learners (high and low achievers) for each of the CAF measures will be presented in an attempt to analyse the impact of CLIL instruction in combination with

the learners' proficiency level at the start of the study on the learners' oral production skills in English. After that, the results obtained from Spearman's rho correlation tests between accuracy-complexity, complexity-fluency and accuracy-fluency at the four times of data collection will be provided with the objective of analysing the relationship between the three dimensions that make up the construct of CAF.

Finally, the results obtained from the qualitative and descriptive analyses will be provided. First, the evolution of 3 CLIL learners (a highly fluent, a fluent and a dysfluent learner) and 3 non-CLIL learners (a highly fluent, a fluent and a dysfluent learner) in the CAF measures will be presented and discussed in relation to the interdependency of the three CAF elements in the interview and narrative tasks. In addition, a description of the language features and structures produced by the selection of young learners in the two tasks at the four data collection times will be presented. Finally, the learners' oral output produced in the narrative task will be also analysed in terms of their narrative competence using Álvarez's (2006) stages of oral narrative development.

6.1 Statistical analysis

6.1.1 Complexity

6.1.1.1 Propositional complexity

In this section the results obtained by CLIL and non-CLIL learners (control group) in the measure used to analyse propositional complexity (i.e. the number of ideas transmitted by the learner) in the interview and picture-elicited narrative will be reported. The measure employed to study propositional complexity is the total number of units (TNU).

6.1.1.1.1 Total number of units

6.1.1.1.1.1 Interview task

As for the total number of units (TNU), Table 6.1 displays the results obtained from the intergroup comparisons between the CLIL and control groups in the interview task. As seen from the table, there are no significant differences between CLIL and control learners at T0 (F= 3.23, p= 0.0755), even though the control group starts with a slight advantage over the CLIL group as regards the total mean score. At T1, however, after 122.5 hours of English input the results do show that the total number of units produced by the control group is significantly higher than that of the CLIL group (F =8.89, p= 0.004). This is also confirmed by the percentages of improvement displayed in Table 6.2, which show that the control group improves 63.2 % from T0 to T1. At T2 and T3 the differences between the two groups disappear, as learners from the CLIL group show greater improvement than non-CLIL learners (see Table 6.2). The mean scores, though, keep showing that control learners produced a higher number of units in general (see Figure 6.1).

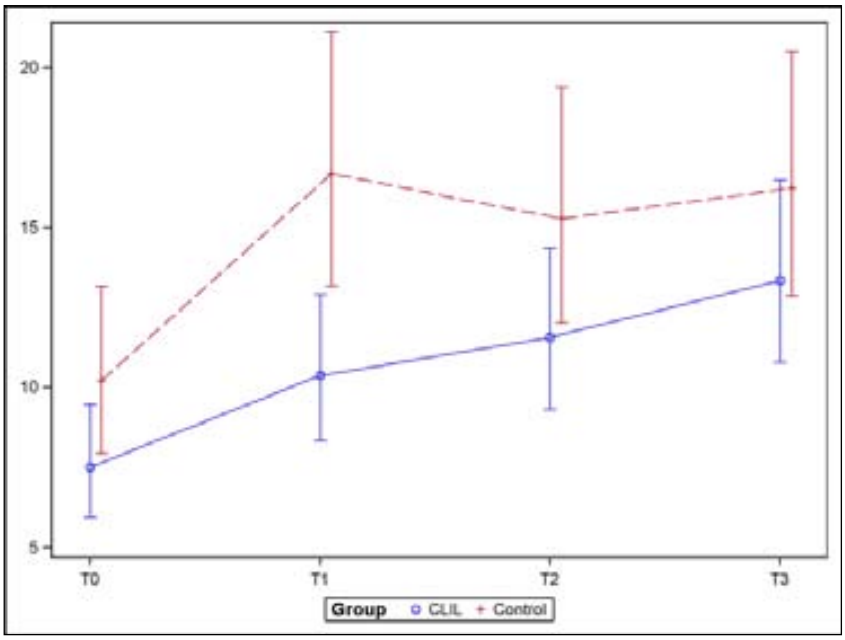
Table 6.1
Intergroup comparisons of the total number of unit in interview task.

	Group	Mean	StdErr	Lower	Upper	F value	p-value
T0	CLIL	7.53	0.88	5.97	9.49	3.23	0.0755
	Control	10.23	1.30	7.95	13.18		
T1	CLIL	10.39	1.14	8.35	12.93	8.89	0.004*
	Control	16.70	1.97	13.19	21.14		
T2	CLIL	11.57	1.25	9.33	14.36	3.12	0.0818
	Control	15.29	1.83	12.03	19.43		
T3	CLIL	13.35	1.41	10.81	16.49	1.62	0.2080
	Control	16.25	1.90	12.86	20.54		

Table 6.2
Percentages of improvement in the total number of units in the interview task.

	I0_1	I1_2	I2_3	I0_3
CLIL	38.0%	11.4%	15.4%	77.4%
Control	63.2%	-8.4%	6.3%	58.8%

Figure 6.1
Achievement mean scores of the total number of units in interview task.



The intragroup comparisons conducted show the degree of improvement learners unfolded between time periods in the TNU produced in the interview task. Overall, the CLIL group shows a linear increase whereas the control group shows a sharp increase from T0 to T1 followed by a slight decrease from T1 to T2. From T2 to T3, the control group displays an even increase (See Figure 6.1). As displayed in Table 6.3, the CLIL group improves significantly from T0 to T1 ($p= 0.0080$) and from T0 to T3 ($p<.0001$), the time period covered by this research project. Likewise, the control group obtains significant improvement from T0 to T1 ($p<.0001$) and from T0 to T3 ($p<.0001$).

Table 6.3
Intragroup comparisons of the total number of units in the interview task.

	Time	Time	t Value	p-value
CLIL	T0	T1	3.25	0.0080*
	T1	T2	1.21	0.6221
	T2	T3	1.69	0.3331
	T0	T3	6.02	<.0001*
Control	T0	T1	4.93	<.0001*
	T1	T2	-0.98	0.7635
	T2	T3	0.68	0.9046
	T0	T3	4.69	<.0001*

As regards the interactions between type of instruction and the learners' English proficiency level, Table 6.4 shows that high achievers from the control group performs significantly better at T0 (F=6.73, p= 0.0113), T1 (F= 6.21, p= 0.0154) and at T2 (F= 8.24, p= 0.0057) than high achievers in the CLIL group, whereas at T3 the differences between high achievers in the CLIL and the control groups disappear. As regards low achievers, Table 6.5 shows that there are no significant differences between the CLIL and control groups in the total number of units produced during the interview task.

Table 6.4
Intergroup comparisons of the total number of units in the interview task. Interactions between high achievers and type of instruction.

	Proficiency	Group	Mean	StdErr	Lower	Upper	F Value	p- value
T0	High	CLIL	7.16	1.08	5.31	9.65	6.73	0.0113*
	High	Control	13.03	2.13	9.40	18.06		
T1	High	CLIL	10.35	1.45	7.83	13.68	6.21	0.0154*
	High	Control	17.85	2.85	12.98	24.56		
T2	High	CLIL	11.77	1.62	8.95	15.49	8.24	0.0057*
	High	Control	21.83	3.39	16.00	29.80		
T3	High	CLIL	14.09	1.89	10.78	18.42	2.75	0.1030
	High	Control	20.06	3.09	14.72	27.33		

Table 6.5
Intergroup comparisons of the total number of units in the interview task. Interactions between low achievers and type of instruction.

	Proficiency	Group	Mean	StdErr	Lower	Upper	F Value	p- value
T0	Low	CLIL	7.92	1.27	5.76	10.89	0.00	0.9512
	Low	Control	8.04	1.64	5.36	12.05		
T1	Low	CLIL	10.43	1.59	7.70	14.14	3.40	0.0689
	Low	Control	15.62	2.92	10.76	22.69		
T2	Low	CLIL	11.38	1.69	8.47	15.29	0.08	0.7804
	Low	Control	10.71	2.02	7.35	15.60		
T3	Low	CLIL	12.65	1.83	9.47	16.90	0.04	0.8499
	Low	Control	13.16	2.40	9.14	18.96		

The intragroup comparisons displayed in Table 6.6 indicate that all four groups (CLIL high achievers, CLIL low achievers, control high achievers and control low achievers) improved significantly from T0 to T3 mainly. Only CLIL high achievers and control low achievers show significant improvement from T0 to T1 as well (see table 6.7 for the percentages of improvement). As seen from Figure 6.2, high and low achievers in the CLIL group perform similarly throughout the two years of this research project in regard to the total number of units produced in the interview task. No significant differences were found between CLIL high and low achievers at any of the four data collection times. High and low achievers from the control group, on the other hand, show different evolutions. High achievers display a regular increase from T0 to T2 followed by a slight decrease while low achievers' evolution includes sharp increases and decreases from T0 to T2. From T2 to T3, however, low achievers in the control group show a moderate increase. As for the differences between control high and low achievers, the statistical analysis conducted shows that significant differences were found within the control group at T2 ($F=8.20$, $p=0.0056$) in favour of control high achievers. At T3, this difference disappears.

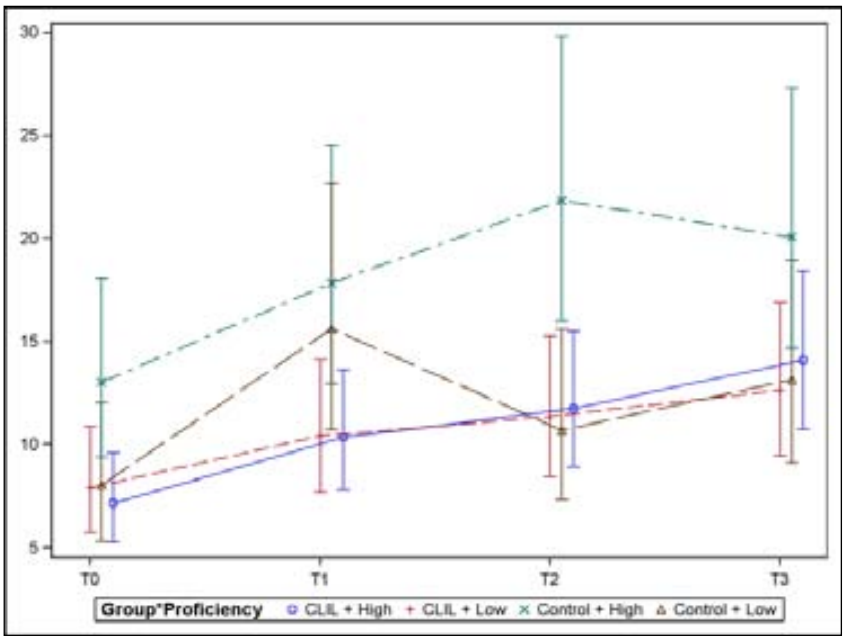
Table 6.6
Intragroup comparisons of the total number of units in interview task. Interactions between instruction type and proficiency level.

	Time	Time	t Value	p-value
CLIL-High	T0	T1	2.91	0.0217*
	T1	T2	1.15	0.6570
	T2	T3	1.72	0.3167
	T0	T3	5.61	<.0001*
CLIL-Low	T0	T1	1.95	0.2111
	T1	T2	0.69	0.9021
	T2	T3	0.90	0.8028
	T0	T3	3.53	0.0032*
Control-High	T0	T1	2.53	0.0599
	T1	T2	1.76	0.2993
	T2	T3	-0.79	0.8566
	T0	T3	3.65	0.0021*
Control-Low	T0	T1	3.86	0.0010*
	T1	T2	-2.46	0.0716
	T2	T3	1.39	0.5049
	T0	T3	2.95	0.0192*

Table 6.7
Percentages of improvement in the total number of units according to proficiency and group in the interview task.

Proficiency	Group	I0_1	I1_2	I2_3	I0_3
Low	CLIL	31.6%	9.1%	11.1%	59.7%
	Control	94.2%	-31.4%	22.8%	63.6%
High	CLIL	44.5%	13.7%	20.3%	96.7%
	Control	37%	22.2%	-5.6%	58%

Figure 6.2
Achievement scores of the total number of units in the interview task. Interactions between Instruction type and proficiency level.



6.1.1.1.1.2 Narrative task

Table 6.8 shows the results obtained from the intergroup comparisons in the narrative task by CLIL and control groups in the total number of units (TNU). As shown in the table, there are no significant differences at T0 between CLIL and non-CLIL learners ($F= 3.77$, $p=0.0548$), although the difference nearly reaches significance. At T1, however, the control group performs significantly better than the CLIL group in this measure ($F=7.43$, $p=0.0078$). As found in the interview task, the differences between CLIL and non-CLIL learners disappear at T2 and T3, as a result of the increase displayed in CLIL learners' mean scores. As seen from Table 6.9, the percentages of improvement are higher for the CLIL group from T1 onwards. Despite that, as Figure 6.3 shows, the mean scores are higher for the non-CLIL group at the four times of data collection.

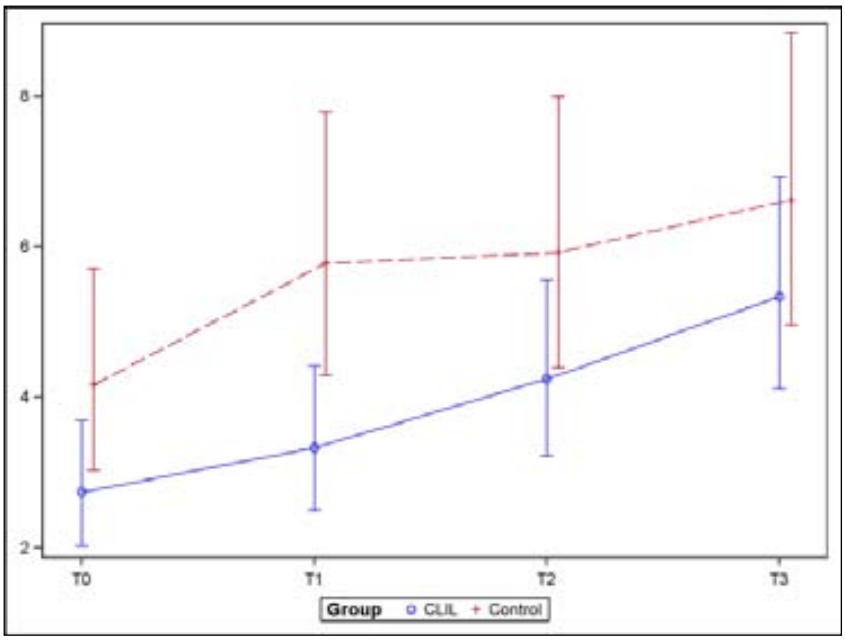
Table 6.8
Intergroup comparisons of the total number of units in the narrative task.

	Group	Mean	StdErr	Lower	Upper	F value	p-value
T0	CLIL	2.74	0.41	2.03	3.70	3.77	0.0548
	Control	4.16	0.66	3.03	5.71		
T1	CLIL	3.33	0.48	2.50	4.42	7.43	0.0078*
	Control	5.78	0.86	4.30	7.79		
T2	CLIL	4.24	0.58	3.23	5.57	2.88	0.0938
	Control	5.93	0.89	4.40	8.00		
T3	CLIL	5.34	0.70	4.11	6.93	1.29	0.2604
	Control	6.63	0.96	4.96	8.85		

Table 6.9
Percentages of improvement in the total number of units in the narrative task.

	I0_1	I1_2	I2_3	I0_3
CLIL	21.6%	27.4%	25.9%	95.1%
Control	38.9%	2.6%	11.7%	59.2%

Figure 6.3
Achievement mean scores of the total number of units in the narrative task.



The results obtained from the intragroup comparisons in the narrative task in terms of the total number of units produced by the learners (see Table 6.10) show that both groups have significant improvement from T0 to T3 (p<.0001 for the CLIL group and p=0.0065 for the control group). As seen from Figure 6.3, the two groups show a linear increase over the two years of the investigation, although the control group displays a sharper increase during the first time period (from T0 to T1).

Table 6.10
Intragroup comparisons of the total number of units in the narrative task.

	Time	Time	t Value	p-value
CLIL	T0	T1	1.30	0.5618
	T1	T2	1.79	0.2812
	T2	T3	1.90	0.2325
	T0	T3	4.86	<.0001*
Control	T0	T1	2.27	0.1104
	T1	T2	0.19	0.9976
	T2	T3	0.85	0.8298
	T0	T3	3.31	0.0065*

In regard to the interaction between type of instruction and the learners' English proficiency level, Table 6.11 shows that high achievers from the control group perform significantly better than high achievers from the CLIL group at T1 (F=5.12, p=0.0265). As for T2 and T3, there are no significant differences between the groups in the total number of units produced in the narrative task. Table 6.12, on the other hand, shows no significant differences between low achievers from the two groups in any of the data collection times.

Table 6.11
Intergroup comparisons of the total number of units in the narrative task. Interactions between high achievers and type of instruction.

	Proficiency	Group	Mean	StdErr	Lower	Upper	F Value	p-value
T0	High	CLIL	2.88	0.56	1.96	4.23	1.76	0.1874
	High	Control	4.27	0.90	2.81	6.50		
T1	High	CLIL	3.36	0.62	2.33	4.84	5.12	0.0265*
	High	Control	6.34	1.27	4.25	9.46		
T2	High	CLIL	4.17	0.74	2.94	5.93	2.66	0.4117
	High	Control	6.52	1.29	4.39	9.69		
T3	High	CLIL	5.27	0.88	3.77	7.37	2.96	0.0908
	High	Control	8.26	1.55	5.67	12.02		

Table 6.12
Intergroup comparisons of the total number of units in the narrative task. Interactions between low achievers and type of instruction.

	Proficiency	Group	Mean	StdErr	Lower	Upper	F Value	p-value
T0	Low	CLIL	2.60	0.56	1.70	3.97	2.21	0.1403
	Low	Control	4.06	1.03	2.45	6.71		
T1	Low	CLIL	3.30	0.66	2.22	4.90	2.79	0.0980
	Low	Control	5.28	1.26	3.28	8.49		
T2	Low	CLIL	4.31	0.81	2.96	6.27	0.68	0.4117
	Low	Control	5.40	1.28	3.37	8.64		
T3	Low	CLIL	5.41	0.96	3.79	7.72	0.00	0.9509
	Low	Control	5.32	1.23	3.35	8.44		

The intragroup comparisons in Table 6.13 show that CLIL high and low achievers as well as control high achievers improved significantly from T0 to T3. As seen from Table 6.14, these three groups of learners display the highest percentages of improvement from T0 to T3, being CLIL low achievers the group with the highest percentage. No significant improvement is found in any other time periods. As seen

from the mean scores obtained in the interview task (see Figure 6.4), high and low achievers from the CLIL group also have similar development in the narrative task as regards the total number of units they produced. Similarly, high and low achievers from the control group display moderate increases from T0 to T2. From T2 to T3 high achievers from the control group show a sharper increase while control low achievers' scores decrease slightly. Despite this, no differences were found between control high and low achievers.

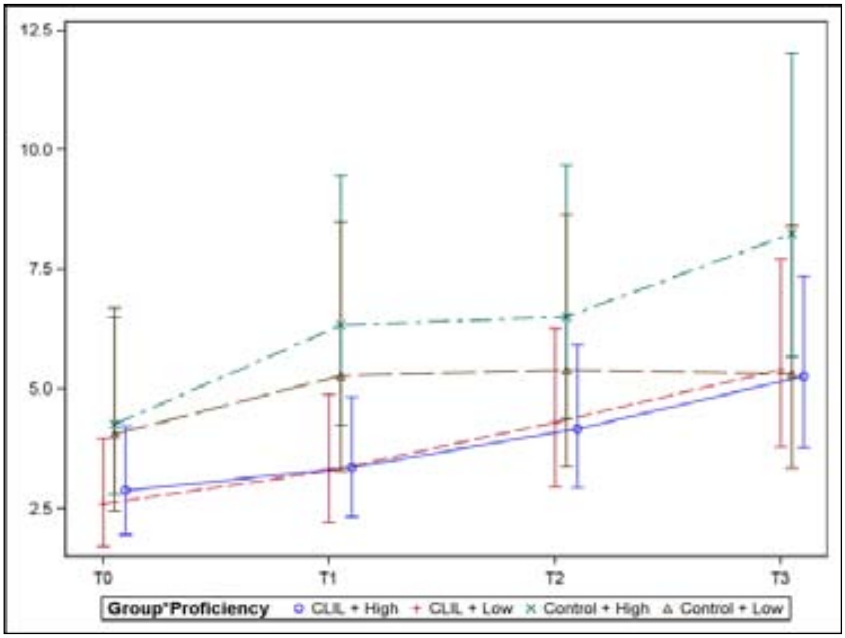
Table 6.13
Intragroup comparisons of the total number of units in the narrative task. Interactions between instruction type and proficiency level.

	Time	Time	t Value	p-value
CLIL -High	T0	T1	0.81	0.8494
	T1	T2	1.25	0.5942
	T2	T3	1.50	0.4389
	T0	T3	3.46	0.0040*
CLIL -Low	T0	T1	1.10	0.6889
	T1	T2	1.40	0.4991
	T2	T3	1.35	0.5350
	T0	T3	3.74	0.0016*
Control -High	T0	T1	2.08	0.1646
	T1	T2	0.16	0.9985
	T2	T3	1.47	0.4561
	T0	T3	3.75	0.0015*
Control -Low	T0	T1	1.07	0.7064
	T1	T2	0.10	0.9996
	T2	T3	-0.07	0.9999
	T0	T3	1.14	0.6640

Table 6.14
Percentages of improvement in the total number of units according to proficiency and group in the narrative task.

Proficiency	Group	I0_1	I1_2	I2_3	I0_3
Low	CLIL	26.9%	30%	25.5%	108%
	Control	30%	2.2%	-1.4%	31%
High	CLIL	16.6%	24.1%	26.3%	82.9%
	Control	48.7%	2.8%	26.6%	93.44%

Figure 6.4
Achievement scores of the total number of units in the narrative task. Interactions between instruction type and proficiency level.



6.1.1.2 Linguistic complexity: syntactic complexity

This section provides the results obtained by the CLIL and the control groups in the syntactic complexity measures employed to examine the learners' oral performance on the interview and narrative task. The measures used to study syntactic complexity are: the percentage of coordinate units (% CU) and the percentage of subordinate units (% SU).

6.1.1.2.1 Percentage of coordinate units

6.1.1.2.1.1 Interview task

As shown in Table 6.15, the statistical analyses indicate a significant difference in favour of the control group at T1 (F= 4.71, P =0.0318) and T2 (F= 5.23, p=0.0240). At T3 the differences between the groups disappear (F= 1.31, p= 0.2553) due to the increase in the production of coordinate units of CLIL learners and the decrease of the control group, as the percentages of improvement displayed in Table 6.16 indicate.

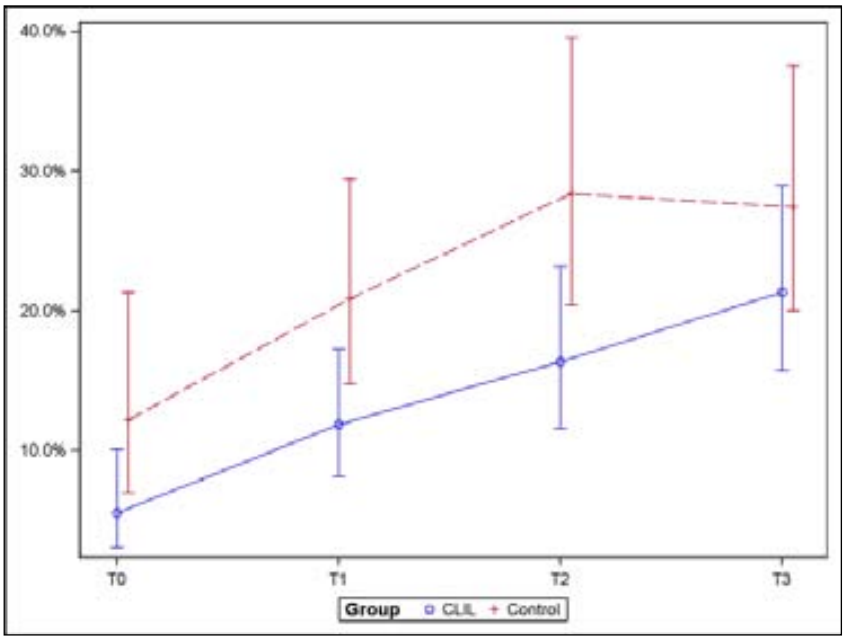
Table 6.15
Intergroup comparisons of the percentages of coordinate units in the interview task.

	Group	Mean	StdErr	Lower	Upper	F value	p-value
T0	CLIL	5.5%	1.7%	3.0%	10.1%	3.37	0.0682
	Control	12.2%	3.5%	7.0%	21.3%		
T1	CLIL	11.9%	2.3%	8.2%	17.3%	4.71	0.0318*
	Control	20.9%	3.6%	14.8%	29.5%		
T2	CLIL	16.4%	2.9%	11.6%	23.2%	5.23	0.0240*
	Control	28.4%	4.7%	20.4%	39.5%		
T3	CLIL	21.3%	3.3%	15.7%	28.9%	1.31	0.2553
	Control	27.4%	4.3%	20.0%	37.6%		

Table 6.16
Percentages of improvement in the percentage of coordinate units in the interview task.

	I0_1	I1_2	I2_3	I0_3
CLIL	114.4%	37.9%	30.3%	285.2%
Control	71.1%	36.2%	-3.5%	124.9%

Figure 6.5
Mean percentages of coordinate units in the interview task.



The results obtained from the intragroup analyses shown in Table 6.17 indicate that both groups improve significantly from T0 to T3 in terms of the percentage of coordinate units produced in the interview task (p= 0.0002 for the CLIL group and p= 0.0297 for the control group). As seen from Figure 6.5, both groups display a linear increase from T0 to T2. At T3, however, the control group shows a slight decrease in the number of coordinate units, whereas the scores obtained by the CLIL group continue increasing.

Table 6.17
Intragroup comparisons of the percentages of coordinate units in the interview task.

	Time	Time	t Value	p-value
CLIL	T0	T1	2.31	0.1018
	T1	T2	1.46	0.4659
	T2	T3	1.38	0.5141
	T0	T3	4.30	0.0002*
Control	T0	T1	1.81	0.2736
	T1	T2	1.63	0.3636
	T2	T3	-0.20	0.9970
	T0	T3	2.80	0.0297*

The analyses conducted to explore the effects of proficiency and instruction type on the percentage of coordinate units produced are shown in Tables 6.18 and 6.19. As regards the differences between high achievers in the CLIL and control groups, Table 6.18 shows no significant differences between the groups over the two years of the study. The data displayed in Table 6.19, on the other hand, does report significant differences between low achievers at T2 in favour of low achievers from the control group (F= 4.87, p= 0.0291). At T3, as seen in previous analyses, this difference between groups disappears.

Table 6.18
Intergroup comparisons of the percentages of coordinate units in the interview task. Interactions between high achievers and type of instruction.

	Proficiency	Group	Mean	StdErr	Lower	Upper	F Value	p-value
T0	High	CLIL	6.8%	2.3%	3.6%	13.1%	2.93	0.0888
	High	Control	14.0%	3.9%	8.0%	24.4%		
T1	High	CLIL	13.5%	3.0%	8.7%	20.9%	3.45	0.0664
	High	Control	24.2%	5.5%	15.4%	38.0%		
T2	High	CLIL	17.6%	3.8%	11.5%	27.0%	1.25	0.2659
	High	Control	24.9%	5.2%	16.4%	37.9%		
T3	High	CLIL	24.0%	4.5%	16.6%	34.8%	0.05	0.8253
	High	Control	25.6%	5.2%	17.0%	38.4%		

Table 6.19
Intergroup comparisons of the percentages of coordinate units in the interview task. Interactions between low achievers and type of instruction.

	Proficiency	Group	Mean	StdErr	Lower	Upper	F Value	p-value
T0	Low	CLIL	4.5%	2.3%	1.7%	12.2%	1.45	0.2304
	Low	Control	10.6%	5.7%	3.7%	30.8%		
T1	Low	CLIL	10.5%	3.0%	5.9%	18.5%	1.91	0.1692
	Low	Control	18.0%	5.5%	9.8%	33.0%		
T2	Low	CLIL	15.2%	3.9%	9.2%	25.1%	4.87	0.0291*
	Low	Control	32.4%	8.8%	18.9%	55.5%		
T3	Low	CLIL	19.0%	4.2%	12.2%	29.4%	2.06	0.1546
	Low	Control	29.4%	7.4%	17.8%	48.5%		

The intragroup comparisons displayed in Table 6.20 show that high and low achievers from the CLIL group improve significantly from T0 to T3 while no significant improvement is found in the control group at any of the time periods analysed. As seen from Figure 6.6, the development of CLIL high and low achievers is quite similar in the production of coordination as both groups show an increase. However, the percentages of improvement displayed in Table 6.21 show that CLIL low achievers usually hold higher percentages of development than CLIL high achievers. As

for the development of high and low achievers from the control group, Figure 6.6 shows linear development of high achievers from the control group, while low achievers from the control group display a rapid increase from T1 to T2 and a slight decrease from T2 to T3.

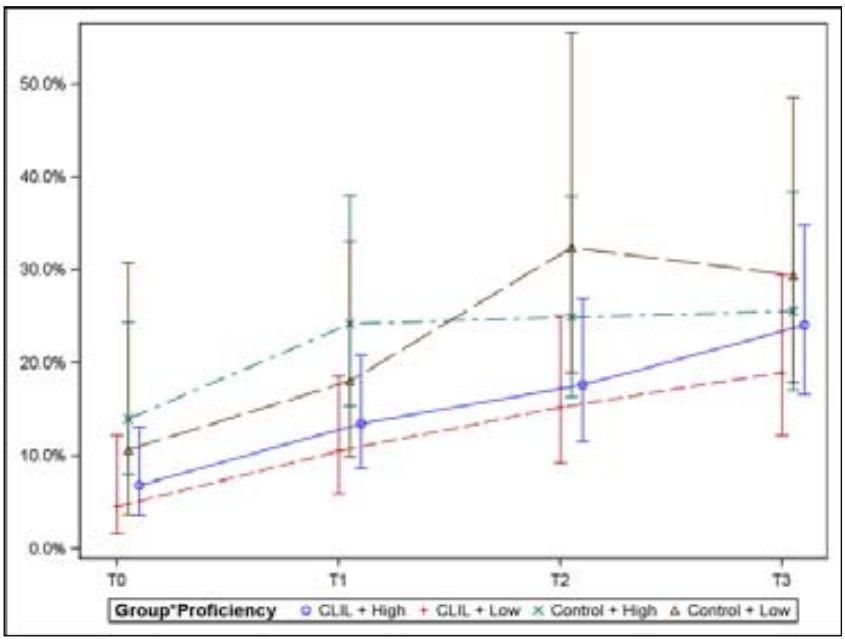
Table 6.20
Intragroup comparisons of the percentages of coordinate units in the interview task. Interactions between instruction type and proficiency level.

	Time	Time	t Value	p-value
CLIL -High	T0	T1	1.90	0.2330
	T1	T2	1.05	0.7186
	T2	T3	1.35	0.5313
	T0	T3	3.72	0.0016*
CLIL -Low	T0	T1	1.53	0.4217
	T1	T2	1.10	0.6884
	T2	T3	0.77	0.8669
	T0	T3	2.75	0.0339*
Control -High	T0	T1	1.80	0.2778
	T1	T2	0.13	0.9992
	T2	T3	0.12	0.9994
	T0	T3	2.10	0.1586
Control -Low	T0	T1	0.92	0.7951
	T1	T2	1.73	0.3131
	T2	T3	-0.34	0.9869
	T0	T3	1.85	0.2537

Table 6.21
Percentages of improvement in the percentage of coordinate units according to proficiency and group in the interview task.

Proficiency	Group	I0_1	I1_2	I2_3	I0_3
Low	CLIL	133.3%	44.7%	25%	322.2%
	Control	69.8%	80%	-9.2%	177.3%
High	CLIL	98.5%	30.3%	36.36%	252.9%
	Control	72.8%	2.8%	2.8%	82.8%

Figure 6.6
Mean percentages of coordinate units in the interview task. Interactions between instruction type and proficiency level.



6.1.1.2.1.2 Narrative task

Regarding the production of coordination in the narrative task and unlike the interview task, Table 6.22 shows that there are no differences between the groups at any of the times analysed (T1, T2 and T3). The instances of coordination found at T0 by CLIL and control learners, on the other hand, were not enough so as to conduct the statistical analyses. In addition, as Figure 6.7 shows, the achievement scores obtained by the control group are higher than the CLIL group at the three times. However, the percentages of improvement displayed in Table 6.23 show that the degree of development is higher for the CLIL group over the three time periods studied. Regarding the intragroup comparisons displayed in Table 6.24, no significant development was found for either of the groups.

Table 6.22
Intergroup comparisons of the percentages of coordinate units in the narrative task.

	Group	Mean	StdErr	Lower	Upper	F value	p-value
T1	CLIL	27.9%	5.5%	18.9%	41.1%	3.05	0.0836
	Control	43.9%	7.8%	30.8%	62.4%		
T2	CLIL	24.2%	4.8%	16.3%	35.8%	1.81	0.1812
	Control	35.2%	7.1%	23.6%	52.6%		
T3	CLIL	31.9%	5.2%	23.0%	44.0%	1.05	0.3083
	Control	40.2%	6.8%	28.7%	56.4%		

Table 6.23
Percentages of improvement in the percentage of coordinate units in the narrative task.

	I1_2	I2_3	I1_3
CLIL	-13.2%	31.8%	14.4%
Control	-19.7%	14.3%	-8.3%

Figure 6.7
Mean percentages of coordinate units in the narrative task.

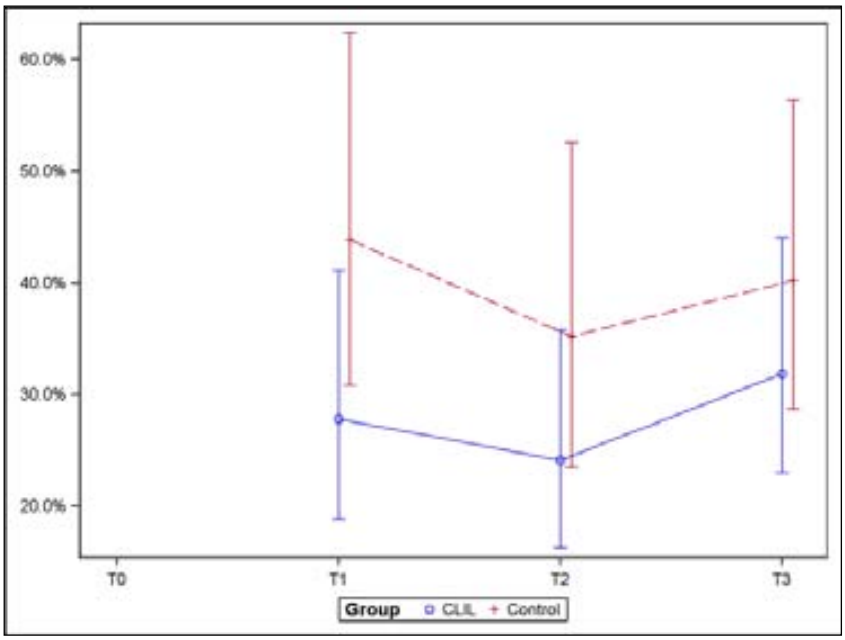


Table 6.24
Intragroup comparisons of the percentage of coordinate units in the narrative task.

	Time	Time	t Value	p-value
CLIL	T1	T2	-0.56	0.8440
	T2	T3	1.20	0.4589
	T1	T3	0.58	0.8298
Control	T1	T2	-0.92	0.6314
	T2	T3	0.57	0.8363
	T1	T3	-0.41	0.9132

As for the interactions between type of instruction and proficiency level in the production of coordination in the narrative task, Table 6.25 shows that even though the mean scores are higher for high achievers from the control group at the three times, these differences do not reach significance until T3 (F= 5.41, p= 0.0232). In contrast, as shown in Table 6.26, no significant differences were found between low achievers from the CLIL and control groups at any of the times. However, it is worth pointing out that despite the fact that low achievers from the CLIL group produced fewer instances of coordination at T1 than low achievers from the control group, low achievers from the CLIL group out perform low achievers from the control group in the coordination measure at T3.

Table 6.25
Intergroup comparisons of the percentages of coordinate units in the narrative task. Interactions between high achievers and type of instruction.

	Proficiency	Group	Mean	StdErr	Lower	Upper	F Value	p-value
T1	High	CLIL	24.5%	6.5%	14.5%	41.5%	3.24	0.0752
	High	Control	45.5%	9.8%	29.5%	70.1%		
T2	High	CLIL	28.5%	6.9%	17.7%	45.9%	2.52	0.1169
	High	Control	47.1%	10.0%	30.7%	72.1%		
T3	High	CLIL	25.4%	5.6%	16.4%	39.4%	5.41	0.0232*
	High	Control	51.1%	9.7%	34.8%	74.9%		

Table 6.26
Intergroup comparisons of the percentages of coordinate units in the narrative task. Interactions between low achievers and type of instruction.

	Proficiency	Group	Mean	StdErr	Lower	Upper	F Value	p-value
T1	Low	CLIL	31.6%	8.9%	18.1%	55.3%	0.58	0.4472
	Low	Control	42.3%	13.1%	22.9%	78.2%		
T2	Low	CLIL	20.5%	6.0%	11.5%	36.7%	0.32	0.5741
	Low	Control	26.3%	9.4%	13.0%	53.2%		
T3	Low	CLIL	39.9%	8.2%	26.5%	60.3%	0.49	0.4871
	Low	Control	31.7%	9.4%	17.7%	56.9%		

The intragroup analyses conducted to explore the evolution of high and low achievers from the two groups (see Table 6.27) at different time periods, show that none of the four groups undergoes significant improvement at any of the time periods analysed. However, as shown in Figure 6.8, while low and high achievers from the CLIL group along with low achievers from the control group show increases and decreases in the number of coordinate units produced from T1 to T3, high achievers from the control group is the only group that displays linear development from T1 to T3. In spite of this, no significant differences were found between high and control

achievers within the control group. In addition, as Table 6.28 displays, CLIL low achievers is the group with the highest percentage of improvement from T0 to T3.

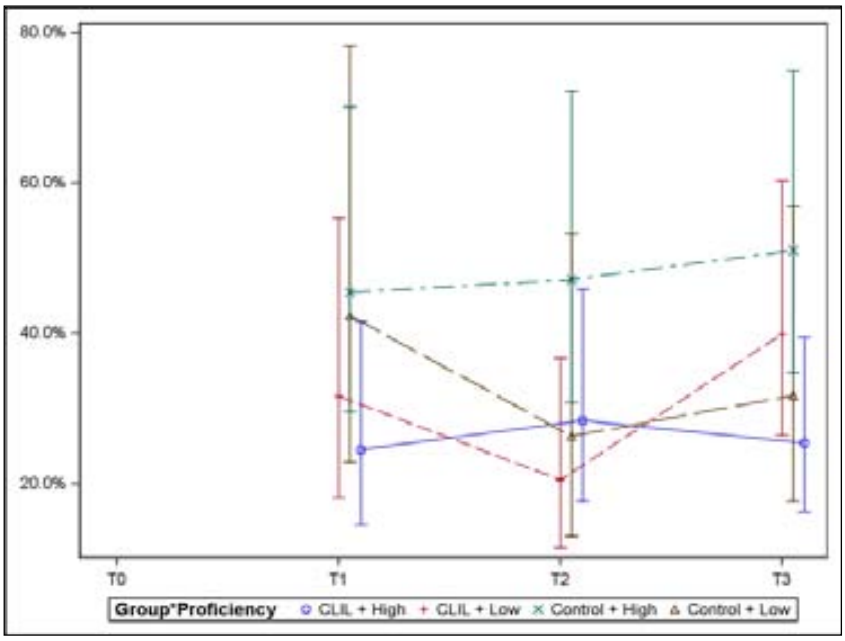
Table 6.27
Intragroup comparisons of the percentages of coordinate units in the narrative task. Interactions between instruction type and proficiency level.

	Time	Time	t Value	p-value
CLIL -High	T1	T2	0.46	0.8909
	T2	T3	-0.39	0.9199
	T1	T3	0.11	0.9932
CLIL -Low	T1	T2	-1.15	0.4890
	T2	T3	2.04	0.1077
	T1	T3	0.74	0.7414
Control -High	T1	T2	0.14	0.9896
	T2	T3	0.35	0.9343
	T1	T3	0.49	0.8741
Control -Low	T1	T2	-1.10	0.5153
	T2	T3	0.44	0.8978
	T1	T3	-0.75	0.7325

Table 6.28
Percentages of improvement in the percentage of coordinate units according to proficiency and group in the narrative task.

Proficiency	Group	I1_2	I2_3	I1_3
Low	CLIL	-35.1%	94.6%	26.26%
	Control	-37.8%	20.53	-25%
High	CLIL	16.32%	-10.8%	3.6%
	Control	3.5%	8.4%	12.3%

Figure 6.8
Mean percentages of coordinate units in the narrative task. Interactions between instruction type and proficiency level.



6.1.1.2.2 Percentage of subordinate units

Due to the limited amount of subordination found in the samples of the interview and narrative task, the statistical analyses could not be conducted for this measure. As seen from the tables provided (Table 6.29 and Table 6.30), there were only three participants out of twenty from the control group who produced subordination in the interview task at T0. In the narrative task, there was only one participant out of twenty who used subordination at T0. Likewise, the instances of subordination found in the samples produced by CLIL learners at T0 were also marginal. Two participants out of thirty-two produced subordination in the interview task while none was found in the narrative task.

Despite the limited amount of subordination produced, both tables show a gradual increase in the number of participants who produce syntactically complex language through subordination. Because of this and in an attempt to carefully examine the emergence and evolution of subordination at incipient levels of language learning in a more qualitative manner, a detailed description of the type of subordinate clauses produced by a selection of CLIL and non-CLIL learners will be provided in section 6.2 of this chapter.

Table 6.29
Number of participants from the control group that produced subordination in the interview and narrative task.

Control group	**T0**	**T1**	**T2**	**T3**
Interview	3 (20)	4 (20)	7 (20)	8 (20)
Narrative	1 (20)	4 (20)	6 (20)	7 (20)

Table 6.30
Number of participants from the CLIL group that produced subordination in the interview and narrative task.

CLIL group	**T0**	**T1**	**T2**	**T3**
Interview	2 (32)	7 (32)	4 (32)	8 (32)
Narrative	0 (32)	1 (32)	4 (32)	7 (32)

6.1.1.3 Lexical complexity: lexical diversity

As mentioned in Chapter 4, the subdimension of lexical complexity that will be examined in this study is lexical diversity. To do so, three ratios presented as percentages will be provided in this section: the percentage of the noun ratio (% NR), the percentage of the lexical verb ratio (% VR) and the percentage of the adjective ratio (% ADJR).

6.1.1.3.1 Percentage of the noun ratio

6.1.1.3.1.1 Interview task

Table 6.31 displays the intergroup group comparisons of the percentages of the noun ratios found in the speech samples of CLIL and control learners in the interview task. As seen from the table and Figure 6.9, the CLIL group produces a significantly higher number of nouns in relation to the total number of words in English (TNWE) than the control group throughout the two years of the study. Despite that, the results from both groups indicate a decrease in % NR from T0 to T3. As for the degree of decrease, Table 6.32 shows that from T0 to T3 the control group decreases the number of nouns produced to a greater extent.

Table 6.31
Intergroup comparisons of the percentages of the noun ratios in the interview task.

	Group	Mean	StdErr	Lower	Upper	F value	p-value
T0	CLIL	28.6%	1.2%	26.3%	31.1%	8.38	0.0043*
	Control	23.4%	1.2%	21.2%	26.0%		
T1	CLIL	27.3%	1.1%	25.3%	29.6%	4.88	0.0287*
	Control	24.1%	1.0%	22.2%	26.2%		
T2	CLIL	25.1%	1.0%	23.2%	27.1%	7.44	0.0071*
	Control	21.4%	1.0%	19.5%	23.3%		
T3	CLIL	26.1%	1.0%	24.2%	28.1%	16.03	<.0001*
	Control	20.8%	0.9%	19.1%	22.6%		

Table 6.32
Percentages of decrease in the percentages of noun ratios in the interview task.

	I0_1	I1_2	I2_3	I0_3
CLIL	-4.3%	-8.3%	4.1%	-8.6%
Control	2.9%	-11.5%	-2.6%	-11.3%

Figure 6.9
Mean percentages of the noun ratios in the interview task.

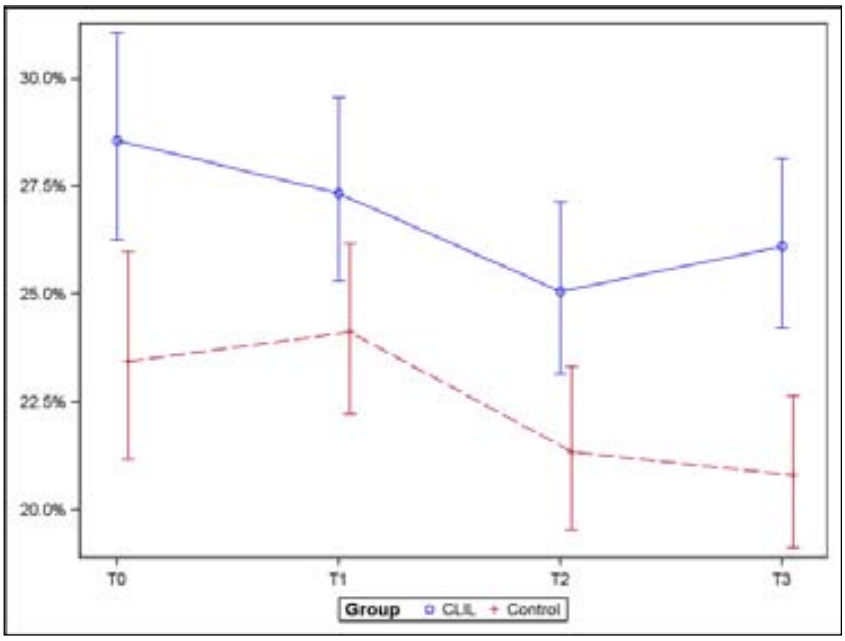


Table 6.33 displays the results obtained from the intragroup comparisons of the percentages of the noun ratios produced by the CLIL and control groups. As seen from the table, no significant increases or decreases were found for any of the time periods studied for either group.

Table 6.33
Intragroup comparisons of the percentages of the noun ratios in the interview task.

	Time	Time	t Value	P-Value
CLIL	T0	T1	-0.86	0.8270
	T1	T2	-1.77	0.2913
	T2	T3	0.84	0.8354
	T0	T3	-1.80	0.2800
Control	T0	T1	0.48	0.9626
	T1	T2	-2.32	0.0978
	T2	T3	-0.48	0.9632
	T0	T3	-2.00	0.1946

With regard to the interactions between instruction type and proficiency level, Table 6.34 shows that high achievers from the CLIL group produce a significantly

higher number of nouns than high achievers from the control group at T0, T1, T2 and T3. As for low achievers, the results displayed in Table 6.35 indicate that the only significant difference found between low achievers from the CLIL and control groups is at T3 in favour of the CLIL group.

Table 6.34
Intergroup comparisons of the percentages of the noun ratios in the interview task. Interactions between high achievers and type of instruction.

	Proficiency	Group	Mean	StdErr	Lower	Upper	F Value	p-value
T0	High	CLIL	31.4%	1.7%	28.2%	35.0%	12.77	0.0005*
	High	Control	23.1%	1.4%	20.4%	26.1%		
T1	High	CLIL	27.8%	1.4%	25.1%	30.7%	8.00	0.0054*
	High	Control	22.3%	1.3%	19.9%	25.0%		
T2	High	CLIL	25.5%	1.3%	23.1%	28.2%	10.11	0.0018*
	High	Control	19.8%	1.2%	17.7%	22.3%		
T3	High	CLIL	25.7%	1.2%	23.4%	28.3%	13.47	0.0003*
	High	Control	19.3%	1.1%	17.3%	21.6%		

Table 6.35
Intergroup comparisons of the percentages of the noun ratios in the interview task. Interactions between low achievers and type of instruction.

	Proficiency	Group	Mean	StdErr	Lower	Upper	F Value	p-value
T0	Low	CLIL	26.0%	1.6%	23.0%	29.4%	0.82	0.3678
	Low	Control	23.8%	2.0%	20.1%	28.1%		
T1	Low	CLIL	26.9%	1.5%	24.1%	30.1%	0.17	0.6770
	Low	Control	26.1%	1.8%	22.8%	29.8%		
T2	Low	CLIL	24.6%	1.4%	22.1%	27.5%	0.74	0.3919
	Low	Control	23.0%	1.6%	20.1%	26.3%		
T3	Low	CLIL	26.5%	1.4%	24.0%	29.4%	4.74	0.0310*
	Low	Control	22.4%	1.5%	19.6%	25.6%		

Table 6.36 shows the results obtained from the intragroup comparisons of the percentages of the noun ratios produced during the interview task to explore whether low and high achievers from the CLIL and control groups increased or decreased significantly in this measure. According to the results, the only significant decrease, in this case, was found in CLIL high achievers from T0 to T3 (F=-3.16, p= 0.0103), which means that the number of nouns CLIL high achievers produce from T0 to T3 decreases significantly. Additionally, it is worth pointing out that all the groups except CLIL low achievers show a decrease in this measure from T0 to T3 (see Figure 6.10). As for the evolution of low achievers, Figure 6.10 shows that low achievers from the two groups display similar trends in the number of nouns produced from T0 to T2. From T2 to T3, on the other hand, low achievers from the CLIL group show a slight increase while low achievers from the control group continue decreasing. See table 6.37 for the percentages of improvement.

Table 6.36
Intragroup comparisons of the percentages of the noun ratios in the interview task. Interactions between instruction type and proficiency level.

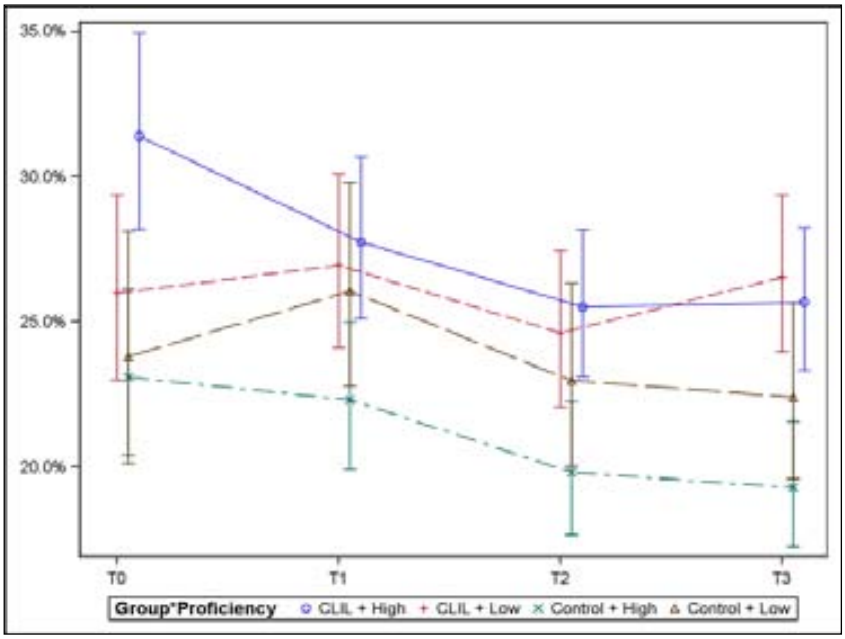
	Time	Time	t Value	p-value
CLIL -High	T0	T1	-1.89	0.2363
	T1	T2	-1.36	0.5290
	T2	T3	0.11	0.9996
	T0	T3	-3.16	0.0103*
CLIL -Low	T0	T1	0.48	0.9631
	T1	T2	-1.30	0.5660
	T2	T3	1.13	0.6689
	T0	T3	0.29	0.9914
Control -High	T0	T1	-0.46	0.9667
	T1	T2	-1.67	0.3425
	T2	T3	-0.39	0.9799
	T0	T3	-2.44	0.0737

	Time	Time	t Value	p-value
Control -Low	T0	T1	0.94	0.7845
	T1	T2	-1.50	0.4387
	T2	T3	-0.29	0.9914
	T0	T3	-0.61	0.9278

Table 6.37
Percentages of improvement in the percentages of noun ratios according to proficiency and group in the interview task.

Proficiency	Group	I0_1	I1_2	I2_3	I0_3
Low	CLIL	3.4%	-8.5%	7.7%	1.9%
	Control	9.6%	-11.8%	-2.6%	-5.8%
High	CLIL	-11.4%	-8.2%	0.7%	-18.1%
	Control	-3.4%	-11.2%	-2.5%	-16.4%

Figure 6.10
Mean percentages of the noun ratios in the interview task. Interactions between instruction type and proficiency level.



6.1.1.3.1.2 Narrative task

Tables 6.38 displays the results of the intergroup comparisons between CLIL and control learners in the percentage of the noun ratios produced in the narrative task. As seen from the table, the CLIL group produces a significantly higher number of nouns than the control group at T0 and T1 (F= 19.07, p< .0001 and F = 4.00, p= 0.0479, respectively). After T1, the differences in this measured disappear between the two groups up to T3. Additionally, as Table 6.39 displays, the CLIL group is the group who decreases the most from T0 to T3. As for the control group, Figure 6.11 shows that the scores obtained by the learners do not show much variability over the two years of the study.

Table 6.38
Intergroup comparisons of the percentages of the noun ratios in the narrative task.

	Group	Mean	StdErr	Lower	Upper	F value	p-value
T0	CLIL	38.9%	1.9%	35.4%	42.8%	19.07	<.0001*
	Control	28.5%	1.6%	25.5%	31.8%		
T1	CLIL	32.7%	1.6%	29.7%	36.0%	4.00	0.0479*
	Control	28.5%	1.5%	25.6%	31.6%		
T2	CLIL	32.0%	1.4%	29.3%	34.9%	2.38	0.1257
	Control	28.8%	1.5%	26.0%	32.0%		
T3	CLIL	30.7%	1.4%	28.0%	33.5%	2.62	0.1087
	Control	27.4%	1.4%	24.8%	30.4%		

Table 6.39
Percentages of improvement in the percentages of noun ratios in the narrative task.

	I0_1	I1_2	I2_3	I0_3
CLIL	-16.0%	-2.2%	-4.1%	-21.2%
Control	-0.1%	1.4%	-4.8%	-3.6%

Figure 6.11
Mean percentages of the noun ratios in the narrative task.

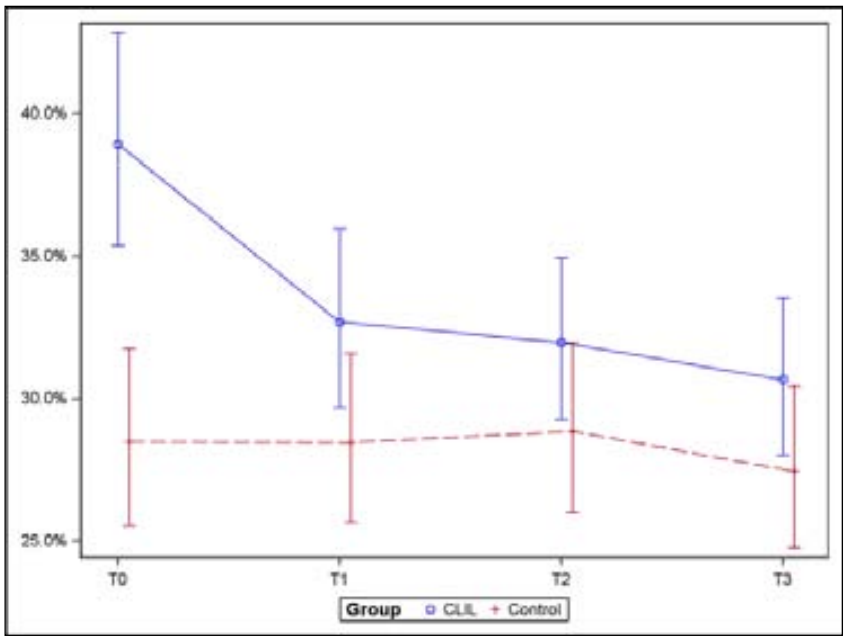


Table 6.40 displays the results obtained from the intragroup comparisons of the percentages of the noun ratios produced by CLIL and control learners for the following time periods: T0-T1, T1-T2, T2-T3 and T0-T3. As seen from the table, the CLIL group shows a significant decrease from T0 to T1 (F= -3.40, p= 0.0049) and from T0 to T3 (F= -4.94. p <.0001). The control group, on the other hand, produces a similar number of nouns at the four times of data collection.

Table 6.40
Intragroup comparisons of the percentages of the noun ratios in the narrative task.

	Time	Time	t Value	p-value
CLIL	T0	T1	-3.40	0.0049*
	T1	T2	-0.46	0.9685
	T2	T3	-0.94	0.7823
	T0	T3	-4.94	<.0001*
Control	T0	T1	-0.02	1.0000
	T1	T2	0.26	0.9941
	T2	T3	-0.94	0.7848
	T0	T3	-0.66	0.9130

Regarding the interactions between instruction type and proficiency level and their effects on the number of nouns produced, Table 6.41 shows that high achievers from the CLIL group produce a significantly higher number of nouns in proportion to the TNWE at T0 (F= 4.98, p=0.0274) than high achievers from the control group. After T0, this difference between the groups disappears gradually. Likewise, low achievers from the CLIL group (see Table 6.42) also produce a significantly higher number of nouns at T0 (F=17.66, p=<.0001) than low achievers from the control group. No other differences were found between low achievers from the CLIL and control groups in this measure at the other data collection times.

Table 6.41
Intergroup comparisons of the percentages of the noun ratios in the narrative task. Interactions between high achievers and type of instruction.

	Proficiency	Group	Mean	StdErr	Lower	Upper	F Value	p-value
T0	High	CLIL	36.0%	2.3%	31.7%	40.8%	4.98	0.0274*
	High	Control	28.6%	2.2%	24.6%	33.2%		
T1	High	CLIL	34.0%	2.1%	30.1%	38.4%	3.36	0.0694
	High	Control	28.4%	2.1%	24.6%	32.8%		
T2	High	CLIL	30.4%	1.8%	27.1%	34.1%	1.25	0.2657
	High	Control	27.4%	1.9%	23.8%	31.5%		
T3	High	CLIL	29.9%	1.7%	26.7%	33.6%	2.31	0.1319
	High	Control	25.9%	1.8%	22.5%	29.8%		

Table 6.42
Intergroup comparisons of the percentages of noun ratios in the narrative task. Interactions between low achievers and type of instruction

	Proficiency	Group	Mean	StdErr	Lower	Upper	F Value	p-value
T0	Low	CLIL	42.1%	2.8%	37.0%	47.9%	17.66	<.0001*
	Low	Control	28.4%	2.4%	24.1%	33.4%		
T1	Low	CLIL	31.4%	2.1%	27.5%	35.9%	1.08	0.3000
	Low	Control	28.5%	2.3%	24.4%	33.4%		
T2	Low	CLIL	33.6%	2.1%	29.7%	38.0%	1.25	0.2664
	Low	Control	30.4%	2.4%	26.0%	35.5%		
T3	Low	CLIL	31.4%	1.9%	27.8%	35.4%	0.68	0.4129
	Low	Control	29.1%	2.4%	24.8%	34.2%		

.

The intragroup comparisons displayed in Table 6.43 indicate that high and low achievers from the CLIL group produce a significantly lower number of nouns at T3 in comparison to T0 (F= -2.90, p= 0.0226 for high achievers and F= -4.57, p<.0001 for low achievers). Additionally, low achievers from the CLIL group also show a significant decrease from T0 to T1 (p=0.0004). As for the control group, no significant decreases were found in this measure for any of the time periods studied. As seen from

Figure 6.12, high achievers from the control group produce a similar number of nouns in proportion to the TNWE throughout the two years of the study. The same applies to los achievers from the control group. Table 6.44 shows that CLIL low achievers are the learners who decrease the most from T0 to T3.

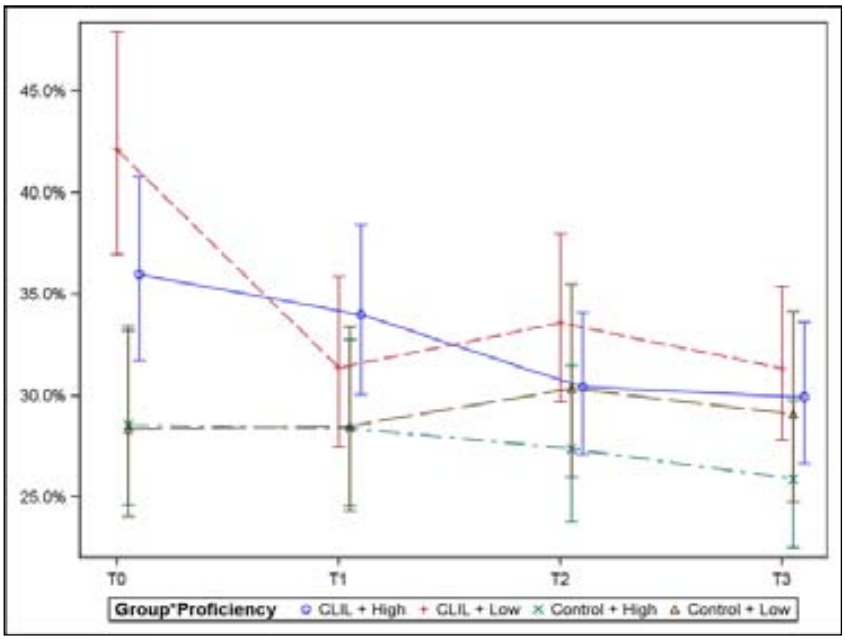
Table 6.43
Intragroup comparisons of the percentages of the noun ratios in the narrative task. Interactions between instruction type and proficiency level.

	Time	Time	t Value	p-value
CLIL -High	T0	T1	-0.85	0.8302
	T1	T2	-1.79	0.2803
	T2	T3	-0.28	0.9926
	T0	T3	-2.90	0.0226*
CLIL -Low	T0	T1	-4.13	0.0004*
	T1	T2	1.00	0.7503
	T2	T3	-1.12	0.6753
	T0	T3	-4.57	<.0001*
Control -High	T0	T1	-0.10	0.9997
	T1	T2	-0.47	0.9649
	T2	T3	-0.78	0.8622
	T0	T3	-1.28	0.5764
Control -Low	T0	T1	0.06	0.9999
	T1	T2	0.78	0.8650
	T2	T3	-0.52	0.9535
	T0	T3	0.29	0.9914

Table 6.44
Percentages of improvement in the percentages of noun ratios according to proficiency and group in the narrative task.

Proficiency	Group	I0_1	I1_2	I2_3	I0_3
Low	CLIL	-25.4%	7%	-6.5%	-25.4%
	Control	0.3%	6.6%	-4.2%	2.4%
High	CLIL	-5.8%	-11.7%	-0.3%	-16.9%
	Control	-0.6%	-3.5%	-5.4%	-9.4%

Figure 6.12
Mean percentages of the noun ratios in the narrative task. Interactions between instruction type and proficiency level.



6.1.1.3.2 Percentage of the verb ratio

6.1.1.3.2.1 Interview task

Table 6.45 displays the results of the intergroup comparisons of the percentage of verbs produced in proportion to the TNWE in the interview task. As shown in the table, no significant differences were found between the groups at any of the data collection times. However, as Figure 6.13 clearly indicates, the control group outperforms the CLIL group in this measure over the two years of the study. Additionally, Table 6.46 shows that the control group shows the highest percentage of improvement from T0 to T3.

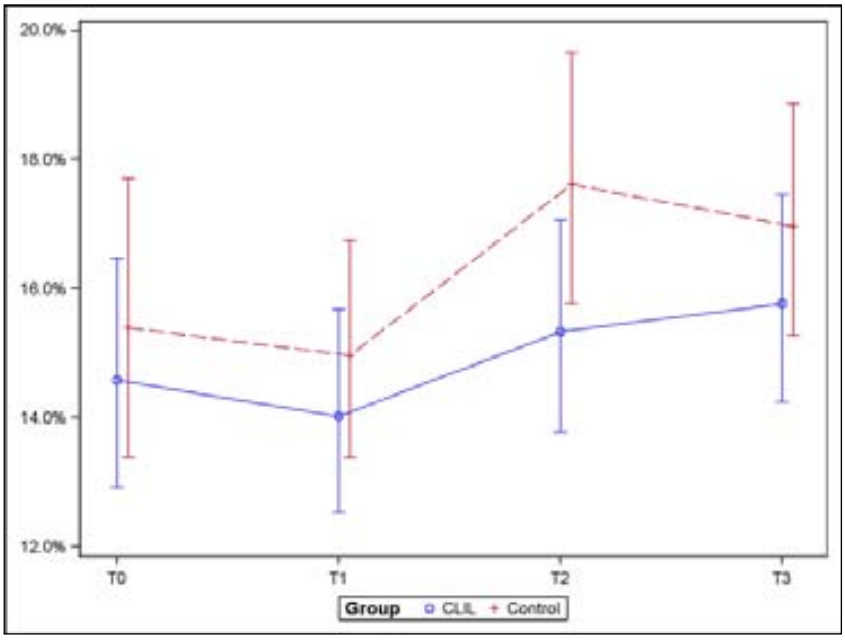
Table 6.45
Intergroup comparisons of the percentages of the verb ratios in the interview task.

	Group	Mean	StdErr	Lower	Upper	F value	p-value
T0	CLIL	14.6%	0.9%	12.9%	16.5%	0.32	0.5727
	Control	15.4%	1.1%	13.4%	17.7%		
T1	CLIL	14.0%	0.8%	12.5%	15.7%	0.68	0.4108
	Control	15.0%	0.9%	13.4%	16.8%		
T2	CLIL	15.3%	0.8%	13.8%	17.1%	3.37	0.0686
	Control	17.6%	1.0%	15.8%	19.7%		
T3	CLIL	15.8%	0.8%	14.2%	17.5%	1.02	0.3149
	Control	17.0%	0.9%	15.3%	18.9%		

Table 6.46
Percentages of improvement in the percentages of the verb ratios in the interview task.

	I0_1	I1_2	I2_3	I0_3
CLIL	-3.8%	9.3%	2.9%	8.1%
Control	-2.7%	17.7%	-3.7%	10.3%

Figure 6.13
Mean percentages of the verb ratios in the interview task.



Regarding the intragroup comparisons of CLIL and control learners for different time periods, the analyses shown in Table 6.47 reveal no significant improvement for either group at any of the time periods.

Table 6.47
Intragroup comparisons of the percentages of the verb ratios in the interview task.

	Time	Time	t Value	P-Value
CLIL	T0	T1	-0.53	0.9526
	T1	T2	1.30	0.5650
	T2	T3	0.44	0.9720
	T0	T3	1.11	0.6851
Control	T0	T1	-0.35	0.9856
	T1	T2	2.43	0.0769
	T2	T3	-0.58	0.9382
	T0	T3	1.25	0.5935

Tables 6.48 and 6.49 display the results obtained from the interactions between proficiency level and instruction type in the percentages of the verb ratios in the interview task. As seen from the tables, no significant differences were found between high and low achievers from the two groups. Likewise, the intragroup comparisons of CLIL high achievers, CLIL low achievers, control high achievers and control low achievers displayed in Table 6.50, show no significant improvement for any of the four groups of participants at any of the four time periods studied. Figure 6.14 shows the similarity is scores.

Table 6.48
Intergroup comparisons of the percentages of the verb ratios in the interview task. Interactions between high achievers and type of instruction.

	Proficiency	Group	Mean	StdErr	Lower	Upper	F Value	p-value
T0	High	CLIL	13.7%	1.1%	11.7%	16.1%	0.94	0.3340
	High	Control	15.4%	1.3%	13.0%	18.1%		
T1	High	CLIL	13.5%	1.0%	11.7%	15.6%	2.20	0.1404
	High	Control	15.8%	1.2%	13.6%	18.3%		
T2	High	CLIL	15.2%	1.0%	13.3%	17.4%	2.58	0.1108
	High	Control	17.8%	1.2%	15.5%	20.5%		
T3	High	CLIL	16.4%	1.0%	14.4%	18.6%	0.00	0.9671
	High	Control	16.4%	1.1%	14.3%	18.9%		

Table 6.49
Intergroup comparisons of the percentages of the verb ratios in the interview task. Interactions between low achievers and type of instruction.

	Proficiency	Group	Mean	StdErr	Lower	Upper	F Value	p-value
T0	Low	CLIL	15.5%	1.4%	13.1%	18.5%	0.00	0.3340
	Low	Control	15.4%	1.8%	12.2%	19.5%		
T1	Low	CLIL	14.6%	1.2%	12.4%	17.1%	0.05	0.8223
	Low	Control	14.2%	1.4%	11.7%	17.2%		
T2	Low	CLIL	15.5%	1.2%	13.4%	17.9%	1.21	0.2728
	Low	Control	17.4%	1.6%	14.6%	20.8%		
T3	Low	CLIL	15.2%	1.1%	13.2%	17.5%	2.02	0.1572
	Low	Control	17.5%	1.5%	14.8%	20.8%		

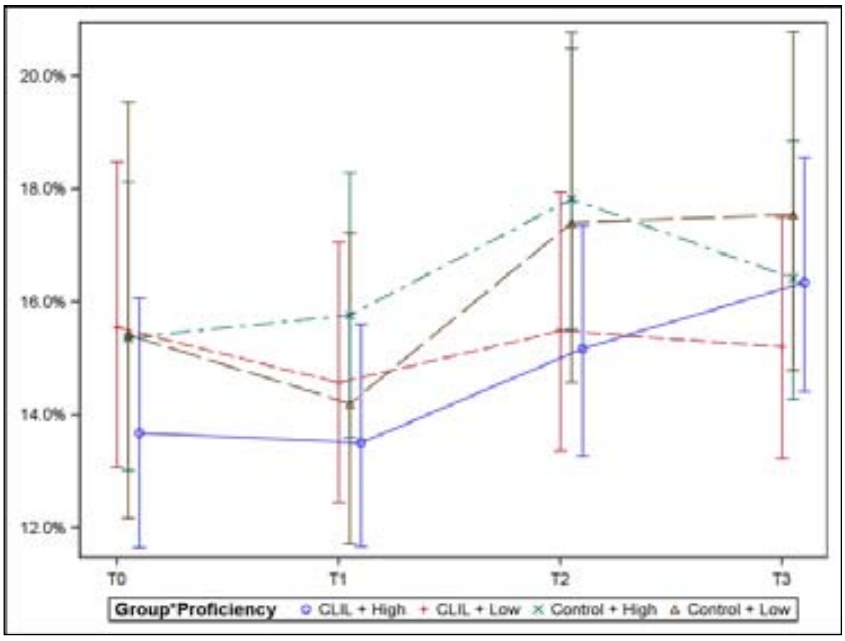
Table 6.50
Intragroup comparisons of the percentages of the verb ratios in the interview task. Interactions between instruction type and proficiency level.

	Time	Time	t Value	p-value
CLIL -High	T0	T1	-0.14	0.9991
	T1	T2	1.35	0.5353
	T2	T3	0.94	0.7848
	T0	T3	1.96	0.2097
CLIL -Low	T0	T1	-0.61	0.9294
	T1	T2	0.64	0.9200
	T2	T3	-0.20	0.9971
	T0	T3	-0.22	0.9965
Control -High	T0	T1	0.27	0.9930
	T1	T2	1.43	0.4812
	T2	T3	-1.01	0.7435
	T0	T3	0.71	0.8930
Control -Low	T0	T1	-0.59	0.9358
	T1	T2	1.77	0.2900
	T2	T3	0.07	0.9999
	T0	T3	0.98	0.7633

Table 6.51
Percentages of improvement in the percentages of the verb rations according to proficiency and group in the interview task.

Proficiency	Group	I0_1	I1_2	I2_3	I0_3
Low	CLIL	-5.8%	6.1%	-1.9%	-1.9%
	Control	-7.7%	22.55	0.5%	13.6%
High	CLIL	3.6%	12.5%	7.8%	19.7%
	Control	2.5%	12.6%	-7.8%	6.4%

Figure 6.14
Mean percentages of the verb ratios in the interview task. Interactions between instruction type and proficiency level.



6.1.1.3.2.2 Narrative task

Table 6.52 displays the results obtained from the intergroup comparisons used to analyse the effects of CLIL on the proportion of verbs produced by the learners in relation to the TNWE in the narrative task. As shown in the table, no differences were found between the CLIL and control groups at any of the four data collection times. As seen from Figure 6.15, both groups produce a similar number of verbs. The control group improves the most from T1 to T3, while the CLIL group starts to improve at T2, after a slight decrease. As for the degree of improvement, the control group shows the highest percentage from T0 to T3 (see Table 6.53).

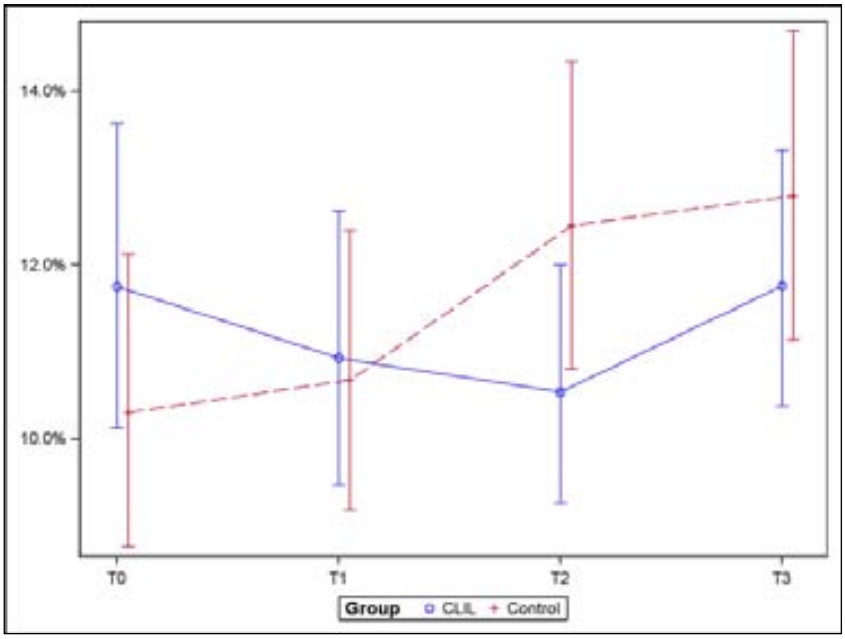
Table 6.52
Intergroup comparisons of the percentages of the verb ratios in the narrative task.

	Group	Mean	StdErr	Lower	Upper	F value	p-value
T0	CLIL	11.8%	0.9%	10.1%	13.6%	1.44	0.2320
	Control	10.3%	0.8%	8.8%	12.1%		
T1	CLIL	10.9%	0.8%	9.5%	12.6%	0.05	0.8194
	Control	10.7%	0.8%	9.2%	12.4%		
T2	CLIL	10.5%	0.7%	9.3%	12.0%	3.15	0.0787
	Control	12.5%	0.9%	10.8%	14.3%		
T3	CLIL	11.8%	0.7%	10.4%	13.3%	0.83	0.3653
	Control	12.8%	0.9%	11.1%	14.7%		

Table 6.53
Percentages of improvement in the percentages of the verb ratios in the narrative task.

	I0_1	I1_2	I2_3	I0_3
CLIL	-7.0%	-3.5%	11.5%	0.1%
Control	3.6%	16.6%	2.8%	24.1%

Figure 6.15
Mean percentages of the verb ratios in the narrative task.



Regarding the intragroup comparisons (see Table 6.54), no significant improvement was found for either of the groups at any of the time periods examined. However, the control group seems to improve the most from T0 to T3, even though the results do not reach significance (p= 0.0954). The CLIL group, on the other hand, does not undergo any improvement at all, since the mean scores obtained at T0 and at T3 are the same (see table 6.52).

Table 6.54
Intragroup comparisons of the percentages of the verb ratios in the narrative task.

	Time	Time	t Value	P-Value
CLIL	T0	T1	-0.79	0.8603
	T1	T2	-0.43	0.9732
	T2	T3	1.44	0.4785
	T0	T3	0.01	1.0000
Control	T0	T1	0.36	0.9839
	T1	T2	1.76	0.2992
	T2	T3	0.33	0.9874
	T0	T3	2.34	0.0954

As for the interactions between instruction type and proficiency level, Table 6.55 displays the results of the analyses for high achievers. Again, no significant differences were found between high achievers from the CLIL and control groups at any of the times analysed. The analyses for low achievers yield similar results (see Table 6.56), since no differences were found between CLIL low achievers and control low achievers.

Table 6.55
Intergroup comparisons of the percentages of verb ratios in the narrative task. Interactions between high achievers and type of instruction.

	Proficiency	Group	Mean	StdErr	Lower	Upper	F Value	p-value
T0	High	CLIL	12.2%	1.1%	10.2%	14.7%	0.20	0.6584
	High	Control	11.4%	1.3%	9.2%	14.2%		
T1	High	CLIL	11.2%	1.0%	9.3%	13.4%	0.01	0.9411
	High	Control	11.1%	1.2%	9.0%	13.6%		
T2	High	CLIL	11.1%	0.9%	9.4%	13.0%	0.74	0.3923
	High	Control	12.4%	1.2%	10.2%	15.0%		
T3	High	CLIL	11.7%	0.9%	9.9%	13.7%	1.63	0.2046
	High	Control	13.7%	1.2%	11.4%	16.4%		

Table 6.56
Intergroup comparisons of the percentages of verb ratios in the narrative task. Interactions between low achievers and type of instruction.

	Proficiency	Group	Mean	StdErr	Lower	Upper	F Value	p-value
T0	Low	CLIL	11.3%	1.2%	9.2%	13.9%	16.7	0.1987
	Low	Control	9.3%	1.2%	7.2%	12.0%		
T1	Low	CLIL	10.7%	1.1%	8.7%	13.1%	0.07	0.7987
	Low	Control	10.3%	1.2%	8.1%	13.1%		
T2	Low	CLIL	10.0%	1.0%	8.3%	12.1%	2.84	0.0945
	Low	Control	12.5%	1.4%	10.1%	15.6%		
T3	Low	CLIL	11.9%	1.0%	10.0%	14.1%	0.00	0.9523
	Low	Control	12.0%	1.4%	9.5%	15.0%		

As for the intragroup comparisons of CLIL high achievers, CLIL low achievers, control high achievers and control low achievers, no significant improvement was found in terms of the percentage of verbs produced during the narrative task for any of the time periods studied (see Table 6.57). As seen from Figure 6.16, the four groups of participants produce fairly similar numbers of verbs at T0 and T3. However, both

control high and low achievers display the highest percentages of improvement (see Table 6.58)

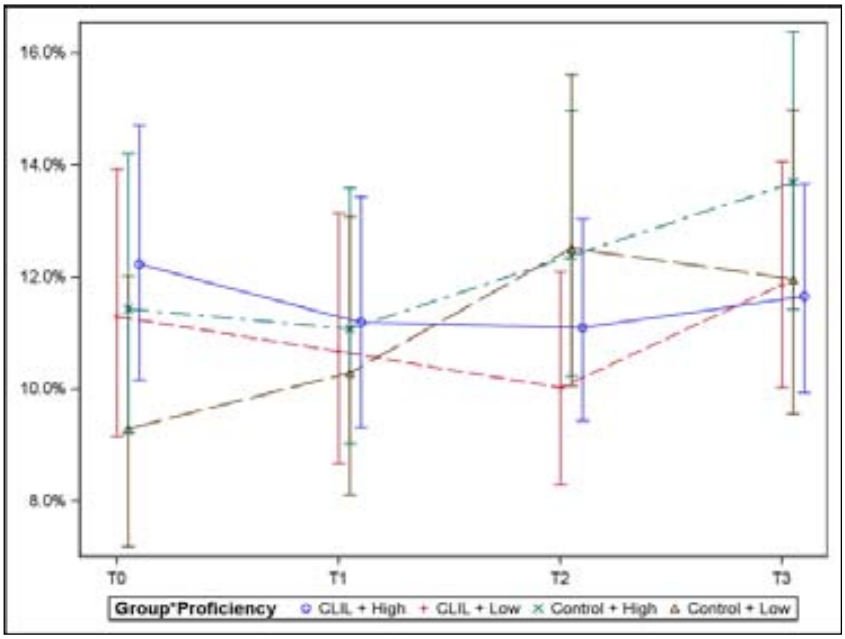
Table 6.57
Intragroup comparisons of the percentages of the verb ratios in the narrative task. Interactions between instruction type and proficiency level.

	Time	Time	t Value	p-value
CLIL -High	T0	T1	-0.77	0.8679
	T1	T2	-0.08	0.9998
	T2	T3	0.51	0.9555
	T0	T3	-0.46	0.9683
CLIL -Low	T0	T1	-0.42	0.9747
	T1	T2	-0.52	0.9553
	T2	T3	1.57	0.4007
	T0	T3	0.42	0.9746
Control -High	T0	T1	-0.25	0.9945
	T1	T2	0.94	0.7857
	T2	T3	0.94	0.7853
	T0	T3	1.49	0.4445
Control -Low	T0	T1	0.66	0.9133
	T1	T2	1.40	0.5016
	T2	T3	-0.34	0.9861
	T0	T3	1.68	0.3410

Table 6.58
Percentages of improvement in the percentages of the verb ratios according to proficiency and group in the narrative task.

Proficiency	Group	I0_1	I1_2	I2_3	I0_3
Low	CLIL	-5.3%	-6.5%	19%	5.3%
	Control	-2.6%	11.7%	10.4%	20.1%
High	CLIL	-8.1%	-0.8%	5.4%	-4%
	Control	10.7%	21.3%	-4%	29%

Figure 6.16
Mean percentages of the verb ratios in the interview task. Interactions between instruction type and proficiency level.



6.1.1.3.3 Percentage of the adjective ratio

6.1.1.3.3.1 Interview task

As for the third measure of lexical complexity, the adjective ratio, Table 6.59 displays the results of the intergroup comparisons in the interview task. As shown in the table, no significant differences were found between the CLIL and control groups at any of the data collection times. Furthermore, the mean percentages displayed in the table as well as the evolution of the two groups (see Figure 6.17) indicate that the percentages of adjectives produced by the two groups of learners were very similar. However, as Table 6.60 shows, CLIL learners' percentage of improvement in this measure from T0 to T3 is higher than that obtained by control learners, who show a decrease.

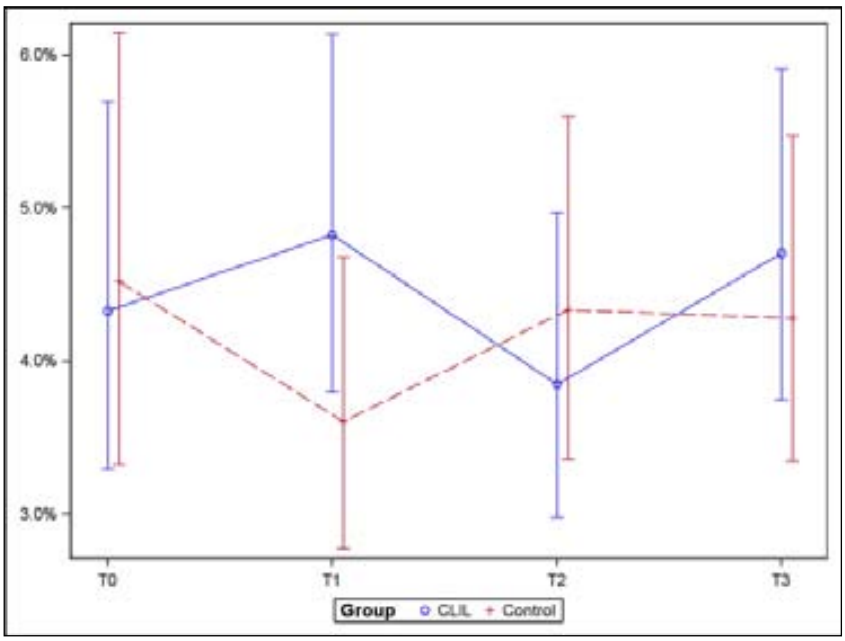
Table 6.59
Intergroup comparisons of the percentages of the adjective ratios in the interview task

	Group	Mean	StdErr	Lower	Upper	F value	p-value
T0	CLIL	4.3%	0.6%	3.3%	5.7%	0.04	0.8329
	Control	4.5%	0.7%	3.3%	6.1%		
T1	CLIL	4.8%	0.6%	3.8%	6.1%	2.61	0.1083
	Control	3.6%	0.5%	2.8%	4.7%		
T2	CLIL	3.8%	0.5%	3.0%	5.0%	0.46	0.4982
	Control	4.3%	0.6%	3.4%	5.6%		
T3	CLIL	4.7%	0.5%	3.8%	5.9%	0.32	0.5705
	Control	4.3%	0.5%	3.3%	5.5%		

Table 6.60
Percentages of improvement in the percentages of the adjective ratios in the interview task.

	I0_1	I1_2	I2_3	I0_3
CLIL	11.5%	-20.3%	22.3%	8.7%
Control	-20.3%	20.3%	-1.3%	-5.3%

Figure 6.17
Mean percentages of the adjective ratios in the interview task.



Concerning the intragroup comparisons of the learners' performance on this measure, the results displayed in Table 6.61 indicate that there is no significant improvement for any of the time periods in either group.

Table 6.61
Intragroup comparisons of the percentages of the adjective ratios in the interview task.

	Time	Time	t Value	P-Value
CLIL	T0	T1	0.65	0.9168
	T1	T2	-1.40	0.5016
	T2	T3	1.28	0.5764
	T0	T3	0.51	0.9569
Control	T0	T1	-1.22	0.6130
	T1	T2	1.12	0.6763
	T2	T3	-0.08	0.9998
	T0	T3	-0.30	0.9903

Regarding the analyses conducted to examine the interaction between proficiency level and instruction type, the results indicate that there are no significant differences between high achievers from the CLIL and control group (Table 6.62) and low achievers from the CLIL and control group (Table 6.63).

Table 6.62
Intergroup comparisons of the percentages of the adjective ratios in the interview task. Interactions between high achievers and type of instruction.

	Proficiency	Group	Mean	StdErr	Lower	Upper	F Value	p-value
T0	High	CLIL	4.4%	0.8%	3.1%	6.2%	0.22	0.6384
	High	Control	5.0%	0.9%	3.5%	7.1%		
T1	High	CLIL	4.9%	0.8%	3.6%	6.7%	2.96	0.0877
	High	Control	3.2%	0.6%	2.3%	4.7%		
T2	High	CLIL	4.1%	0.7%	3.0%	5.7%	0.07	0.7896
	High	Control	4.4%	0.7%	3.2%	6.1%		
T3	High	CLIL	4.3%	0.7%	3.2%	5.8%	0.36	0.5477
	High	Control	4.9%	0.7%	3.7%	6.7%		

Table 6.63
Intergroup comparisons of the percentages of the adjective ratios in the interview task. Interactions between low achievers and type of instruction.

	Proficiency	Group	Mean	StdErr	Lower	Upper	F Value	p-value
T0	Low	CLIL	4.2%	0.8%	2.9%	6.2%	0.01	0.9210
	Low	Control	4.1%	1.1%	2.4%	6.9%		
T1	Low	CLIL	4.8%	0.8%	3.3%	6.8%	0.44	0.5086
	Low	Control	4.0%	0.9%	2.6%	6.2%		
T2	Low	CLIL	3.6%	0.6%	2.5%	5.1%	0.51	0.4774
	Low	Control	4.3%	0.9%	2.8%	6.4%		
T3	Low	CLIL	5.1%	0.8%	3.8%	6.9%	1.93	0.1672
	Low	Control	3.7%	0.8%	2.5%	5.6%		

As for the intragroup analyses that explored the improvement of CLIL high and low achievers and control high and low achievers between different time periods, Table 6.64 shows that none of the four groups improves significantly in the percentage of adjectives used in the interview task between any of the time periods examined. As Figure 6.18 shows, the four groups obtain approximately the same percentages at T0

and T3. Surprisingly, the group which obtains the highest percentage of improvement from T0 to T3 in this measure is CLIL low achievers (see Table 6.65)

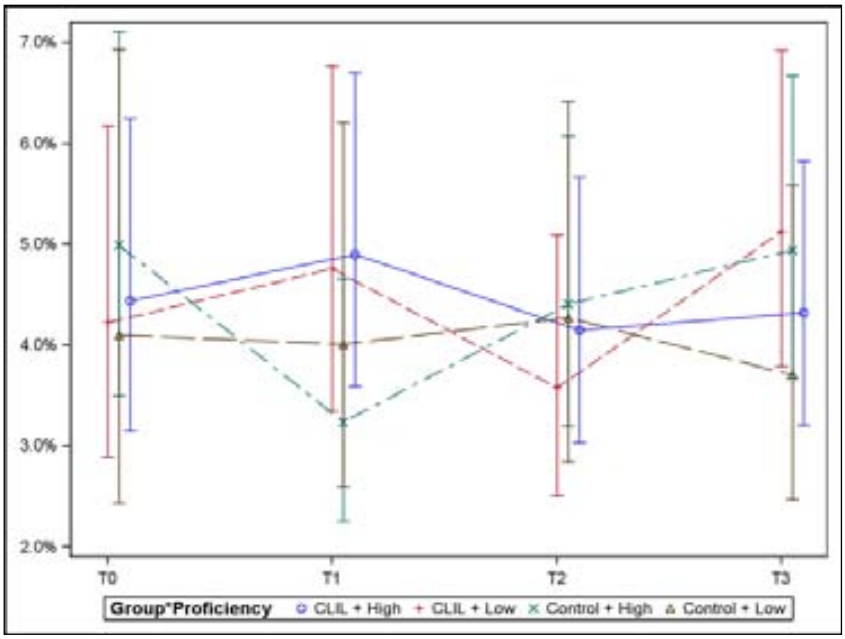
Table 6.64
Intragroup comparisons of the percentages of the adjective ratios in the interview task. Interactions between instruction type and proficiency level.

	Time	Time	t Value	p-value
CLIL -High	T0	T1	0.47	0.9661
	T1	T2	-0.83	0.8415
	T2	T3	0.22	0.9965
	T0	T3	-0.12	0.9993
CLIL -Low	T0	T1	0.49	0.9620
	T1	T2	-1.23	0.6100
	T2	T3	1.69	0.3331
	T0	T3	0.86	0.8255
Control -High	T0	T1	-1.87	0.2454
	T1	T2	1.41	0.4962
	T2	T3	0.60	0.9323
	T0	T3	-0.04	1.0000
Control -Low	T0	T1	-0.07	0.9999
	T1	T2	0.23	0.9959
	T2	T3	-0.54	0.9501
	T0	T3	-0.32	0.9883

Table 6.65
Percentages of improvement in the percentages of the adjectives ratios according to proficiency and group in the interview task.

Proficiency	Group	I0_1	I1_2	I2_3	I0_3
Low	CLIL	14.2%	-25%	41.66%	21.4%
	Control	-2.4%	7.5%	-13.9%	-9.7%
High	CLIL	11.3%	-16.3%	4.8%	-2.2%
	Control	-36%	37.5%	11.3%	-2%

Figure 6.18
Mean percentages of the adjective ratios in the interview task. Interactions between instruction type and proficiency level.



6.1.1.3.3.2 Narrative task

The percentages of the adjective ratios used in the narrative task could not be analysed in the statistical analysis due to the limited number of adjectives produced by the learners. See Table 6.66 for the number of participants from the CLIL (N=32) and control groups (N=20) who produced adjectives.

Table 6.66
Number of participants that produced adjectives in the narrative task.

Narrative task	T0	T1	T2	T3
CLIL group	6 (32)	4 (32)	7 (32)	5 (32)
Control group	4 (20)	2 (20)	4 (20)	6 (20)

6.1.2 Summary complexity results

In this section, the most relevant results concerning the measures used to analyse propositional, syntactic and lexical complexity will be highlighted and summarised. First, the intergroup analyses between CLIL and non-CLIL learners will be provided for each of the measures. After that, the intragroup comparisons which analyse the learners' improvement from T0 to T3 will be presented. Finally, the interactions between instruction type and proficiency level will be provided.

The intergroup analyses conducted to analyse the effects of CLIL instruction on the learners performance on propositional and syntactic complexity measures reveal that (see Table 6.67), even though there are no differences between the CLIL and control groups at T0, the onset of this investigation, the control group performs significantly better in the TNU at T1, in the % CU at T1 and T2 in the interview task and in the TNU at T1 in the narrative task. As seen from the table, differences between the groups disappear either at T2 or T3.

Table 6.67
Propositional and syntactic complexity. Summary of the significance values obtained in the intergroup comparisons of CLIL and control groups in the interview and narrative task.

	INTERVIEW				NARRATIVE			
	T0	**T1**	**T2**	**T3**	**T0**	**T1**	**T2**	**T3**
TNU		p=0.004* Control				p=0.0078* Control		
% CU		P=0.0318* Control	P=0.0240* Control		Unable to conduct statistical analyses			
% SU	Unable to conduct statistical analyses							

Regarding lexical complexity, the only difference found is in the percentages of the noun ratios. As table 6.68 shows, the CLIL group produces a significantly higher

number of nouns in proportion to TNWE than the control group at T0, T1, T2 and T3 in the interview task and at T0 and T1 in the narrative task. As for the other ratios, no differences were found between the CLIL and control groups.

Table 6.68
Lexical complexity. Summary of the significance values obtained in the intergroup comparisons of CLIL and control groups in the interview and narrative task.

	INTERVIEW				NARRATIVE			
	T0	**T1**	**T2**	**T3**	**T0**	**T1**	**T2**	**T3**
% NR	p=0.0043* CLIL	p=0.0287* CLIL	p=0.0071* CLIL	p<.0001* CLIL	p<.0001* CLIL	p=0.0479* CLIL		
% VR								
% ADJR					Unable to conduct statistical analyses			

Concerning the intragroup comparisons in the propositional and syntactic complexity measures, Table 6.69 shows that both CLIL and control learners display significant development from T0 to T3 in the TNU and in the % CU in the interview task. As for the narrative task, the only significant improvement was found in the TNU for both the CLIL and control group.

Table 6.69
Propositional and syntactic complexity. Summary of the significance values obtained in the intragroup comparisons (from T0 to T3) of CLIL and control groups in the interview and narrative task.

	INTERVIEW		NARRATIVE	
	CLIL	**CONTROL**	**CLIL**	**CONTROL**
TNU	p<.0001*	p<.0001*	p<.0001*	p=0.0065*
% CU	p=0.0002*	p=0.0297*		
% SU	Unable to conduct statistical analyses			

In regard to lexical complexity, the results of the intragroup comparisons (see Table 6.70) show that the only significant development from T0 to T3, in this case a significant decrease, was found in the % NR in the narrative task in favour of the CLIL group. No other significant development was found for any of the measures used to analyse lexical complexity.

Table 6.70
Lexical complexity. Summary of the significance values obtained in the intragroup comparisons (from T0 to T3) of CLIL and control groups in the interview and narrative task.

	INTERVIEW		**NARRATIVE**	
	CLIL	**CONTROL**	**CLIL**	**CONTROL**
% NR			p=0.0001* (decrease)	
% LVR				
% ADJR			Unable to conduct statistical analyses	

Table 6.71 shows the significant differences found in the propositional and syntactic complexity measures between CLIL and control high achievers. The results displayed in table 6.71 indicate the control high achievers significantly outperform CLIL high achievers in the TNU at T0, T1 and T2 in the interview task. As reported previously, however, these differences disappear at T3. In the narrative task, control high achievers also obtain significant results in the TNU at T1 and in the % CU at T3. As for CLIL and control low achievers, the intergroup comparisons shown in Table 6.72 only indicate a significant difference in favour of the control group in the % CU in the interview task at T2.

Table 6.71
Propositional and syntactic complexity. Summary of the significance values in the intergroup comparisons of CLIL and control high achievers in the interview and narrative task.

	INTERVIEW				NARRATIVE			
	T0	**T1**	**T2**	**T3**	**T0**	**T1**	**T2**	**T3**
TNU	p=0.0113* Control-high	p=0.0154* Control-high	p=0.0057* Control high			p=0.0265* Control-high		
% CU					Unable to conduct statistical analyses			p=0.0232* Control-high
% SU	Unable to conduct statistical analyses							

Table 6.72
Propositional and syntactic complexity. Summary of the significance values in the intergroup comparisons of CLIL and control low achievers in the interview and narrative task.

	INTERVIEW				NARRATIVE			
	T0	**T1**	**T2**	**T3**	**T0**	**T1**	**T2**	**T3**
TNU								
% CU			p=0.0291* Control-low		Unable to conduct statistical analyses			
% SU	Unable to conduct statistical analyses							

In regard to the interactions between proficiency level and instruction type in the lexical complexity measures, Table 6.73 shows that the % NR is significantly higher for CLIL high achievers than control high achievers in the interview task at the four times of data collection and at T0 in the narrative task. The results in Table 6.74 also indicate that CLIL low achievers obtain significantly higher scores than control low achievers in the % NR at T3 in the interview task and at T0 in the narrative task.

Table 6.73
Lexical complexity. Summary of the significance values in the intergroup comparisons of CLIL and control high achievers in the interview and narrative task.

	INTERVIEW				NARRATIVE			
	T0	**T1**	**T2**	**T3**	**T0**	**T1**	**T2**	**T3**
% NR	p=0.0005* CLIL- high	p=0.0054* CLIL- high	p=0.0018* CLIL- high	p=0.0003* CLIL- high	p=0.0274* CLIL- high			
% VR								
% ADJR					Unable to conduct statistical analyses			

Table 6.74
Lexical complexity. Summary of the significance values in the intergroup comparisons of CLIL and control low achievers in the interview and narrative task.

	INTERVIEW				NARRATIVE			
	T0	**T1**	**T2**	**T3**	**T0**	**T1**	**T2**	**T3**
% NR				p=0.0310* CLIL- low	p<.0001* CLIL-low			
% LVR								
% ADJR					Unable to conduct statistical analyses			

Additionally, the intragroup comparisons displayed in Table 6.75 indicate that CLIL high achievers improve significantly from T0 to T3 in the TNU and in the % CU in the interview task while control high achievers only improved significantly in the TNU. As for the narrative task, both CLIL and control high achievers improve significantly in the TNU. Likewise, the results in Table 6.76 show that CLIL and control low achievers display significant development in the same areas as their peers in the high proficiency group in the interview task. In the narrative task, only CLIL low achievers undergo significant improvement in the TNU from T0 to T3.

Table 6.75
Propositional and syntactic complexity. Summary of the significance values of the intragroup comparisons (from T0 to T3) of CLIL and control high achievers in the interview and narrative task.

	INTERVIEW		NARRATIVE	
	CLIL-high	**CONTROL-high**	**CLIL-high**	**CONTROL-high**
TNU	p<.0001*	p=0.0021*	p=0.0040*	p=0.0015*
% CU	p=0.0016*			
% SU	Unable to conduct statistical analyses			

Table 6.76
Propositional and syntactic complexity. Summary of the significance values of the intragroup comparisons (from T0 to T3) of CLIL and control low achievers in the interview and narrative task.

	INTERVIEW		NARRATIVE	
	CLIL-low	**CONTROL-low**	**CLIL-low**	**CONTROL-low**
TNU	p=0.0032*	p=0.0192*	p=0.0016*	
% CU	p=0.0339*			
% SU	Unable to conduct statistical analyses			

Finally, the intragroup comparisons in Tables 6.77 and 6.78 show the instances of significant development of CLIL high/low achievers and control high/low achievers. As seen form Table 6.77, CLIL high achievers display a significant decrease in the % NR from T0 to T3 in the two tasks. CLIL low achievers, on the other hand, only show a significant decrease in the % NR in the narrative task (see Table 6.78). Control high and low achievers do not show any instances of significant development in any of the measures of lexical complexity.

Table 6.77
Lexical complexity. Summary of the significance values of the intragroup comparisons (from T0 to T3) of CLIL and control high achievers in the interview and narrative task.

	INTERVIEW		NARRATIVE	
	CLIL-high	**CONTROL-high**	**CLIL-high**	**CONTROL-high**
% NR	p=0.0103* (decrease)		p=0.0004* (decrease)	
% VR				
% ADJR			Unable to conduct statistical analyses	

Table 6.78
Lexical complexity. Summary of the significance values of the intragroup comparisons (from T0 to T3) of CLIL and control low achievers in the interview and narrative task.

	INTERVIEW		NARRATIVE	
	CLIL-low	**CONTROL-low**	**CLIL-low**	**CONTROL-low**
% NR			p<.0001* (decrease)	
% VR				
% ADJR			Unable to conduct statistical analyses	

6.1.3 Accuracy

This section presents the accuracy results obtained by CLIL and control learners in the interview and narrative tasks. As mentioned in the previous chapter, two measures were selected for the analysis of accuracy: the percentage of error-free units (% EFU) and the percentage of correct verb forms (% CV). The % EFU was chosen as a global accuracy measure while specific accuracy was measured by the % CV.

6.1.3.1 Global accuracy: percentage of error-free units

6.1.3.1.1 Interview task

Table 6.79 displays the percentages and significance values of the intergroup comparisons of the % EFU between CLIL and control leaners in the interview task. As can be seen from the table, there are no significant differences between the groups since both groups display similar percentages of EFU over the two years of the study. In addition, as Figure 6.19 shows, both groups experience a decrease in the number of error-free units produced. According to the percentages of improvement (decrease in this case) displayed in Table 6.80, the CLIL group has the highest percentages of decrease from T1 to T2 and from T2 to T3, when compared to the control group.

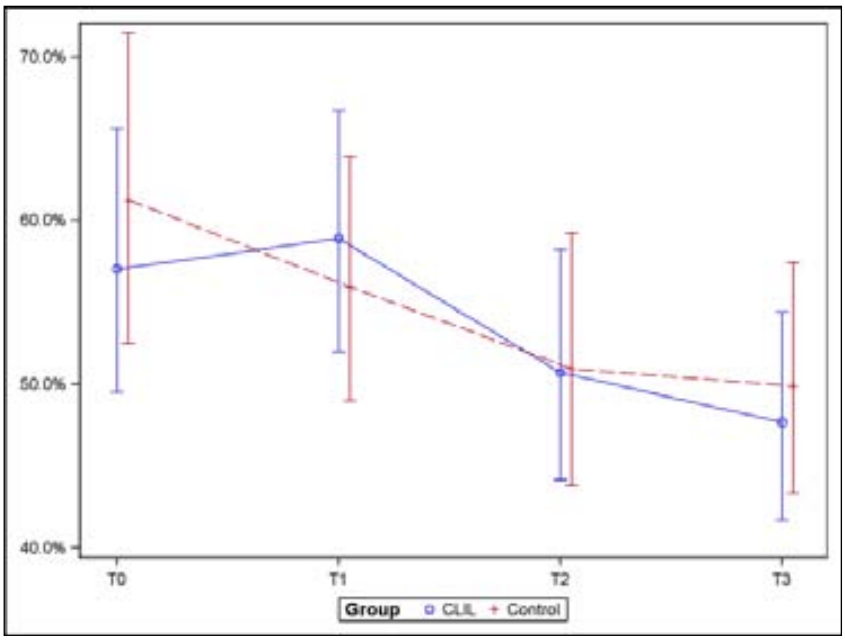
Table 6.79
Intergroup comparisons of the percentages of error-free units in the interview task.

	Group	Mean	StdErr	Lower	Upper	F value	p-value
T0	CLIL	57.0%	4.1%	49.5%	65.6%	0.48	0.4903
	Control	61.3%	4.8%	52.5%	71.5%		
T1	CLIL	58.9%	3.7%	52.0%	66.8%	0.32	0.5736
	Control	56.0%	3.8%	49.0%	63.9%		
T2	CLIL	50.7%	3.6%	44.2%	58.2%	0.00	0.9622
	Control	51.0%	3.9%	43.8%	59.2%		
T3	CLIL	47.7%	3.2%	41.7%	54.4%	0.24	0.6284
	Control	49.9%	3.6%	43.3%	57.4%		

Table 6.80
Percentages of improvement in the percentages of error-free units in the interview task.

	I0_1	I1_2	I2_3	I0_3
CLIL	3.3%	-13.9%	-6.0%	-16.4%
Control	-8.7%	-8.9%	-2.1%	-18.6%

Figure 6.19
Mean percentages of error-free units in the interview task.



Regarding the evolution of both groups, Table 6.81 shows the results from the intragroup comparisons of the % EFU. As seen from the table, no significant development was found in this measure for any of the time periods studied.

Table 6.81
Intragroup comparisons of the percentages of error-free units in the interview task.

	Time	Time	t Value	P-Value
CLIL	T0	T1	0.38	0.9809
	T1	T2	-1.77	0.2937
	T2	T3	-0.71	0.8941
	T0	T3	-2.02	0.1874
Control	T0	T1	-1.00	0.7521
	T1	T2	-1.04	0.7274
	T2	T3	-0.23	0.9959
	T0	T3	-2.17	0.1366

As for the interactions between type of instruction and proficiency level, Tables 6.82 and 6.83 show no significant impact of type of instruction according to the learners' proficiency level on the % EFU in the interview task. Both high and low achievers have quite similar percentages of error-free units regardless of the type of instruction received.

Table 6.82
Intergroup comparisons of the percentages of error-free units in the interview task. Interactions between high achievers and type of instruction.

	Proficiency	Group	Mean	StdErr	Lower	Upper	F Value	p-value
T0	High	CLIL	58.9%	5.4%	49.3%	70.5%	0.18	0.6739
	High	Control	62.4%	5.7%	52.0%	74.8%		
T1	High	CLIL	57.9%	4.6%	49.5%	67.9%	0.58	0.4485
	High	Control	52.6%	5.0%	43.5%	63.5%		
T2	High	CLIL	50.6%	4.5%	42.4%	60.4%	0.29	0.5938
	High	Control	47.1%	4.5%	39.0%	56.9%		
T3	High	CLIL	43.6%	3.9%	36.5%	52.0%	2.76	0.0990
	High	Control	54.0%	4.7%	45.5%	64.2%		

Table 6.83
Intergroup comparisons of the percentages of error-free units in the interview task. Interactions between low achievers and type of instruction.

	Proficiency	Group	Mean	StdErr	Lower	Upper	F Value	p-value
T0	Low	CLIL	55.2%	5.5%	45.3%	67.1%	0.34	0.5610
	Low	Control	60.2%	7.8%	46.6%	77.8%		
T1	Low	CLIL	59.9%	5.3%	50.3%	71.3%	0.00	0.9631
	Low	Control	59.6%	6.4%	48.2%	73.6%		
T2	Low	CLIL	50.9%	4.7%	42.3%	61.1%	0.33	0.5646
	Low	Control	55.2%	6.7%	43.4%	70.1%		
T3	Low	CLIL	52.1%	4.6%	43.8%	62.1%	0.90	0.3439
	Low	Control	46.1%	5.3%	36.7%	57.8%		

The intragroup comparisons in Table 6.84 indicate that the only a significant instance of development, in this case a significant decrease in the % EFU, was found for CLIL high achievers from T0 to T3. As Figure 6.20 displays, CLIL high achievers is the group which decreased the most in this measure (as Table 6.85 also shows). No other significant development was found in the other groups.

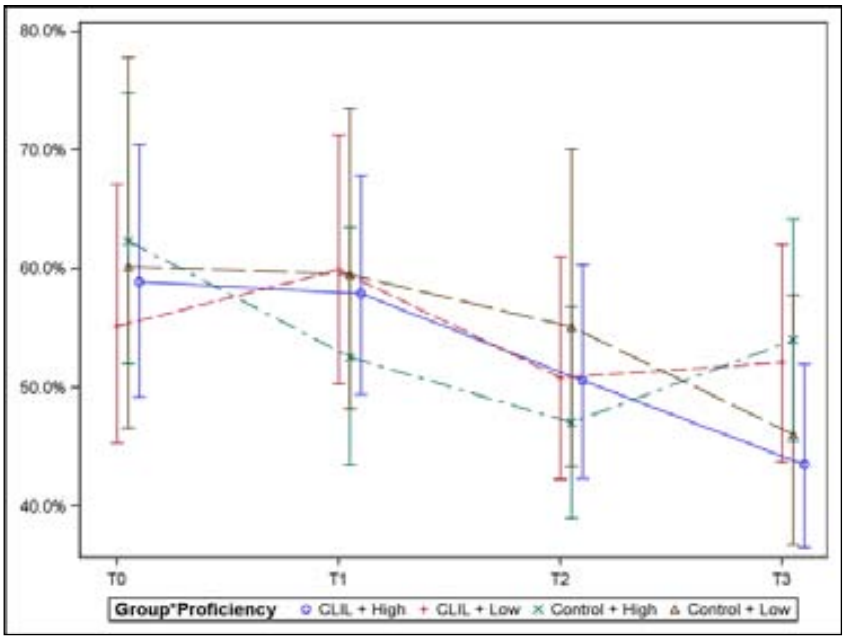
Table 6.84
Intragroup comparisons of the percentages of error-free units in the interview task. Interactions between instruction type and proficiency level.

	Time	Time	t Value	p-value
CLIL -High	T0	T1	-0.16	0.9987
	T1	T2	-1.26	0.5920
	T2	T3	-1.30	0.5659
	T0	T3	-2.60	0.0495*
CLIL -Low	T0	T1	0.69	0.9023
	T1	T2	-1.42	0.4863
	T2	T3	0.21	0.9965
	T0	T3	-0.47	0.9657
Control -High	T0	T1	-1.47	0.4565
	T1	T2	-0.92	0.7925
	T2	T3	1.22	0.6154
	T0	T3	-1.32	0.5546
Control -Low	T0	T1	-0.07	0.9999
	T1	T2	-0.53	0.9511
	T2	T3	-1.19	0.6346
	T0	T3	-1.69	0.3307

Table 6.85
Percentages of improvement in the percentages error-free units according to proficiency and group in the interview task.

Proficiency	Group	I0_1	I1_2	I2_3	I0_3
Low	CLIL	8.5%	-15%	2.3%	-5.6%
	Control	-0.9%	-7.3%	-16.4%	-23.4%
High	CLIL	-1.6%	-13.6%	-13.8%	-25.9%
	Control	-15.7%	-10.4%	14.6%	-13.4%

Figure 6.20
Mean percentages of error-free units in the interview task. Interactions between instruction type and proficiency level.



6.1.3.1.2 Narrative task

Table 6.86 displays the results of the intergroup comparisons between the CLIL and control groups in the % EFU in the narrative task. As seen from the table, no differences were found between the groups at any of the data collection times. The control group starts with a slight advantage at T0 in comparison to the CLIL group. However, at T1 and T3 the CLIL group has a higher percentage of error-free units. Furthermore, as can be seen from Figure 6.21, the CLIL group shows a more regular trend from T1 to T3 than the control group, which displays a higher variability of mean percentages. Despite that, as Table 6.87 shows, CLIL learners have the highest percentage of improvement from T0 to T3. In regard to the intragroup comparisons of CLIL and non-CLIL learners, Table 6.88 indicates that neither of the groups shows significant development at any of the time periods analysed.

Table 6.86
Intergroup comparisons of the percentages of error-free units in the narrative task.

	Group	Mean	StdErr	Lower	Upper	F value	p-value
T0	CLIL	7.6%	2.8%	3.7%	15.7%	0.40	0.5280
	Control	10.1%	3.5%	5.1%	20.1%		
T1	CLIL	15.1%	3.8%	9.1%	24.8%	2.20	0.1394
	Control	6.7%	3.2%	2.6%	17.4%		
T2	CLIL	11.4%	3.0%	6.8%	19.1%	1.31	0.2547
	Control	16.7%	4.0%	10.4%	26.9%		
T3	CLIL	14.1%	2.9%	9.4%	21.2%	0.75	0.3863
	Control	9.1%	4.3%	3.6%	22.9%		

Table 6.87
Percentages of improvement in the percentages of error-free units in the narrative task.

	I0_1	I1_2	I2_3	I0_3
CLIL	97.8%	-24.4%	23.8%	85.2%
Control	-33.9%	149.4%	-45.7%	-10.5%

Figure 6.21
Mean percentages of error-free units in the narrative task.

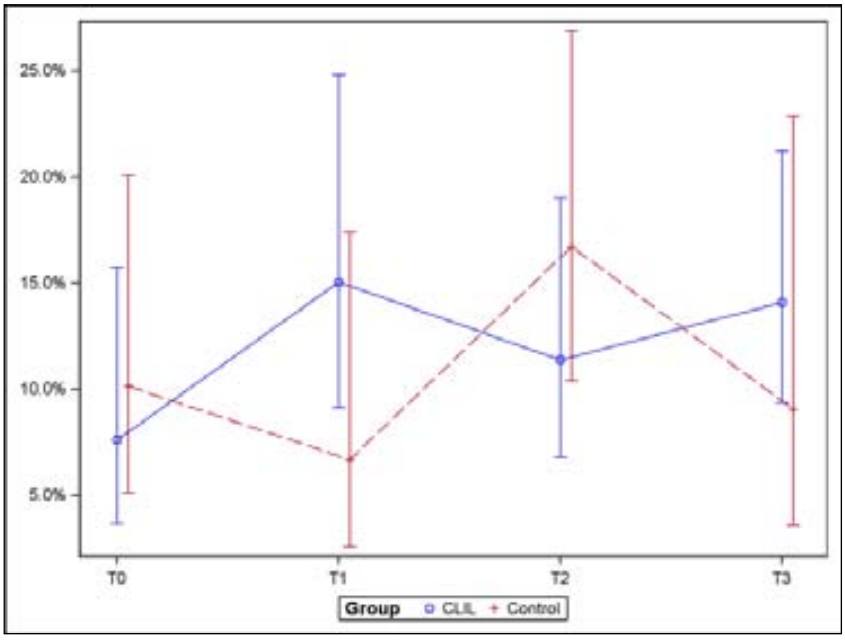


Table 6.88
Intragroup comparisons of the percentages of error-free units in the narrative task.

	Time	Time	t Value	P-Value
CLIL	T0	T1	1.55	0.4131
	T1	T2	-0.78	0.8621
	T2	T3	0.66	0.9132
	T0	T3	1.49	0.4489
Control	T0	T1	-0.70	0.8962
	T1	T2	1.71	0.3217
	T2	T3	-1.17	0.6448
	T0	T3	-0.19	0.9975

The results from the interactions between type of instruction and proficiency level are shown in Tables 6.89 and 6.90. Table 6.89 displays the results from the intergroup comparisons of CLIL high achievers and control high achievers at the four times of data collection. As seen from the table, the only significant difference between CLIL and control high achievers was found at T3 (F=9.81, p=0.0024) when the average % EFU of the control group is 34.5%, while the CLIL group has an average of 11.8%. Likewise, Table 6.90 shows significant differences between CLIL and control low achievers at T3 as well (F=4.24, p=0.0410). In this case, the mean % EFU is 16.9% for CLIL learners and 2.4% for non-CLIL learners. It is also important to note that significant differences were found between control high and low achievers in favour of high achievers at T3 (p=0.0058).

Table 6.89
Intergroup comparisons of the percentages of error-free units in the narrative task. Interactions between high achievers and type of instruction.

	Proficiency	Group	Mean	StdErr	Lower	Upper	F Value	p-value
T0	High	CLIL	8.6%	3.6%	3.8%	19.5%	0.38	0.390
	High	Control	12.1%	4.9%	5.4%	27.0%		
T1	High	CLIL	18.2%	5.5%	10.0%	33.1%	0.00	0.951
	High	Control	17.7%	5.2%	9.9%	31.6%		
T2	High	CLIL	13.8%	4.3%	7.4%	25.4%	0.02	0.8877
	High	Control	12.9%	4.3%	6.7%	24.9%		
T3	High	CLIL	11.8%	3.3%	6.8%	20.3%	9.81	0.0024*
	High	Control	34.5%	6.6%	23.4%	50.9%		

Table 6.90
Intergroup comparisons of the percentages of error-free units in the narrative task. Interactions between low achievers and type of instruction.

	Proficiency	Group	Mean	StdErr	Lower	Upper	F Value	p-value
T0	Low	CLIL	6.8%	3.5%	2.4%	18.7%	0.10	0.7465
	Low	Control	8.5%	5.0%	2.7%	27.4%		
T1	Low	CLIL	12.5%	5.1%	5.6%	27.8%	2.56	0.1114
	Low	Control	2.5%	2.4%	0.4%	16.2%		
T2	Low	CLIL	9.4%	3.6%	4.4%	20.0%	3.08	0.0812
	Low	Control	21.7%	7.7%	10.8%	43.8%		
T3	Low	CLIL	16.9%	4.5%	10.0%	28.6%	4.24	0.0410*
	Low	Control	2.4%	2.2%	0.4%	14.9%		

The intragroup comparisons displayed in Table 6.91 show that the only group which shows significant development in the % EFU in the narrative task is high achievers from the control group (F=2.66, p= 0.0425) from T2 to T3. The second group which shows the greatest improvement is CLIL low achievers from T0 to T3 (see Table 6.92). As can be seen from Figure 6.22, control high achievers is the group that improves the most from T2 to T3. Control low achievers, however, show the greatest decrease in the % EFU from T2 to T3. CLIL high and low achievers, on the other hand, show similar mean percentages at the four times of data collection.

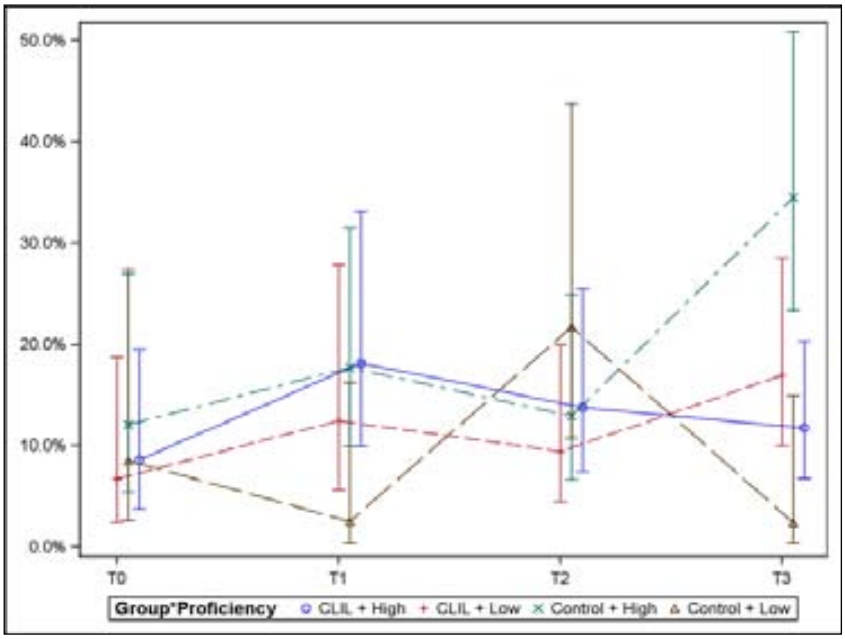
Table 6.91
Intragroup comparisons of the percentages of error-free units in the narrative task. Interactions between instruction type and proficiency level.

	Time	Time	t Value	p-value
CLIL -High	T0	T1	1.48	0.4520
	T1	T2	-0.65	0.9143
	T2	T3	-0.39	0.9802
	T0	T3	0.64	0.9183
CLIL -Low	T0	T1	0.94	0.7809
	T1	T2	-0.51	0.9564
	T2	T3	1.28	0.5746
	T0	T3	1.61	0.3776
Control -High	T0	T1	0.78	0.8627
	T1	T2	-0.73	0.8837
	T2	T3	2.66	0.0425*
	T0	T3	2.41	0.0803
Control -Low	T0	T1	-1.10	0.6914
	T1	T2	2.15	0.1411
	T2	T3	-2.24	0.1177
	T0	T3	-1.16	0.6505

Table 6.92
Percentages of improvement in the percentages error-free units according to proficiency and group in the narrative task.

Proficiency	Group	I0_1	I1_2	I2_3	I0_3
Low	CLIL	83.8%	-24.8%	79.7%	148.5%
	Control	-70.5%	768%	-88.9%	-71.7%
High	CLIL	111.6%	-24.1%	-14.4%	37.2%
	Control	46.2%	-27.1%	167.4%	185.1%

Figure 6.22
Mean percentages of error-free units in the narrative task. Interactions between instruction type and proficiency level



6.1.3.2 Specific accuracy: percentage of correct verb forms

6.1.3.2.1 Interview task

Table 6.93 displays the results of the intergroup comparisons between the CLIL and control groups in the percentage of correct verb forms produced in the interview task. As shown in the table, no significant differences were found between the groups at any of the data collection times. In addition, it is worth noting that both groups display similar trends in this measure inasmuch as both groups obtain similar mean percentages

at Time 0 and time 3 (see Figure 6.23). Despite that, as Table 6.94 displays, the CLIL group obtains a higher percentage of improvement from T0 to T3. As for the intragroup comparisons, Table 6.95 shows no statistically significant differences in development for either group at any of the time periods analysed.

Table 6.93
Intergroup comparisons of the percentages of correct verb forms in the interview task.

	Group	Mean	StdErr	Lower	Upper	F value	p-value
T0	CLIL	55.6%	4.4%	47.5%	65.1%	1.02	0.3128
	Control	62.6%	5.4%	52.8%	74.1%		
T1	CLIL	66.5%	4.5%	58.1%	76.1%	0.05	0.8237
	Control	65.1%	4.5%	56.8%	74.7%		
T2	CLIL	58.2%	4.1%	50.7%	66.8%	0.12	0.7332
	Control	60.2%	4.2%	52.4%	69.1%		
T3	CLIL	61.0%	4.0%	53.7%	69.4%	0.13	0.7231
	Control	59.1%	4.0%	51.7%	67.5%		

Table 6.94
Percentages of improvement in the percentages of correct verb forms in the interview task.

	I0_1	I1_2	I2_3	I0_3
CLIL	19.6%	-12.5%	4.8%	9.7%
Control	4.1%	-7.6%	-1.9%	-5.6%

Figure 6.23
Mean percentages of correct verb forms in the interview task.

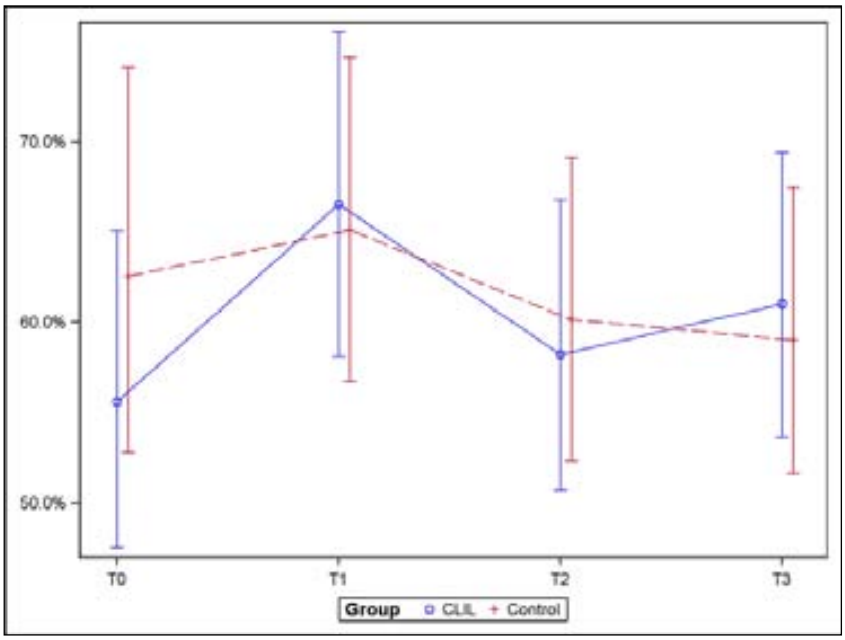


Table 6.95
Intragroup comparisons of the percentages of correct verb forms in the interview task.

	Time	Time	t Value	p-value
CLIL	T0	T1	1.90	0.2336
	T1	T2	-1.54	0.4175
	T2	T3	0.56	0.9439
	T0	T3	1.00	0.7479
Control	T0	T1	0.41	0.9765
	T1	T2	-0.94	0.7848
	T2	T3	-0.22	0.9960
	T0	T3	-0.60	0.9328

Regarding the interactions between proficiency level and instruction type, Tables 6.96 and 6.97 display the results obtained by high and low achievers in the % CV in the interview task. As in the previous analyses, no significant differences were found between CLIL and control learners. Furthermore, as Figure 6.24 shows, CLIL and

control high achievers as well as control low achievers display a decreasing tendency, while CLIL low achievers obtained the highest score at T3 (see Table 6.99 for the percentages of improvement. Finally, the results in Table 6.98 indicate that none of the four groups (CLIL high achievers, CLIL low achievers, control high achievers or control low achievers) developed significantly in this measure at any of the time periods.

Table 6.96
Intergroup comparisons of the percentages of correct verb forms in the interview task. Interactions between high achievers and type of instruction.

	Proficiency	Group	Mean	StdErr	Lower	Upper	F Value	p-value
T0	High	CLIL	60.0%	6.2%	49.0%	73.5%	0.52	0.4717
	High	Control	66.7%	6.8%	54.5%	81.7%		
T1	High	CLIL	64.7%	5.9%	54.1%	77.3%	0.23	0.6294
	High	Control	68.7%	6.1%	57.7%	81.9%		
T2	High	CLIL	57.0%	5.0%	48.0%	67.7%	0.58	0.4489
	High	Control	62.7%	5.4%	52.8%	74.3%		
T3	High	CLIL	54.7%	4.6%	46.4%	64.6%	0.99	0.3210
	High	Control	62.0%	5.3%	52.3%	73.4%		

Table 6.97
Intergroup comparisons of the percentages of correct verb forms in the interview task. Interactions between low achievers and type of instruction.

	Proficiency	Group	Mean	StdErr	Lower	Upper	F Value	p-value
T0	Low	CLIL	51.6%	5.9%	41.1%	64.7%	0.58	0.4485
	Low	Control	58.6%	8.7%	43.7%	78.7%		
T1	Low	CLIL	68.4%	6.6%	56.5%	82.9%	0.56	0.4541
	Low	Control	61.7%	7.5%	48.6%	78.4%		
T2	Low	CLIL	59.5%	5.6%	49.4%	71.6%	0.05	0.8276
	Low	Control	57.8%	6.6%	46.1%	72.4%		
T3	Low	CLIL	68.1%	5.9%	57.3%	80.9%	2.18	0.1417
	Low	Control	56.3%	6.3%	45.2%	70.1%		

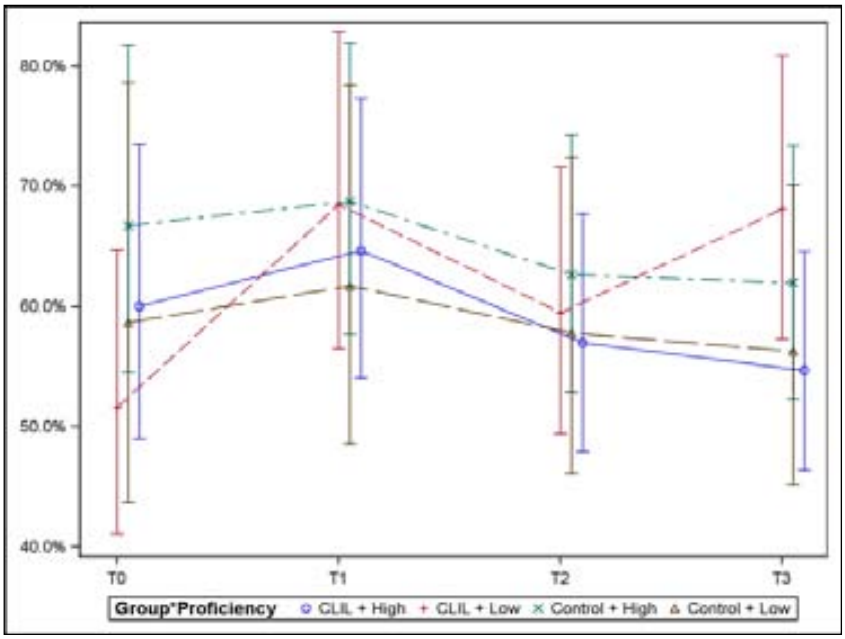
Table 6.98
Intragroup comparisons of the percentages of correct verb forms in the interview task. Interactions between instruction type and proficiency level.

	Time	Time	t Value	p-value
CLIL -High	T0	T1	0.61	0.9302
	T1	T2	-1.13	0.6698
	T2	T3	-0.38	0.9807
	T0	T3	-0.78	0.8651
CLIL -Low	T0	T1	2.06	0.1704
	T1	T2	-1.16	0.6509
	T2	T3	1.20	0.6306
	T0	T3	2.12	0.1515
Control -High	T0	T1	0.25	0.9944
	T1	T2	-0.90	0.8068
	T2	T3	-0.11	0.9995
	T0	T3	-0.65	0.9169
Control -Low	T0	T1	0.29	0.9914
	T1	T2	-0.45	0.9699
	T2	T3	-0.19	0.9975
	T0	T3	-0.25	0.9948

Table 6.99
Percentages of improvement in the percentages correct verb forms according to proficiency and group in the interview task.

Proficiency	Group	I0_1	I1_2	I2_3	I0_3
Low	CLIL	32.5%	-13%	14.4%	31.9%
	Control	5.2%	-6.3%	-2.5%	-3.9%
High	CLIL	7.8%	-11.9%	-4%	-8.8%
	Control	2.9%	-8.7%	-1.1%	-7%

Figure 6.24
Mean percentages of correct verb forms in the interview task. Interactions between instruction type and proficiency level



6.1.3.2.2 Narrative task

With regard to the results obtained in the narrative task, Table 6.100 shows that the control group performs significantly better in the % CV than the CLIL group at T2 (F= 5.84, p= 0.0170) and at T3 (F=6.31, p= 0.0131). As shown in Figure 6.25, the

control group obtains higher mean percentages in this measure already at T0, however, it is not until T2 that this difference reaches statistical difference in favour of the control group (see Table 6.101 for the percentages of improvement)

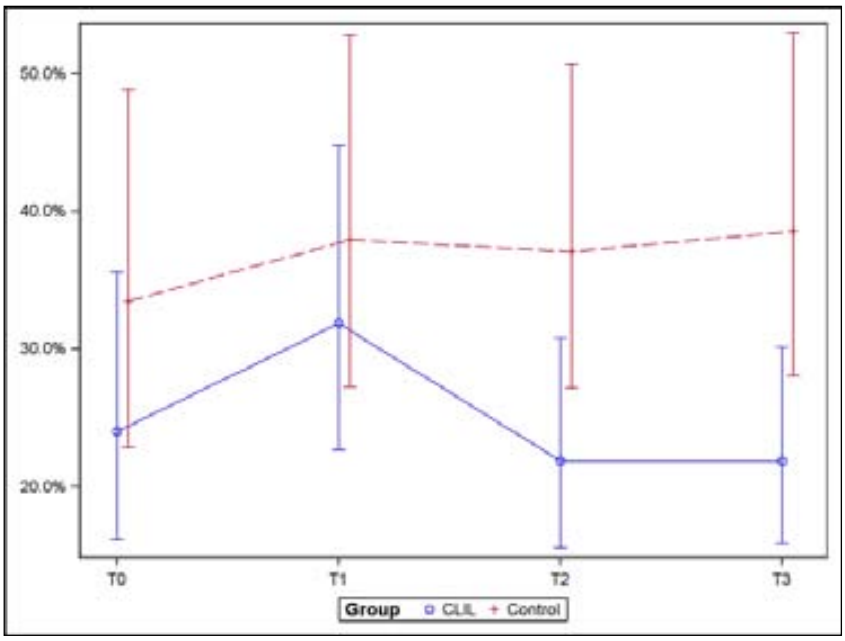
Table 6.100
Intergroup comparisons of the percentages of correct verb forms in the narrative task.

	Group	Mean	StdErr	Lower	Upper	F value	p-value
T0	CLIL	24.0%	4.8%	16.2%	35.5%	1.46	0.2293
	Control	33.4%	6.4%	22.9%	48.9%		
T1	CLIL	31.9%	5.5%	22.7%	44.8%	0.54	0.4635
	Control	37.9%	6.3%	27.2%	52.8%		
T2	CLIL	21.9%	3.8%	15.5%	30.8%	5.84	0.0170*
	Control	37.1%	5.8%	27.1%	50.7%		
T3	CLIL	21.9%	3.6%	15.9%	30.2%	6.31	0.0131*
	Control	38.6%	6.2%	28.1%	53.0%		

Table 6.101
Percentages of improvement in the percentages of correct verb forms in the narrative task.

	I0_1	I1_2	I2_3	I0_3
CLIL	32.7%	-31.3%	-0.0%	-8.9%
Control	13.4%	-2.2%	4.0%	15.3%

Figure 6.25
Mean percentages of correct verb forms in the narrative task.



The results obtained from the intragroup comparisons are displayed in Table 6.102. As can be seen from the table, no significant development was found in any of the groups or time periods analysed.

Table 6.102
Intragroup comparisons of the percentages of correct verb forms in the narrative task.

	Time	Time	t Value	p-Value
CLIL	T0	T1	1.16	0.6518
	T1	T2	-1.69	0.3339
	T2	T3	-0.00	1.0000
	T0	T3	-0.40	0.9788
Control	T0	T1	0.55	0.9472
	T1	T2	-0.11	0.9995
	T2	T3	0.20	0.9971
	T0	T3	0.63	0.9210

As for the interactions between proficiency level and instruction type, the results in Table 6.103 show a significant difference in favour of control high achievers at T2 (F=6.00, p= 0.0157) and at T3 (F=7.01, p= 0.0094). The analysis with low achievers' scores shown in Table 6.104, on the other hand, indicate that control low achievers produce a significantly higher percentage of correct verb forms in the narrative task than CLIL low achievers (F=3.99, p=0.0473) at T0,which disappeared from T1 onwards.

Table 6.103
Intergroup comparisons of the percentages of correct verb forms in the narrative task. Interactions between high achievers and type of instruction.

	Proficiency	Group	Mean	StdErr	Lower	Upper	F Value	p-value
T0	High	CLIL	31.5%	6.9%	20.5%	48.5%	0.18	0.6723
	High	Control	27.2%	7.2%	16.1%	46.0%		
T1	High	CLIL	33.3%	7.0%	21.9%	50.6%	0.76	0.3834
	High	Control	43.5%	9.4%	28.3%	66.7%		
T2	High	CLIL	22.8%	4.8%	15.0%	34.6%	6.00	0.0157*
	High	Control	46.2%	9.0%	31.4%	68.0%		
T3	High	CLIL	24.9%	4.9%	16.9%	36.7%	7.01	0.0094*
	High	Control	51.2%	9.1%	36.0%	73.0%		

Table 6.104
Intergroup comparisons of the percentages of correct verb forms in the narrative task. Interactions between low achievers and type of instruction.

	Proficiency	Group	Mean	StdErr	Lower	Upper	F Value	p-value
T0	Low	CLIL	18.3%	5.8%	9.8%	34.2%	3.99	0.0473*
	Low	Control	41.1%	12.7%	22.3%	75.8%		
T1	Low	CLIL	30.5%	7.6%	18.6%	49.9%	0.06	0.8105
	Low	Control	33.1%	9.4%	18.9%	57.9%		
T2	Low	CLIL	21.0%	5.2%	13.0%	34.1%	1.21	0.2738
	Low	Control	29.8%	7.5%	18.0%	49.1%		
T3	Low	CLIL	19.2%	4.4%	12.2%	30.1%	1.45	0.2297
	Low	Control	29.0%	8.3%	16.5%	51.1%		

Finally, as Table 6.105 shows, no significant development was found in the intragroup comparisons. As Figure 6.26 displays, CLIL high and low achievers along with control low achievers display a similar trend in the % CV throughout the whole study. Control high achievers, on the other hand, is the group that improves the most. As for the percentages of improvement shown in Table 6.106, control high achievers and CLIL low achievers are the groups who improve the most.

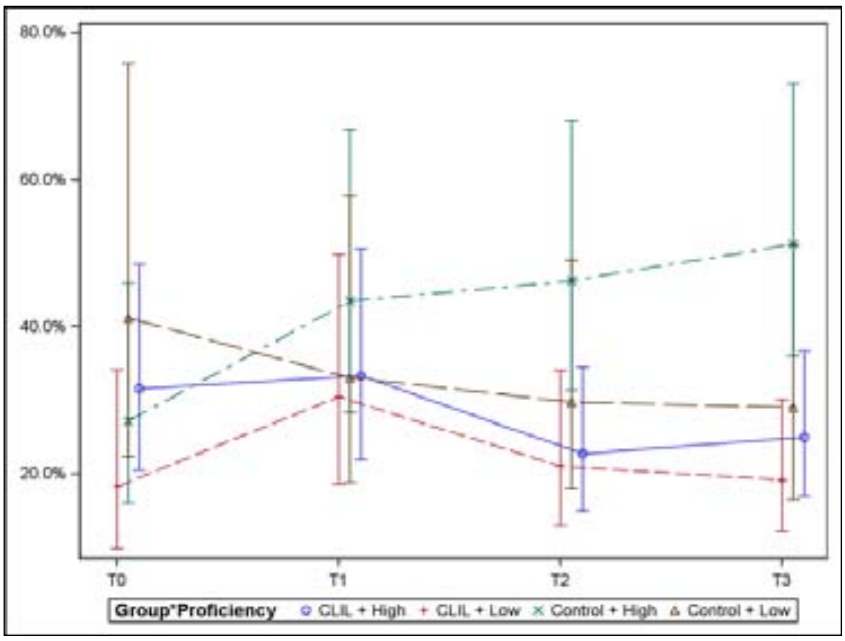
Table 6.105
Intragroup comparisons of the percentages of correct verb forms in the narrative task. Interactions between instruction type and proficiency level.

	Time	Time	t Value	p-value
CLIL -High	T0	T1	0.20	0.9970
	T1	T2	-1.40	0.4986
	T2	T3	0.35	0.9848
	T0	T3	-0.89	0.8084
CLIL -Low	T0	T1	1.35	0.5329
	T1	T2	-1.16	0.6534
	T2	T3	-0.30	0.9903
	T0	T3	0.13	0.9992
Control -High	T0	T1	1.53	0.4250
	T1	T2	0.24	0.9949
	T2	T3	0.48	0.9636
	T0	T3	2.26	0.1135
Control -Low	T0	T1	-0.57	0.9416
	T1	T2	-0.31	0.9894
	T2	T3	-0.07	0.9999
	T0	T3	-0.90	0.8042

Table 6.106
Percentages of improvement in the percentages correct verb forms according to proficiency and group in the narrative task.

Proficiency	Group	I0_1	I1_2	I2_3	I0_3
Low	CLIL	66.6%	-31.1%	-8.5%	4.9%
	Control	-19.4%	-9.9%	-2.6%	-29.4%
High	CLIL	5.7%	-31.5%	9.2%	-20.9
	Control	59.9%	6.2%	10.8%	88.2%

Figure 6.26
Mean percentages of correct verb forms in the narrative task. Interactions between instruction type and proficiency level.



6.1.4 Summary accuracy results

This section will summarise the most relevant findings derived from the analyses of accuracy. To start with, as shown in Table 6.107, it is worth highlighting that the analyses of accuracy show very few significant differences between CLIL and control learners in % EFU and % CV in the two tasks. In the intergroup comparisons, for instance, no differences were found between CLIL and control learners in % EFU in either of the tasks at any of the four data collection times. In the interview task, a non-

significant decrease in the % EFU is observed for the two groups from T0 to T3, while in the narrative task the evolution is more irregular as the % EFU increases and decreases throughout the two years of the study for CLIL and non-CLIL learners. In regard to the second measure used to analyse accuracy (% CV), no differences were found between CLIL and non-CLIL learners in the intergroup comparisons at any of the time periods in the interview task whereas in the narrative task the control group performs significantly better at T2 (p=0.0170) and T3 (p=0.0131) than the CLIL group (see Table 6.107). As for the intragroup analyses, no instances of significant development from T0 to T3 were found in any of the tasks or measures.

Table 6.107
Accuracy. Summary of the significance values obtained in the intergroup comparisons of CLIL and control groups in the interview and narrative task.

	INTERVIEW				NARRATIVE			
	T0	**T1**	**T2**	**T3**	**T0**	**T1**	**T2**	**T3**
% EFU								
% CV							p=0.0170* Control	p=0.0131* Control

The interactions between proficiency level and instruction type yield similar results. The intergroup comparisons between CLIL/control high achievers and CLIL/control low achievers show no differences between the groups according to instruction type in % EFU and % CV in the interview task. In the narrative task, on the other hand, significant differences were found in favour of control high achievers (p=0.0024) and CLIL low achievers (p=0.0410) at T3 in % EFU. As for % CV, control high achievers produce a significantly higher % CV at T2 (p=0.0157) and T3 (p=0.0094) than CLIL high achievers in the narrative task. As for low achievers, the

analyses report a significant difference in favour of control low achievers at T0 (p=0.0473) in % CV in the narrative task. See Tables 6.108 and 6.109.

Table 6.108
Accuracy. Summary of the significance values in the intergroup comparisons of CLIL and control high achievers in the interview and narrative task.

	INTERVIEW				NARRATIVE			
	T0	**T1**	**T2**	**T3**	**T0**	**T1**	**T2**	**T3**
% EFU								p=0.0024* Control high
% CV							p=0.0157* Control high	p=0.0094* Control high

Table 6.109
Accuracy. Summary of the significance values in the intergroup comparisons of CLIL and control low achievers in the interview and narrative task.

	INTERVIEW				NARRATIVE			
	T0	**T1**	**T2**	**T3**	**T0**	**T1**	**T2**	**T3**
% EFU								p=0.0410* CLIL low
% CV					p=0.0473* Control low			

The intragroup comparisons of CLIL high achievers, control high achievers, CLIL low achievers and control low achievers report a significant decrease in the % EFU of CLIL high achievers from T0 to T3 (p=0.0495) in the interview task while no significant development was found in % CV. In the narrative task, no significant differences in development were found in either of the two accuracy measure. See Table 6.110.

Table 6.110
Accuracy. Summary of the significance values of the intragroup comparisons (from T0 to T3) of CLIL and control high achievers in the interview and narrative task.

	INTERVIEW		NARRATIVE	
	CLIL-high	**CONTROL-high**	**CLIL-high**	**CONTROL-high**
% EFU	p=0.0495* (decrease)			
% CV				

In short, as seen from this summary, accuracy, especially when measured by error-free units, seems to be one of the most unaffected areas by instruction, proficiency level or time. However, the percentages of improvement do seem to suggest a slight advantage for CLIL learners.

6.1.5 Fluency

In this section, the results obtained by the CLIL and control groups in the fluency measures in the interview and narrative tasks will be presented. One of the subdimensions of fluency under analysis in this study is speed fluency, which will be measured by speech rate in words (SPRW). In addition, the results obtained from an L1 word ratio (L1WR) will also be provided in this section.

6.1.5.1 Speed fluency: speech rate in words

6.1.5.1.1 Interview task

Table 6.111 displays the results obtained from the intergroup comparisons between CLIL and control learners in speech rate in words in the interview task. As the table shows, the control group was significantly better at T0 (F=4.67, p=0.0341), T1 (F=4.84, p= 0.0311) and T2 (F=11.95, p= 0.0009). At T3, the differences between the

groups in this fluency measure disappear. As seen from Figure 6.27 and Table 6.112, the control group obtained similar results at T2 and T3, while the CLIL group improved considerably from T2 to T3.

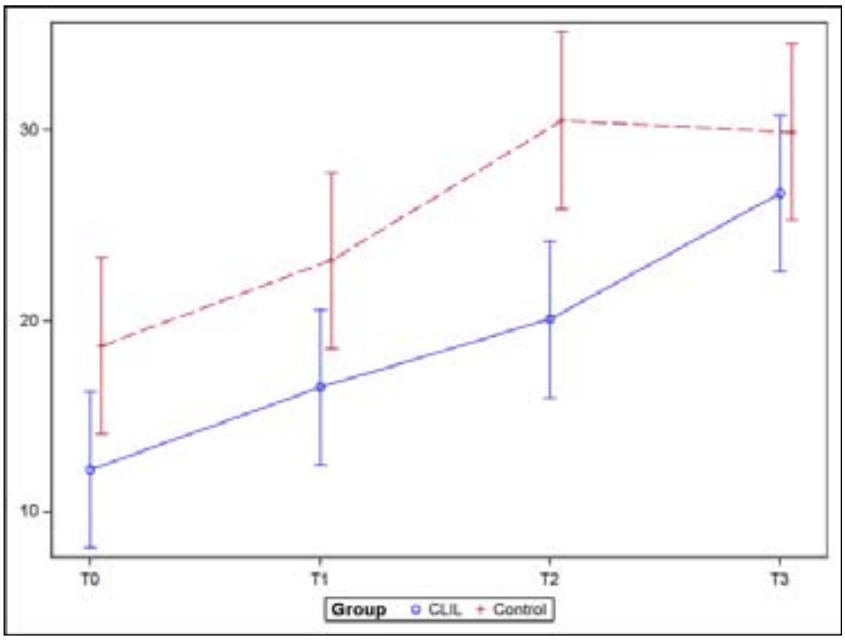
Table 6.111
Intergroup comparisons of speech rate in words in the interview task.

	Group	Mean	StdErr	Lower	Upper	F value	p-value
T0	CLIL	12.20	2.05	8.11	16.30	4.67	0.0341*
	Control	18.70	2.31	14.09	23.32		
T1	CLIL	16.53	2.05	12.43	20.63	4.84	0.0311*
	Control	23.17	2.32	18.55	27.80		
T2	CLIL	20.08	2.06	15.98	24.18	11.95	0.0009*
	Control	30.52	2.34	25.86	35.18		
T3	CLIL	26.68	2.05	22.59	30.77	1.14	0.2886
	Control	29.90	2.31	25.29	34.51		

Table 6.112
Percentages of improvement in speech rate in words in the interview task.

	I0_1	I1_2	I2_3	I0_3
CLIL	35.5%	21.5%	32.9%	118.7%
Control	23.9%	31.7%	-2.0%	59.8%

Figure 6.27
Mean scores of speech rate in words in the interview task.



The results of the intragroup comparisons displayed in Table 6.113 show that both groups improve significantly from T0 to T3 (p<.0001) in speech rate in words. In addition, the CLIL group also improves significantly from T0 to T1 (p=0.0246) and from T2 to T3 (p=0.0001). Similarly, the control group obtains significant results in their improvement from T0 to T1 (p=0.0485) and from T1 to T2 (p=0.0003).

Table 6.113
Intragroup comparisons of speech rate in words in the interview task.

	Time	Time	t Value	p-value
CLIL	T0	T1	2.87	0.0246*
	T1	T2	2.35	0.0930
	T2	T3	4.36	0.0001*
	T0	T3	9.60	<.0001*
Control	T0	T1	2.61	0.0485*
	T1	T2	4.21	0.0003*
	T2	T3	-0.36	0.9845
	T0	T3	6.58	<.0001*

With regard to the interactions between proficiency level and instruction type, Table 6.114 shows that high achievers from the control group perform significantly better at T0 (F=5.05, p= 0.0278), T1 (F=7.42, p= 0.0081) and T2 (F= 15.25, p=0.0002) than the CLIL group in speech rate in words. In spite of this, the significance difference between the two groups of high achievers disappears at T3. As for low achievers (see Table 6.115), no differences were found between CLIL low achievers and control low achievers at any of the data collection times.

Table 6.114
Intergroup comparisons of speech rate in words in the interview task. Interactions between high achievers and type of instruction.

	Proficiency	Group	Mean	StdErr	Lower	Upper	F Value	p-value
T0	High	CLIL	11.46	2.63	6.23	16.70	5.05	0.0278*
	High	Control	21.03	3.13	14.78	27.28		
T1	High	CLIL	16.30	2.63	11.06	21.54	7.42	0.0081*
	High	Control	28.03	3.20	21.66	34.39		
T2	High	CLIL	20.96	2.63	15.72	26.20	15.25	0.0002*
	High	Control	37.69	3.18	31.35	44.02		
T3	High	CLIL	27.14	2.63	21.90	32.38	2.79	0.0994
	High	Control	34.25	3.13	28.00	40.50		

Table 6.115
Intergroup comparisons of speech rate in words in the interview task. Interactions between low achievers and type of instruction.

	Proficiency	Group	Mean	StdErr	Lower	Upper	F Value	p-value
T0	Low	CLIL	12.94	2.80	7.36	18.51	0.73	0.3948
	Low	Control	16.38	3.52	9.35	23.40		
T1	Low	CLIL	16.76	2.80	11.18	22.34	0.15	0.6980
	Low	Control	18.32	3.53	11.28	25.36		
T2	Low	CLIL	19.20	2.80	13.62	24.79	1.06	0.3058
	Low	Control	23.35	3.52	16.32	30.37		
T3	Low	CLIL	26.22	2.80	20.65	31.79	0.03	0.8672
	Low	Control	25.55	3.52	18.52	32.57		

The intragroup comparisons displayed in Table 6.116 show that all four groups (CLIL high achievers, CLIL low achievers, control high achievers and control low achievers) improved significantly from T0 to T3 in speech rate in words. In addition, CLIL high and low achievers show significant improvement from T2 to T3, while control high achievers improve significantly from T0 to T1 and from T1 to T2. As Figure 6.28 shows, control high achievers improve noticeably from T0 to T2, while CLIL high, CLIL low and control low achievers show more regular improvement from T0 to T3. According to the analyses, control high achievers are significantly better in speech rate in words than control low achievers at T2 (p=0.0040). As for the percentages of improvement (Table 6.117), the results show that CLIL learners display higher percentages of improvement from T2 to T3 and from T0 to T3.

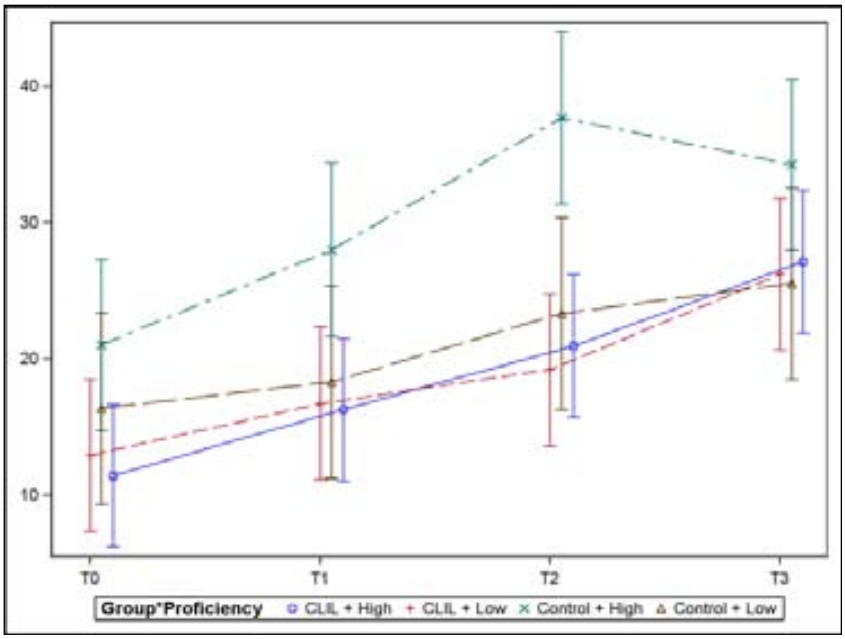
Table 6.116
Intragroup comparisons of speech rate in words in the interview task. Interactions between instruction type and proficiency level.

	Time	Time	t Value	p-value
CLIL -High	T0	T1	2.51	0.0635
	T1	T2	2.41	0.0797
	T2	T3	3.20	0.0092*
	T0	T3	8.12	<.0001*
CLIL -Low	T0	T1	1.85	0.2525
	T1	T2	1.19	0.6370
	T2	T3	3.41	0.0048*
	T0	T3	6.47	<.0001*
Control -High	T0	T1	2.93	0.0207*
	T1	T2	3.94	0.0008*
	T2	T3	-1.45	0.4706
	T0	T3	5.74	<.0001*
Control -Low	T0	T1	0.75	0.8770
	T1	T2	1.93	0.2205
	T2	T3	0.85	0.8309
	T0	T3	3.54	0.0030*

Table 6.117
Percentages of improvement in speech rate in words according to proficiency and group in the interview task.

Proficiency	Group	I0_1	I1_2	I2_3	I0_3
Low	CLIL	29.5%	14.55	36.5%	102.6%
	Control	11.8%	27.4%	9.4%	55.9%
High	CLIL	42.2%	28.5%	29.4%	136.8%
	Control	33.2%	34.4%	-9.1%	62.8%

Figure 6.28
Mean scores of speech rate in words in the interview task. Interactions between instruction type and proficiency level



6.1.5.1.2 Narrative task

Regarding the results obtained in speech rate in words in the narrative task, Table 6.118 displays the intergroup comparisons between CLIL and control learners. As can be seen from the table, a significant difference in favour of the control group was found at T2 (F= 4.86, p = 0.0300). At T3, however, the difference between the groups did not reach significance due to the increase in speech rate in words of the CLIL group from T2 to T3 (see Table 6.119). Despite this, as Figure 6.29 shows, the control group obtains higher mean scores at the four data collection times.

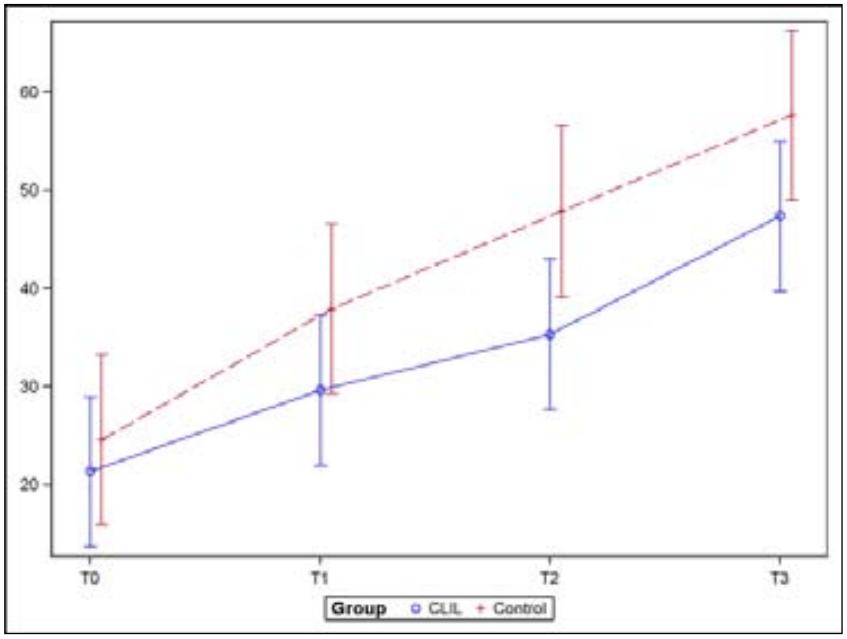
Table 6.118
Intergroup comparisons of speech rate in words in the narrative task.

	Group	Mean	StdErr	Lower	Upper	F value	p-value
T0	CLIL	21.37	3.85	13.72	29.02	0.34	0.5641
	Control	24.63	4.34	16.02	33.25		
T1	CLIL	29.64	3.85	21.99	37.30	2.16	0.1455
	Control	37.97	4.36	29.32	46.63		
T2	CLIL	35.37	3.86	27.70	43.04	4.86	0.0300*
	Control	47.87	4.41	39.11	56.62		
T3	CLIL	47.40	3.85	39.75	55.05	3.32	0.0719
	Control	57.68	4.34	49.06	66.30		

Table 6.119
Percentages of improvement in speech rate in words in the narrative task.

	I0_1	I1_2	I2_3	I0_3
CLIL	38.7%	19.3%	34.0%	121.8%
Control	54.2%	26.0%	20.5%	134.2%

Figure 6.29
Mean scores of speech rate in words in the narrative task.



As for the intragroup comparisons, the results displayed in Table 6.120 indicate that both groups improve significantly from T0 to T3 in the narrative task. In addition, the CLIL group shows significant development from T2 to T3 (p=0.0042), while the control group improves significantly from T0 to T1 (p=0.0052).

Table 6.120
Intragroup comparisons of speech rate in words in the narrative task.

	Time	Time	t Value	p-value
CLIL	T0	T1	2.37	0.0873
	T1	T2	1.64	0.3607
	T2	T3	3.44	0.0042*
	T0	T3	7.47	<.0001*
Control	T0	T1	3.38	0.0052*
	T1	T2	2.46	0.0716
	T2	T3	2.45	0.0722
	T0	T3	8.42	<.0001*

The results obtained from the interactions between proficiency level and instruction type are shown in Tables 6.121 and 6.122. As seen from Table 6.121, significant differences were found between CLIL and control high achievers at T1 (F=4.44, =0.0377), T2 (F=5.00, p=0.0278) and T3 (F=10.03, p=0.0021) in favour of control high achievers. As seen from the F values, the differences between the two groups grows gradually over time. As for low achievers, no differences were found between CLIL and control low achievers (see Table 6.122).

Table 6.121
Intergroup comparisons of speech rate in words in the narrative task. Interactions between high achievers and type of instruction.

	Proficiency	Group	Mean	StdErr	Lower	Upper	F Value	p-value
T0	High	CLIL	19.88	4.92	10.09	29.66	0.85	0.3582
	High	Control	27.25	5.88	15.58	38.93		
T1	High	CLIL	29.61	4.92	19.82	39.40	4.44	0.0377*
	High	Control	46.72	6.05	34.71	58.73		
T2	High	CLIL	37.49	4.93	27.69	47.28	5.00	0.0278*
	High	Control	55.50	6.01	43.57	67.42		
T3	High	CLIL	45.71	4.92	35.92	55.49	10.03	0.0021*
	High	Control	70.99	5.88	59.32	82.67		

Table 6.122
Intergroup comparisons of speech rate in words in the narrative task. Interactions between low achievers and type of instruction.

	Proficiency	Group	Mean	StdErr	Lower	Upper	F Value	p-value
T0	Low	CLIL	22.86	5.24	12.44	33.27	0.01	0.9115
	Low	Control	22.02	6.60	8.90	35.13		
T1	Low	CLIL	29.67	5.25	19.24	40.10	0.00	0.9532
	Low	Control	29.23	6.63	16.06	42.40		
T2	Low	CLIL	33.26	5.25	22.82	43.69	0.86	0.3563
	Low	Control	40.24	6.61	27.10	53.37		
T3	Low	CLIL	49.10	5.24	38.69	59.51	0.39	0.5316
	Low	Control	44.37	6.60	31.25	57.49		

Finally, the intragroup comparisons of CLIL high and low achievers and control high and low achievers are displayed in Table 6.123. As seen from the table, the four groups improve significantly from T0 to T3. In addition, significant improvement was

also found for CLIL low achievers from T2 to T3 and control high achievers from T0 to T1 and from T2 to T3. As Figure 6.30 and Table 6.124 show, control high achievers is the group who improves the most from T0 to T3 whereas the other three groups develop steadily from T0 to T3. A statistically significant difference was found between control high and low achievers in favour of high achievers at T3 (p=0.0040).

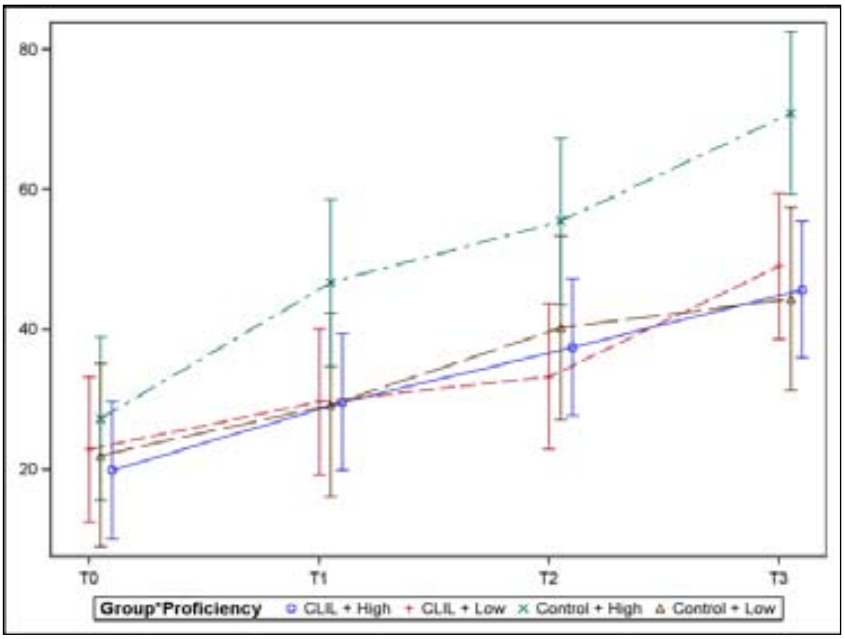
Table 6.123
Intragroup comparisons of speech rate in words in the narrative task. Interactions between instruction type and proficiency level.

	Time	Time	t Value	p-value
CLIL -High	T0	T1	2.18	0.1329
	T1	T2	1.77	0.2944
	T2	T3	1.84	0.2585
	T0	T3	5.80	<.0001*
CLIL -Low	T0	T1	1.43	0.4804
	T1	T2	0.75	0.8756
	T2	T3	3.33	0.0060*
	T0	T3	5.53	<.0001*
Control -High	T0	T1	3.53	0.0031*
	T1	T2	1.55	0.4101
	T2	T3	2.84	0.0267*
	T0	T3	8.23	<.0001*
Control -Low	T0	T1	1.20	0.6268
	T1	T2	1.83	0.2631
	T2	T3	0.69	0.9002
	T0	T3	3.74	0.0015*

Table 6.124
Percentages of improvement in speech rate in words according to proficiency and group in the narrative task.

Proficiency	Group	I0_1	I1_2	I2_3	I0_3
Low	CLIL	29.7%	12%	47.6%	114.7%
	Control	32.7%	37.6%	10.2%	101.4%
High	CLIL	48.9%	26.6%	21.6%	129.9%
	Control	71.6%	18.6%	27.9%	160.5%

Figure 6.30
Mean scores of speech rate in words in the narrative task. Interactions between instruction type and proficiency level.



6.1.5.2 Breakdown (dys)fluency: percentage of L1 word ratio

6.1.5.2.1 Interview task

Table 6.125 shows the results of the intergroup comparisons between CLIL and control learners in the percentage of L1 words produced during the interview task. As seen from the table, no differences were found between the groups in this measure as both groups produced a similar number of L1 words throughout the study. In addition, Table 6.126 and Figure 6.31 show that both groups decrease noticeably in this measure from T0 to T3. According to the results shown in Table 6.127, both groups decrease significantly the number of L1 words produced in the interview task from T0 to T3 ($p<.0001$). In addition, the CLIL group also decreases significantly from T0 to T1 (t=

-3.39, p= 0.0049), from T1 to T 2 (t=-2.81, p= 0.0291) and from T 2 to T 3 (t=-2.81, p=0.0286). Likewise, the control group shows a significant decrease in the % L1WR from T0 to T1 (t=-2.94, p=0.0198) and from T1 to T2 (t=-3.62, p=0.0023).

Table 6.125
Intergroup comparisons of the percentages of L1words in the interview task.

	Group	Mean	StdErr	Lower	Upper	F value	p-value
T0	CLIL	22.0%	4.5%	14.5%	33.2%	1.02	0.3174
	Control	16.1%	3.9%	9.9%	26.1%		
T1	CLIL	14.3%	3.0%	9.3%	21.9%	1.39	0.2423
	Control	9.9%	2.4%	6.0%	16.1%		
T2	CLIL	8.8%	2.1%	5.6%	14.0%	3.94	0.0503
	Control	4.5%	1.3%	2.6%	7.9%		
T3	CLIL	4.7%	1.2%	2.8%	7.8%	0.02	0.8772
	Control	4.4%	1.3%	2.5%	7.8%		

Table 6.126
Percentages of improvement (decrease) in the percentage of L1 words in the interview task.

	I0_1	I1_2	I2_3	I0_3
CLIL	-35.1%	-38.1%	-47.3%	-78.8%
Control	-38.5%	-54.5%	-2.1%	-72.6%

Figure 6.31
Mean percentages of L1 words in the interview task.

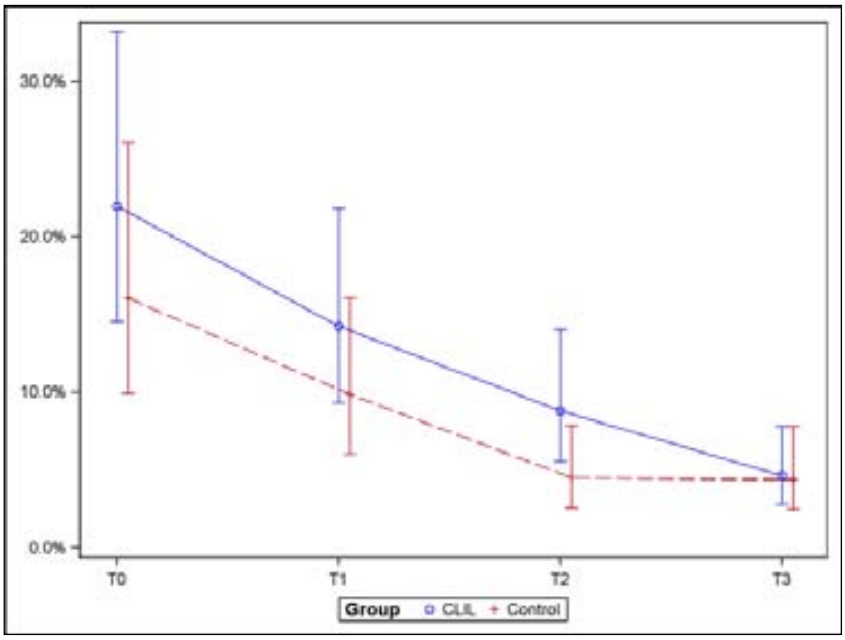


Table 6.127
Intragroup comparisons of the percentages of L1words in the interview task.

	Time	Time	t Value	p-value
CLIL	T0	T1	-3.39	0.0049*
	T1	T2	-2.81	0.0291*
	T2	T3	-2.81	0.0286*
	T0	T3	-7.84	<.0001*
Control	T0	T1	-2.94	0.0198*
	T1	T2	-3.62	0.0023*
	T2	T3	-0.08	0.9998
	T0	T3	-5.67	<.0001*

The results of the intergroup comparisons between CLIL high achievers and control high achievers are displayed in Table 6.128. As seen from table, CLIL high achievers produce higher L1 word ratios than the control group at T0 (F=4.38, p=0.0402), T1 (F=4.07, p=0.0473) and at T2 (F=5.63, p=0.0195). At T3, no differences were found between the two groups of high achievers in this measure. As for low

achievers, the results in Table 6.129 show no differences between CLIL and control low achievers at any of the four times of data collection.

Table 6.128
Intergroup comparisons of the percentages of L1 words in the interview task. Interactions between high achievers and type of instruction.

	Proficiency	Group	Mean	StdErr	Lower	Upper	F Value	p-value
T0	High	CLIL	24.3%	6.4%	14.3%	41.2%	4.38	0.0402*
	High	Control	9.4%	3.3%	4.7%	18.7%		
T1	High	CLIL	14.9%	4.1%	8.6%	25.8%	4.07	0.0473*
	High	Control	5.8%	2.1%	2.9%	11.8%		
T2	High	CLIL	8.7%	2.6%	4.8%	15.7%	5.63	0.0195*
	High	Control	2.6%	1.0%	1.1%	5.8%		
T3	High	CLIL	4.2%	1.4%	2.2%	8.2%	0.77	0.3829
	High	Control	2.6%	1.1%	1.1%	6.0%		

Table 6.129
Intergroup comparisons of the percentages of L1words in the interview task. Interactions between low achievers and type of instruction.

	Proficiency	Group	Mean	StdErr	Lower	Upper	F Value	p-value
T0	Low	CLIL	19.9%	5.5%	11.4%	34.8%	0.65	0.4299
	Low	Control	27.4%	9.5%	13.7%	54.8%		
T1	Low	CLIL	13.6%	3.8%	7.8%	24.0%	0.26	0.6103
	Low	Control	16.7%	5.8%	8.3%	33.4%		
T2	Low	CLIL	8.9%	2.7%	4.9%	16.3%	0.09	0.7597
	Low	Control	7.9%	3.0%	3.7%	16.7%		
T3	Low	CLIL	5.1%	1.7%	2.7%	9.8%	0.75	0.3887
	Low	Control	7.5%	2.9%	3.4%	16.4%		

Concerning the intragroup comparisons of CLIL high achievers, control high achievers, CLIL low achievers and control low achievers, the results in Table 6.130 and Figure 6.32 indicate that the four groups decrease significantly in the production of L1 words during the interview task from T0 to T3. However, as Table 6.131 displays, CLIL high achievers is the group who decrease the most.

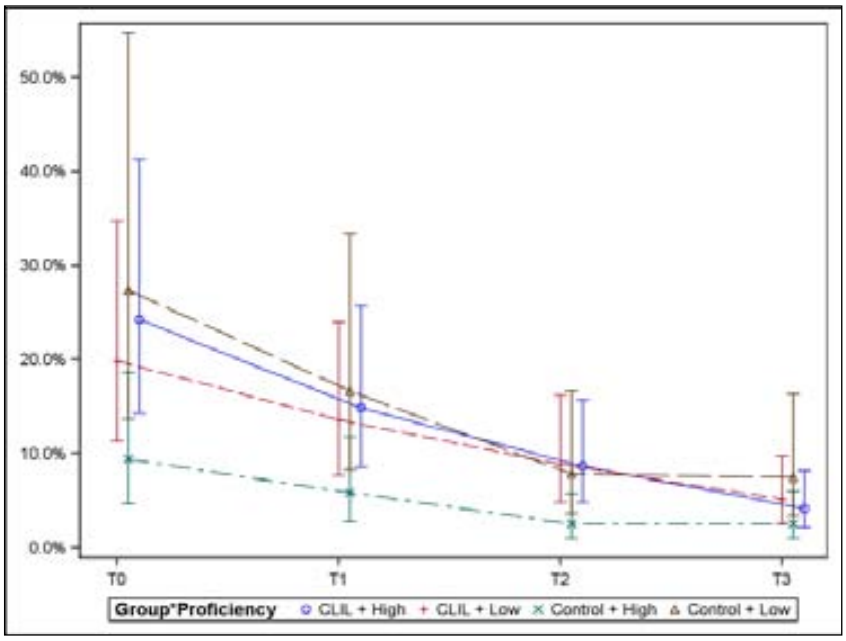
Table 6.130
Intragroup comparisons of the percentages of L1 words in the interview task. Interactions between instruction type and proficiency level.

	Time	Time	t Value	p-value
CLIL -High	T0	T1	-2.90	0.0221*
	T1	T2	-2.43	0.0769
	T2	T3	-2.48	0.0680
	T0	T3	-6.75	<.0001*
CLIL -Low	T0	T1	-2.31	0.1018
	T1	T2	-2.13	0.1502
	T2	T3	-2.10	0.1577
	T0	T3	-5.73	<.0001*
Control -High	T0	T1	-1.86	0.2490
	T1	T2	-2.46	0.0714
	T2	T3	0.01	1.0000
	T0	T3	-3.67	0.0020*
Control -Low	T0	T1	-2.55	0.0569
	T1	T2	-3.02	0.0156*
	T2	T3	-0.15	0.9988
	T0	T3	-4.78	<.0001*

Table 6.131
Percentages of improvement (decrease) in the percentage of L1 words according to proficiency and group in the interview task.

Proficiency	Group	I0_1	I1_2	I2_3	I0_3
Low	CLIL	-31.6%	-34.5%	-42.6%	-74.3%
	Control	-39%	-52.6%	-5%	-72.6%
High	CLIL	-38.6%	-41.6%	-51.7%	-82.7%
	Control	-38.2%	-55.1%	0%	-72.3%

Figure 6.32
Mean percentages of L1 words in the interview task. Interactions between instruction type and proficiency level.



6.1.5.2.2 Narrative task

As for the percentages of L1 words produced in the narrative task by the CLIL and control groups, Table 6.132 shows no significant differences between the groups at any of the four data collection times. As can be seen from Table 6.133, Figure 6.33 and Table 6.134, both groups decrease significantly ($p<.0001$) in the production of L1 words in this task over the two years of the study (from T0 to T3). In addition, the CLIL group also shows a significant decrease from T2 to T3 ($p=0.0073$).

Table 6.132
Intergroup comparisons of the percentages of L1words in the narrative task.

	Group	Mean	StdErr	Lower	Upper	F value	p-value
T0	CLIL	13.5%	3.1%	8.5%	21.4%	0.19	0.6671
	Control	11.6%	3.1%	6.9%	19.7%		
T1	CLIL	10.9%	2.5%	6.8%	17.3%	1.35	0.2485
	Control	7.2%	2.0%	4.1%	12.5%		
T2	CLIL	7.7%	1.8%	4.8%	12.4%	3.10	0.0819
	Control	4.0%	1.2%	2.2%	7.3%		
T3	CLIL	3.4%	1.0%	1.9%	6.1%	0.09	0.7700
	Control	3.0%	1.0%	1.6%	5.9%		

Table 6.133
Percentages of improvement (decrease) in the percentage of L1 words in the narrative task.

	I0_1	I1_2	I2_3	I0_3
CLIL	-19.4%	-28.8%	-55.8%	-74.6%
Control	-38.3%	-43.9%	-24.6%	-73.9%

Figure 6.33
Mean percentages of L1 words in the narrative task.

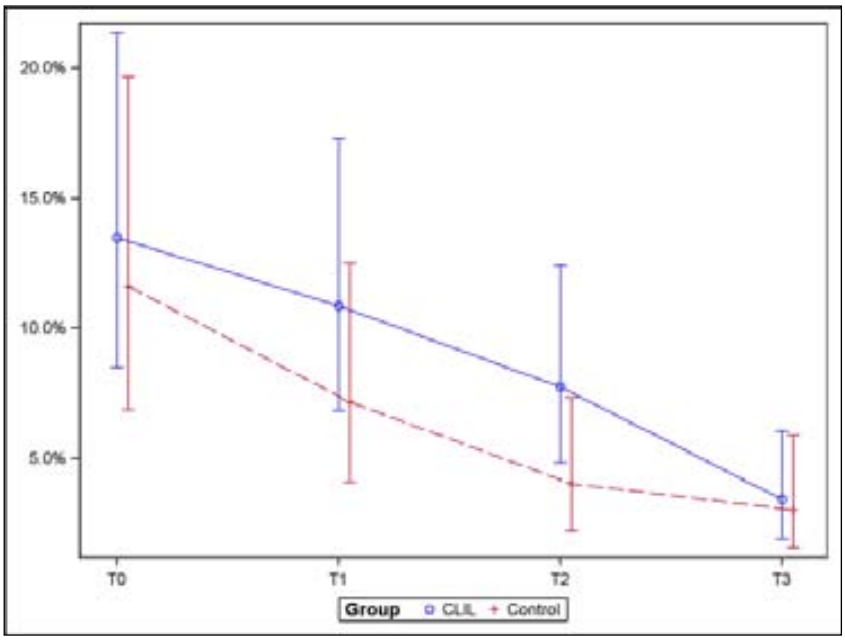


Table 6.134
Intragroup comparisons of the percentages of L1words in the narrative task.

	Time	Time	t Value	p-value
CLIL	T0	T1	-1.28	0.5795
	T1	T2	-1.91	0.2282
	T2	T3	-3.28	0.0073*
	T0	T3	-5.68	<.0001*
Control	T0	T1	-2.20	0.1299
	T1	T2	-2.17	0.1362
	T2	T3	-0.87	0.8178
	T0	T3	-4.68	<.0001*

In regard to the interactions between proficiency level and instruction type, Tables 6.135 and 6.136 show no significant differences between CLIL and control high achievers or CLIL and control low achievers at any of the four data collection times.

Table 6.135
Intergroup comparisons of the percentages of L1 words in the narrative task. Interactions between high achievers and type of instruction.

	Proficiency	Group	Mean	StdErr	Lower	Upper	F Value	p-value
T0	High	CLIL	14.1%	4.1%	7.8%	25.3%	1.66	0.2012
	High	Control	7.3%	2.8%	3.4%	15.8%		
T1	High	CLIL	10.2%	3.1%	5.6%	18.7%	2.75	0.1007
	High	Control	4.2%	1.8%	1.8%	9.6%		
T2	High	CLIL	8.4%	2.5%	4.6%	15.3%	3.05	0.0837
	High	Control	3.2%	1.4%	1.4%	7.6%		
T3	High	CLIL	3.2%	1.2%	1.5%	6.5%	0.57	0.4517
	High	Control	2.0%	1.0%	0.7%	5.3%		

Table 6.136
Intergroup comparisons of the percentages of L1words in the narrative task. Interactions between low achievers and type of instruction.

	Proficiency	Group	Mean	StdErr	Lower	Upper	F Value	p-value
T0	Low	CLIL	12.9%	3.9%	7.0%	23.8%	0.70	0.4081
	Low	Control	18.5%	6.8%	8.8%	38.9%		
T1	Low	CLIL	11.6%	3.6%	6.3%	21.4%	0.02	0.8929
	Low	Control	12.3%	4.7%	5.7%	26.5%		
T2	Low	CLIL	7.2%	2.2%	3.8%	13.4%	0.56	0.4571
	Low	Control	5.0%	2.1%	2.2%	11.6%		
T3	Low	CLIL	3.7%	1.3%	1.8%	7.4%	0.26	0.6094
	Low	Control	4.7%	2.0%	2.0%	11.1%		

Finally, as Table 6.137 shows, the four groups (CLIL high achievers, control high achievers, CLIL low achievers and control low achievers) decrease significantly in their production of L1 words from T0 to T3, among other periods (see Table 6.138 for the percentages of decrease). Control high achievers, however, is the group that shows the smallest decrease due to their already low percentage of L1 words at T0 (see Figure 6.34).

Table 6.137
Intragroup comparisons of the percentages of L1 words in the narrative task. Interactions between instruction type and proficiency level.

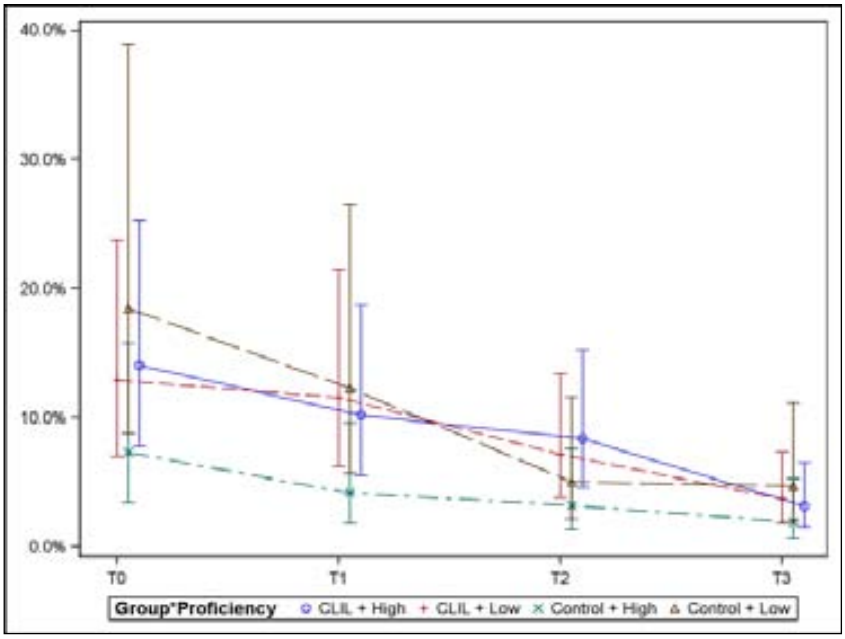
	Time	Time	t Value	p-value
CLIL -High	T0	T1	-1.38	0.5114
	T1	T2	-0.81	0.8478
	T2	T3	-3.07	0.0136*
	T0	T3	-4.84	<.0001*
CLIL -Low	T0	T1	-0.55	0.9456
	T1	T2	-2.27	0.1108
	T2	T3	-2.42	0.0780
	T0	T3	-4.82	<.0001*

	Time	Time	t Value	p-value
Control -High	T0	T1	-1.53	0.4241
	T1	T2	-0.63	0.9231
	T2	T3	-1.00	0.7521
	T0	T3	-2.85	0.0255*
Control -Low	T0	T1	-1.59	0.3894
	T1	T2	-2.74	0.0345*
	T2	T3	-0.17	0.9982
	T0	T3	-4.17	0.0003

Table 6.138
Percentages of improvement (decrease) in the percentage of L1 words according to proficiency and group in the narrative task.

Proficiency	Group	I0_1	I1_2	I2_3	I0_3
Low	CLIL	-10%	-37.9%	-48.6%	-71.3%
	Control	-33.5%	-59.3%	-6%	-74.5%
High	CLIL	-27.6%	-17.6%	-61.9%	-77.3%
	Control	-42.4%	-23.8%	-37.5%	-72.6%

Figure 6.34
Mean percentages of L1 words in the narrative task. Interactions between instruction type and proficiency level.



6.1.6 Summary fluency results

This section will provide a summary of the main findings related to fluency in the interview and narrative. As in previous summaries, the results obtained from the intergroup comparisons between CLIL and control learners will be presented first, followed by the intragroup comparisons of CLIL and control learners. Finally, this section will offer a summary of the most relevant results obtained from the interactions between type of instruction and proficiency level.

Table 6.139 shows the significant values obtained from the two fluency measures in the interview and narrative tasks. As shown in the table, the only significant differences were found in SPRW in favour of the control group in the interview task at T0 (p=0.0341), T 1 (p=0.0311) a nd T 2 (p=0.0009) a nd i n t he na rrative t ask a t T 2 (p=0.0300). Interestingly, all the differences disappear at T3.

Table 6.139
Fluency. Summary of the significance values obtained in the intergroup comparisons of CLIL and control groups in the interview and narrative task.

	INTERVIEW				NARRATIVE			
	T0	**T1**	**T2**	**T3**	**T0**	**T1**	**T2**	**T3**
SPRW	p=0.0341* Control	p=0.0311* Control	p=0.0009* Control				p=0.0300* Control	
% L1WR								

As for the intragroup comparisons, Table 6.140 shows that both CLIL and control groups developed significantly from T0 to T3 in the two fluency measures and in the two tasks. In the case of % L1WR, significant development refers to decreases in the number of L1words produced by the learners from T0 to T3.

Table 6.140
Fluency. Summary of the significance values obtained in the intragroup comparisons (from T0 to T3) of CLIL and control groups in the interview and narrative task.

	INTERVIEW		NARRATIVE	
	CLIL	**CONTROL**	**CLIL**	**CONTROL**
SPRW	p<.0001*	p<.0001*	p<.0001*	p<.0001*
% L1WR	p<.0001* (decrease)	p<.0001* (decrease)	p<.0001* (decrease)	p<.0001* (decrease)

The intergroup comparisons of CLIL and control high achievers displayed in Table 6.141 indicate that the control group is significantly better in SPRW at T0 (p=0.0278), T1 (p=0.0081) and T2 (p=0.0002) in the interview task. In addition, the results in the table also show that the CLIL group produces a significantly higher number of L1 words at T0 (p=0.0402), T1 (p=0.0473) and T2 (p=0.0195) in the interview task. At T3, however, no differences were found between the groups in either measure in the interview task. In the narrative task, the only significant differences found are in SPRW in favour of control high achievers at T1 (p=0.0377), T2 (p=0.0278) and T3 (p=0.0021). Concerning CLIL and control low achievers, no differences were found between the groups in any of the two tasks and measures.

Table 6.141
Fluency. Summary of the significance values in the intergroup comparisons of CLIL and control high achievers in the interview and narrative task.

	INTERVIEW				NARRATIVE			
	T0	**T1**	**T2**	**T3**	**T0**	**T1**	**T2**	**T3**
SPRW	p=0.0278* Control-high	p=0.0081* Control-high	p=0.0002* Control-high			p=0.0377* Control-high	p=0.0278* Control-high	p=0.0021* Control-high
% L1WR	p=0.0402* CLIL-high	p=0.0473* CLIL-high	p=0.0195* CLIL-high					

Finally, the intragroup comparisons of CLIL and control high achievers displayed in Table 6.142 indicate that both groups develop significantly in the two measures and tasks from T0 to T3. Similarly, as the results in Table 6.143 show, significant development was found for CLIL and control low achievers in the two fluency measures and tasks from T0 to T3. In the case of the % L1 words, as indicated previously, the results show a significant decrease in the amount of L1 language produced by the learners.

Table 6.142
Fluency. Summary of the significance values of the intragroup comparisons (from T0 to T3) of CLIL and control high achievers in the interview and narrative task.

	INTERVIEW		**NARRATIVE**	
	CLIL-high	**CONTROL-high**	**CLIL-high**	**CONTROL-high**
SPRW	p<.0001*	p<.0001*	p<.0001*	p<.0001*
% L1WR	p<.0001* (decrease)	p=0.0020* (decrease)	p<.0001* (decrease)	p=0.0255* (decrease)

Table 6.143
Fluency. Summary of the significance values of the intragroup comparisons (from T0 to T3) of CLIL and control low achievers in the interview and narrative task.

	INTERVIEW		**NARRATIVE**	
	CLIL-low	**CONTROL-low**	**CLIL-low**	**CONTROL-low**
SPRW	p<.0001*	p=0.0030*	p<.0001*	p=0.0015*
% L1WR	p<.0001* (decrease)	p<.0001* (decrease)	p<.0001* (decrease)	p=0.0003* (decrease)

6.1.7 Correlation results

In order to explore the relationship between the three elements that make up the CAF construct, Spearman's rho correlation tests were conducted between accuracy-syntactic complexity, syntactic complexity-fluency and accuracy-fluency. The measures used to run the correlation tests were % CU for complexity, % EFU for accuracy and SPRW for fluency. The correlation tests were carried out with the results obtained from the interview task only, as the results from the narrative task were too low to examine the relationships between the measures.

The strength of the correlations was determined using Cohen's (1988) guidelines, which suggest that r values between .10 and .29 are considered to indicate a small or weak correlation between two variables. R values between .30 and .49 mark a medium or moderate correlation while r values between .50 and 1.0 reflect a large or strong correlation between variables.

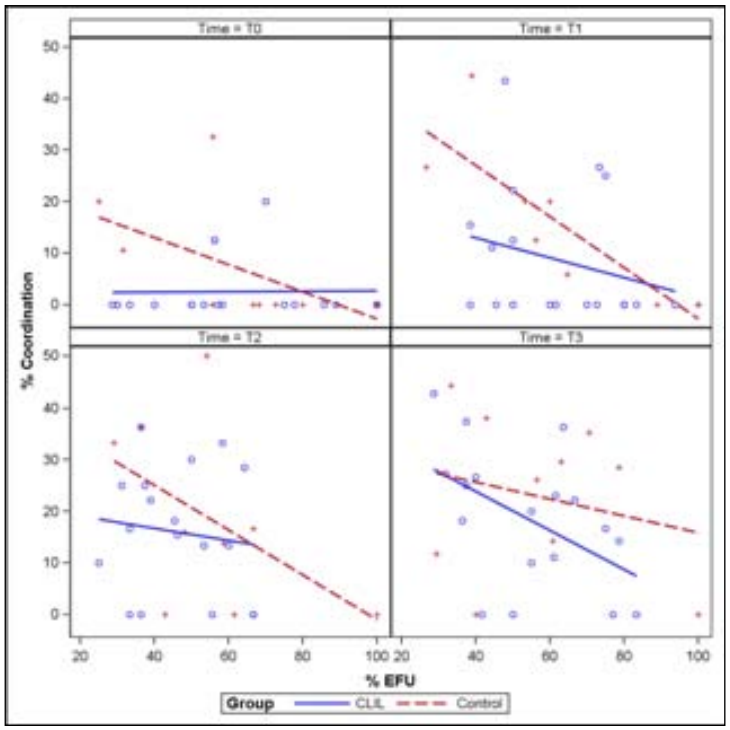
6.1.7.1 Accuracy (% EFU) and complexity (% CU)

This section will provide the graphs and results obtained from the correlation tests between accuracy and complexity using the scores learners achieved in % EFU and % CU at T0, T1, T2 and T3. The results will be presented for both CLIL and control groups.

Figure 6.35 shows the relationship between accuracy and complexity at the four times of data collection in CLIL + EFL settings (CLIL group) and in EFL settings (control group). As seen from the graphs, the results obtained by the control group in % EFU and % CU reflect a clearer negative relationship between the variables than the results from the CLIL group. The scores obtained by the control group in the two measures reveal that learners who produce the highest percentages of coordination also

produce the lowest percentages of error-free units at the four times of data collection. The results obtained by the CLIL group, on the other hand, show that at T0 learners do not produce any coordinate units at all, whereas their scores in accuracy vary from 0% EFU to 100% EFU. At T1, T2 and T3, the situation changes, as the results from the CLIL group start to resemble those obtained by the control group. That is to say, the results indicate that, in general terms, CLIL learners whose percentages in coordination are high also show low percentages in accuracy. As for the statistical analysis, the results from the correlation test show a negative correlation between the two measures for both groups. However, in the case of the CLIL group the strength of the correlation is weak (r(30)=-0.29, p =0.0121) whereas the results from the control group show a strong negative correlation (r(18)=-0.56, p=0.0002) between accuracy and complexity.

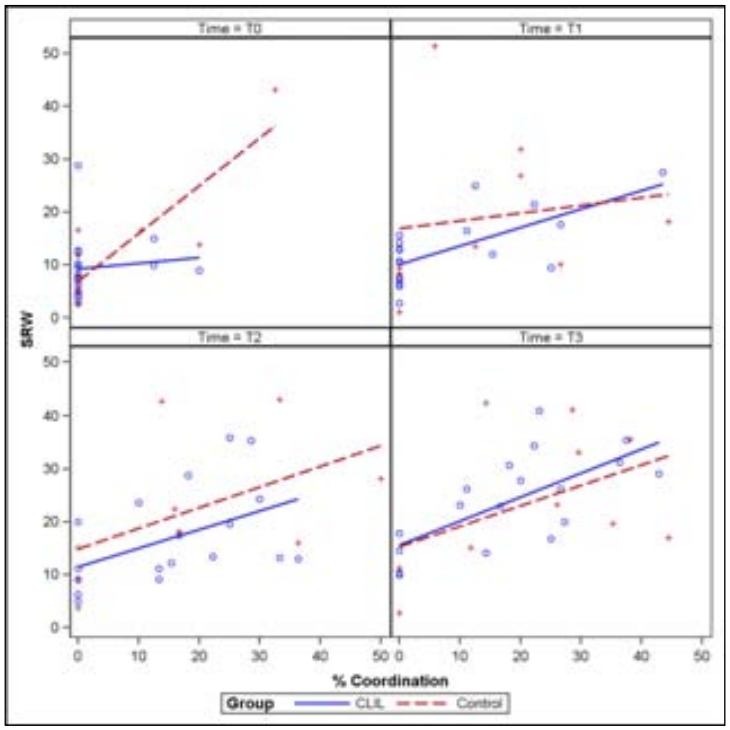
Figure 6.35
Correlations between accuracy and complexity in the interview task.



6.1.7.2 Complexity (% CU) and fluency (SPRW)

The results of the correlation test between complexity and fluency are presented in Figure 6.36. The relationship between the two measures is positive for the two groups. At T0, however, results show that the participants from both the CLIL and the control groups hardly produce any coordinate units and their fluency scores are pretty low. The situation changes at T1, when the results start to suggest that fluency and coordination scores increase simultaneously. Finally, the results from T2 and T3 corroborate the tendency found at T1 as learners who show the highest scores in fluency also obtain high percentages in coordination in general. As for the statistical results obtained, the tests conducted indicate that there is a strong positive correlation between complexity and fluency for the CLIL ($r(30)=0.67$, $p< .0001$) and control groups ($r(18)=0.71$, $p<.0001$).

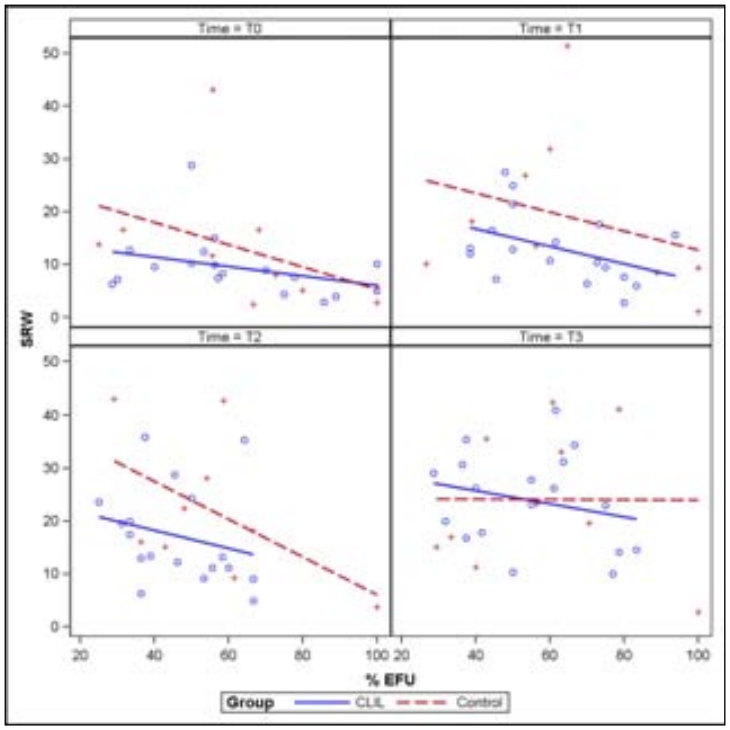
Figure 6.36
Correlations between complexity and fluency in the interview task.



6.1.7.3 Accuracy (% EFU) and fluency (SPRW)

The results obtained from the correlation tests between accuracy and fluency in the interview task are presented in Figure 6.37. As in the previous correlation test between accuracy and complexity, a negative correlation is also observed between accuracy and fluency. In this correlation, both groups display a negative relationship between the variables at T0 since the learners whose fluency scores are high have the lowest scores in accuracy generally. This tendency is still visible at T1 and T2 for both groups. At T3, however, the trend changes slightly as learners from the control group start to obtain relatively high scores in fluency and accuracy simultaneously. The results from the CLIL group, on the other hand, continue to reflect a clear negative correlation between the two measures. Regarding the strength of the correlations, the results indicate that a medium negative correlation was found between accuracy and fluency in the CLIL ($r(30)=-0.37$, $p=0.0015$) and control groups ($r(18)=-0.41$, $p=0.0097$).

Figure 6.37
Correlations between accuracy and fluency in the interview task.



To sum up, the Spearman's rho correlation tests reveal that there is a strong positive correlation between complexity and fluency ($p<.0001$) regardless of the instructional context (CLIL+EFL versus EFL). The correlation between accuracy and complexity, on the other hand, does seem to vary according to instruction type. In the case of the control group a strong negative correlation between the measures was found ($p=0.0002$), while in the CLIL group the results indicate that the strength of the negative correlation is rather weak ($p=0.0121$). As for the relationship between accuracy and fluency a medium negative correlation was found for both the CLIL ($p=0.0015$) and control ($p=0.0097$) groups.

6.2 Qualitative analysis

In this section, a description of the learners' output and evolution will be provided from a more qualitative perspective at an individual level with the aim of highlighting relevant aspects of language development which have been inevitably overlooked in the previous quantitative statistical analyses. As Larsen-Freeman (2006: 612) states, " micro-level description of the individual's development is not always addressed". According to her, individual descriptions are also needed to understand the mechanisms behind language acquisition. Thus, in an attempt to provide detailed descriptions of language development, the following subsections will focus on the oral output of six learners with different fluency levels in the interview and narrative task at the four times of data collection (three CLIL learners and three control learners).

6.2.1 CAF evolution: group and individual trends

The qualitative analyses presented in the following subsections will provide a description of the evolution of the most representative CAF measures used in this

dissertation in the CLIL and control groups in the interview and narrative tasks. The measures selected to carry out this description are percentage of coordinate units (% CU) to analyse complexity, percentage of error-free units (% EFU) to examine accuracy and speech rate in words (SPRW) to analyse fluency. The mean scores obtained by the CLIL and control groups in these measures in the interview and narrative tasks will be plotted in graphs for the four data collection times in an attempt to offer a complete picture of the evolution and relationship of CAF in the two instructional settings.

The criteria used for the selection of these measures were related to the type of data that young learners produced and the evolution of the measures throughout the two years of the study. In the case of complexity, for instance, coordination was thought to be more representative of the learners' syntactic complexity than subordination due to the fact that the instances of subordination found in the samples were too limited. Coordinate units, on the other hand, were much more common at the four times of data collection. As for accuracy, the reason why error-free units was chosen over percentage of correct verb forms was related to the regularity this measure showed over the two years of the study. Finally, the fluency measure selected, speech rate in words, was considered to be more complete than the percentage of L1 words (% L1) due to the fact that it takes into account both the total number of words in English (TNWE) and total task time (TTT).

In addition to the description of the evolution of CAF measures in the two settings, this section will also present the individual evolution of CAF of a highly fluent CLIL learner, a highly fluent control learner, a fluent CLIL learner, a fluent control learner, a dysfluent CLIL learner and a dysfluent control learner in the two tasks. In the case of the interview task, a selection of questions and their corresponding answers has been made. The criterion used to classify the learners according to their fluency level

was their scores in speech rate in words at T0 in the narrative task. The CLIL and control learners who obtained the highest scores in speech rate in words are labelled as 'highly fluent'. Those learners whose scores were the lowest will be referred to as 'dysfluent learners'. Finally, those learners whose scores in speech rate in words were the closest to the average group score are named 'fluent learners'. The analysis of the individual evolutions will start with a description of the development of the CAF measures selected and will be complemented with a thorough account of the development of the other CAF measures used in the quantitative data analysis (% VR, % ADJR, % SU and % L1WR) and relevant language features produced by the learners (i.e. use and evolution of formulaic language, attempts to produce subordination or question formation among others). Besides these descriptions, the oral output produced by the learners in the narrative task will also be examined using Álvarez's (2006) description of narrative stages.

6.2.1.1 Interview task

The evolution of CAF of CLIL and control learners in the interview task is shown in Figures 6.38 and 6.39. As seen from the two graphs, the evolution of CAF is very similar between the two groups. Complexity and fluency seem to progress simultaneously while accuracy decreases from T0 to T3. The only minor difference between the two groups is that while complexity and fluency increase steadily from T0 to T3 in the CLIL group, in the control group complexity and fluency have a tendency to increase gradually from T0 to T2. From T2 to T3, on the other hand, complexity and fluency show a slight decrease.

Figure 6.38
Evolution of CAF in the CLIL group. Interview task.

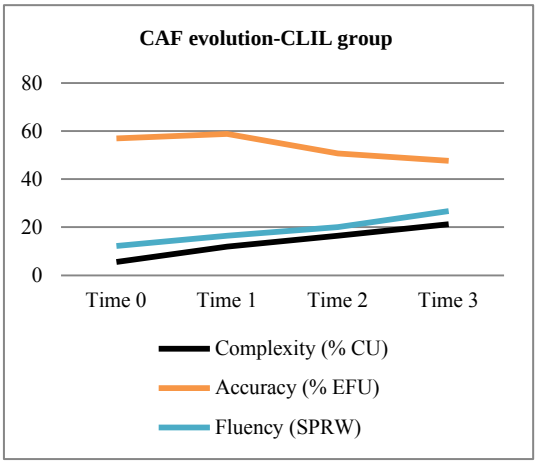
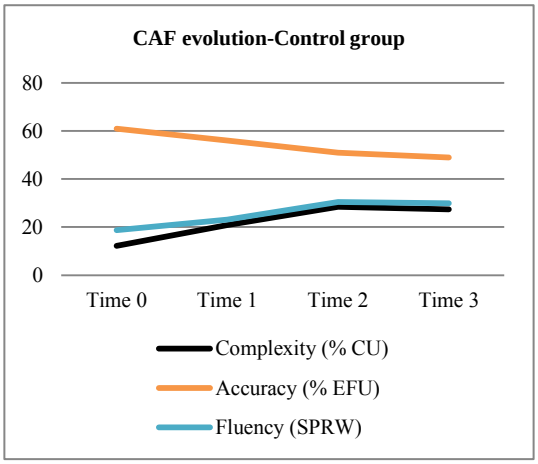


Figure 6.39
Evolution of CAF in the control group. Interview task.



6.2.1.1.1 Highly fluent learners from the CLIL and control groups

The evolution of CAF of the two highly fluent learners selected (student A and B) is displayed in Figures 6.40 and 6.41 below. Additionally, Tables 6.144 and 6.145 display the answers provided by these learners to five questions from the interview task at the four times of data collection (see Appendix C for the complete transcipts). The coding symbols used in the tables and transcripts are explained in Appendix D.

Figure 6.40
Evolution of CAF in the interview task. Highly fluent CLIL learner (student A).

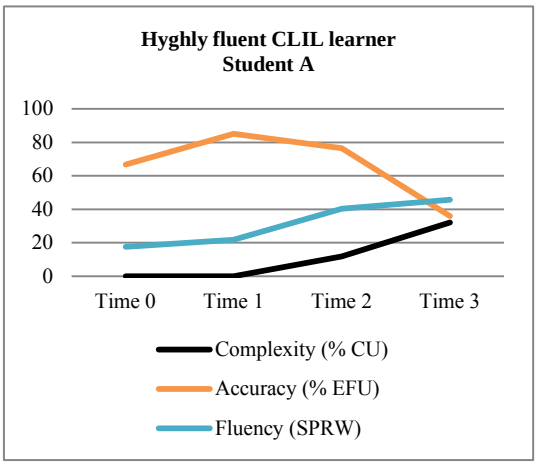


Figure 6.41
Evolution of CAF in the interview task. Highly fluent control learner (student B).

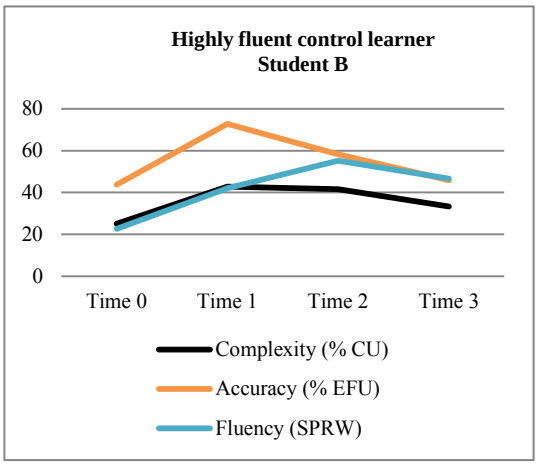


Table 6.144
Highly fluent CLIL learner. Student A's answers in the interview task.

Questions from the interview	**T0**	**T1**	**T2**	**T3**
1.How old are you?	Ten years old.	Eleven.	I'm el even years old.	I ha ve t welve years o ld an d I have t he b irthday last month.
2.Tell me about your family? **Any brothers or sisters?**	One sister and one father an d o ne mother.	One sister and one father an d o ne mother.	I ha ve two brothers [//] bueno@s:c o ne sister a nd o ne brother, one father one mother a nd six more parents@s:c b ut I don't k now t he name of tiet@s:c.	I have m y f ather works in a f actory and he had I think so he had fifty [//] forty-nine y ears old. And my mother works i n a s chool is a maths i s a prof@s:c [//] sorry teacher of maths.
3.What are your parents' jobs?	[2nd attempt] My f ather i s a teacher. My father no@s:c s é@s:c com@s:c es @s:c diu@s:c.	Teacher.	My f ather i s a boss of empresa@s:c a nd my mother i s a teacher of maths.	

4.What do you usually do at the weekends	No.	In Saturday I go to play football a t eleven o'clock and the other time no I +....	I d on't kno w. I t's different. E very day is d ifferent. Sometimes I go to the mountain t o run bicycle.	I p lay football with my t eam o r I do my homework. Nothing more.
5.Now it's your turn ask me a few questions.	How old are you? And your family? Who i s t he j ob of your fathers?	How old are you? What is your job? And your family? Sisters brothers? How years h ave he?	What a bout your parents? And your job? What's t he j ob of your parents?	Which i s your work? And a bout y our family, do you have some o ther teachers i n your family? Is i t b ecause you like teachers work or because your fathers te ll you d o teachers work?

Table 6.145
Highly fluent control learner. Student B's answers in the interview task.

Questions from the interview	**T0**	**T1**	**T2**	**T3**
1.How old are you?	I ten years old.	I am ten years old.	I'm el even years old.	I a m t welve years old.
2.Tell me about your family? **Any brothers or sisters?**	My mum his name is … a nd my father … Yes, i t o ne si ster and his name is …	My father i s … and he ha ve forty five years o ld an d I ha ve my grandfathers and grandmothers, aunts and uncles. Yes, I h ave a sister. His name is … a nd he [//] s he have s even y ears and my mo ther is… a nd s he have forty years old.	My f ather ar e … and he works in a hospital a nd my mother is a dentist and his name is ... And I ha ve o ne sister t hat his name is [//] she's name is …	My father is a [//] he works i n t he hospital a nd his name i s … , … i n English. And my mother is a d octor. A ll my family doctors and his name is …. And I have a small sister y ounger than me. His name is … and he do third course.
3.What are your parents' jobs?	No an swer from the participant.	My mother's work in a d entist a nd my father's i s a doctor.		
4.What do you usually do at the weekends	I do [//] I r ide the horse a nd do t o the horse and nothing.	I go to ride a horse and we went to … with a c ousins. We stay in his house.	[2nd attempt] Normally we went to t he house o f … This is a farm there ar e h orses and we +…	In Saturday I go to ride my h orse and Sunday I do what I want.

5.Now it's your turn ask me a few questions.	What i s your name? What i s your favourite subject? What i s your mother name's? How old are she? What i s your favourite sport?	How are you? How old are you? What i s your favourite sport? When you born? Today y ou go to more schools or +…? What i s your favourite animal?	How old are you? What i s your favourite colour? What i s your favourite sport?	How old are you? What i s your favourite sport? Interesting, yo u relax? You go t o different s chools to do this? Where you live?

According to the graphs, the most remarkable difference in terms of syntactic complexity between the two participants is that student A, the highly fluent CLIL learner, does not produce coordination until T2 while his peer in the control group, student B, already does so at T0. As seen from Table 6.144, however, student A produced a to-infinitive subordinate clause already at T1 'In Saturday I go to play football a t e leven o 'clock'. Additionally, the learner attempts to complement that sentence by adding ' and the other time no I +…'. Even though the learner fails to produce a coordinate unit here, an attempt is made to produce coordination. At T2, the instance of coordination found is in the answer given to question 2 'I have two brothers [//] *bueno* a sister and a brother, one father, one mother and six more *parents* but I don't know the name of *tiet*'. The coordinate unit produced here ('but I don't know the name of *tiet*') lets the researcher know that the learner would like to name more relatives but he does not know what *tiet*, the Catalan word for uncle, is in English. Besides, student A also uses coordination to talk about his parents' jobs: 'My father is a boss of *empresa* and my mother is a teacher of maths'. In addition to coordination, student A's output also be comes m ore c omplex t hrough t he us e of a dverbial s ubordination a t T 2 'Sometimes I go to the mountain to run bicycle'. Even though the utterance contains lexical and grammatical errors, the learner succeeds in producing a subordinate clause to indicate purpose. At T3, as displayed in Figure 6.40 and Table 6.144, student A's speech becomes even more complex due to the increase in the number of coordinate and

subordinate units. The answer student A gives to question 2, for instance, exemplifies that increase in syntactic complexity: 'I have my father works in a factory and he had I think so he had fifty [//] forty-nine years old'. As seen from this utterance, student A provides the interlocutor with a greater amount of information than in previous data collection times which in turn leads to the production of more elaborate output. Firstly, the learner is clearly using coordination. In addition, it seems as if the learner is attempting to produce a relative clause ('I have a father *[who]* works…') to join clauses. Finally, the learner also produces subordination within the second coordinate unit by using 'I think so'. Despite the fact that the learner uses 'I think so' as an unanalysed chunk when he intends to use *I think*, the learner attempts to produce a nominal subordinate that-clause. Besides the use of coordination and subordination in student A's answer to question 2 at T3, other instances of coordinate and subordinate units are found in the other student's answers. Student A's answers to question 1 and 3, for example, also consist of two coordinate units 'I have twelve years old and I have the birthday last month' and 'I play football with my team or I do my homework'. Finally, another example of the degree of syntactic complexity produced by student A at T3 is seen in one of the questions the learner asks the researcher ' Is it because you like teachers work or because your fathers tell you do teachers work?'. In this utterance, the learner has produced two coordinate units and three subordinate units (two with the subordinating conjunction *because* and an attempt to produce a nominal to-infinitive clause). All in all, the speech samples displayed in Table 6.14 4 corroborate and complement the increase in complexity shown in Figure 6.40 from T2 to T3.

In regard to student B (the highly fluent control learner), as mentioned earlier and as shown in Figure 6.41, the learner already produces coordination at T0. From T0 to T1 there is an increase in coordination which is followed by a slight decrease from T1

to T3. The instances of coordination produced at T0, which are displayed in Table 6.145, are 'My mum his name is and my dad…' and 'It one sister and his name is ...'. In spite of the formulaic use of 'it' at the beginning of the sentence instead of *I have* and other grammatical errors, student B succeeds in joining units by means of coordination. At T1, this same answer evolves into 'I have a sister. His name is…and he [//] she have seven years old'. The coordination pattern found in this utterance (name of the relative and age) is repeated in two more answers at T1 'My father is …and he have forty five years old' and 'My mother is…and she have forty years old'. Additionally, student B at T2 also produces coordination when talking about his parents' jobs: 'My mother's work in a dentist and my father's is a doctor' while at T0 the learner was unable to answer the question about his parents' jobs. Moreover, the learner also produces coordination in his answer to the question related to weekend activities by saying 'I go to ride a horse and we went to…'. As seen from this description the high number of coordinate units found at T1 explains the sharp increase described in the graph from T0 to T1. According to Figure 6.41, the use of coordinate units decreases slightly from T1 onwards. However, that does not negatively affect the learners' syntactic complexity, since at T2 and T3, instances of the emergence of subordination are detected in Table 6.145 through attempts to produce relative clauses, such as 'And I have one sister that his name is [//] she's name is…' and ' This is a farm there are horses and we +… ', which increase undoubtedly the degree of syntactic complexity unfolded by the learner. In the first example, student B is trying to produce a relative clause with *that* and in the second example the learners fails to use the relative pronoun *where* to join the two units. Sentence word order as well as the lack of pauses between the two units seem to indicate that the learner is trying to produce a relative clause. At T3, the percentage of coordinate units continues decreasing slightly at the same time as attempts to produce

subordinate units are maintained from T2. Another example of student B's attempt to produce a relative clause at T3 is: 'And I have a small sister younger than me'. In this utterance, the learner failed to use a relative pronoun (*who*), however, the structure of the sentence seems to reveal the learner's attempt to produce a relative clause. In addition to the use of coordination and the attempts to produce relative clauses, it is worth highlighting that student B also produces other types of complex sentences. At T1, for example, the learner produces a to-infinitive embedded clause, as in 'I go to ride a horse'. A similar version of this utterance is also used at T3 in 'In Saturday I go to ride my horse', which is followed by a coordinate unit which contains a nominal relative subordinate clause 'and Sunday I do what I want'.

As for lexical diversity, the subdimension of lexical complexity examined in this study, Table 6.144 shows an increase in the number of verbs from T2 onwards in student A's output. According to the language samples in the table, at T0 the only verb used by student A is the verb 'be' in sentences like 'My father is a teacher', 'How old are you?' and 'Who is the job of your fathers?'. At T1, the learner also uses the verb 'go', 'play' and 'have' in the following utterances: 'In Saturday I go to play football' and 'How years have he?'. At T2, the verbs used are the same as the ones used at T0 and T1, however, the learner uses them in a wider variety of contexts. For example, at T2 the learner uses the verb 'be' for the first time to talk about his age 'I'm twelve years old'. In addition, it is also the first time that the learner uses 'have' to talk about his family 'I have two brothers [//] *bueno* one sister and one brother'. At T3, the learner uses the following verbs: 'have', 'work', 'play', 'do', 'like' and 'tell'. One of the most relevant aspects to highlight is the use of 'tell' and the learner's attempt to use it appropriately in the coordinate unit 'or because your fathers tell you do teachers work?'. Concerning the use of adjectives, at T0 the only adjective used in the samples provided

in Table 5.144 is 'old' in the utterance 'Ten years old' and the formulaic question 'How old are you?'. The reason why this question is believed to be formulaic at T0 is because at T1 the learner uses both 'How old are you?' and 'How years have he?', which indicates that as the learners' linguistic knowledge develops and increases, he starts to combine already learnt words or the structure of already learnt chunks to formulate questions ('How years have he?') in addition to the use of unanalysed chunks of language ('How old are you?'). At T2, the learner also uses the adjective 'different' in the utterance 'It's different. Every day is different'. At T3, the two adjectives used by the learner are 'old' and 'last' in the sentence 'I have twelve years old and I have the birthday last month'.

Student B's lexical diversity also increases from T2 onwards in terms of the number of adjectives and verbs used. As seen from Table 6.145, at T0 the only verbs used are 'be' and 'ride'. The former is used in sentences that describe the learner's family, as in 'My mum his name is …' or 'It one sister and his name is…' and in all the questions asked to the researcher, as in 'What is your name?', 'What is your favourite subject?', 'What is your mother name's?' and 'How old are she?'. The latter is used in the sentence 'I ride the horse', which is produced at three data collection times. At T1, the number of verbs used by the learner increases. The learner introduces the verb 'be' to talk about his age 'I am ten years old'. Besides, the verb 'have' is used for the first time in sentences like 'He have forty-five years old' or 'I have a sister'. At T0, this same learner failed to use the verb 'have' when talking about his family 'It one sister'. In addition to the verb 'have', student B also uses the verbs 'go' and 'stay' at T1. Interestingly and as a further instance of increased lexical complexity, the learner uses the past and present tense of 'go' in the following utterances: 'I go to ride a horse and we went to…' and 'Today you go to more schools?'. Additionally, as already

mentioned, student B uses the verb 'stay' in 'We stay in his house, which shows the growth of the learner's lexicon. At T2, in addition to 'be', 'have' and 'go', the learner also produces the verb 'work' instead of 'be' to talk about his parents' jobs 'My fathers are…and he works in a hospital'. From this time onwards, student B uses the verb 'work' to talk about his parents' jobs. Finally, at T3 the learner adds the verbs 'do', 'relax' and 'live' to his verb repertoire. The verb 'do' is used in two utterances 'He do third course' meaning that his sister is in year 3. The use of 'do' in this context is clearly an instance of L1 transfer as this would be the verb used in Catalan. Despite this, the learner manages to use the verb in the right context in 'I do what I want' and 'You go to different schools to do this?'. As for the verbs 'go' and 'live', they are used when the learner asks the researcher some questions, as in 'You go to different schools to do this?' and 'Where you live?'. In regard to the use of adjectives, Table 6.145 shows that the only adjectives used by the learner at T0, T1 and T2 are 'old' and 'favourite' in formulaic utterances like 'I ten years old', 'I am eleven years old', 'What is your favourite subject', 'What is your favourite sport?' or 'What is your favourite colour?'. At T3, the learner also uses 'small', 'younger', 'interesting' and 'different' in addition to 'old' and 'favourite'. One of the most relevant aspects worth commenting on regarding the use of adjectives by student B is that the adjectives used from T0 to T2 were included in formulaic language while at T3 the learner starts to use a wider variety of adjectives in different contexts indicating that the use of adjectives is no longer associated with formulaic language.

As for accuracy, Figures 6.40 and 6.41 also show the evolution of the percentage of error-free units for the highly fluent CLIL learner, student A, and the highly fluent control learner, student B, in the interview task. Student A's evolution of error-free units displayed in Figure 6.40 indicates a slight increase from T0 to T1 followed by a modest

decrease from T1 to T2 and a substantial decrease from T2 to T3. As for student B's evolution of the accuracy measure, Figure 6.41 displays a sharp increase from T0 to T1 and a gradual decrease from T1 to T3 in the percentage of error-free units. At this point of the qualitative analysis, it is worth pointing out that the percentage of error-free units produced by student A decreases as the percentage of coordinate units increases up to the point that at T3, student A displays the highest score in complexity and the lowest score in accuracy. As for the relationship between the percentages of coordinate units and error-free units obtained by student B, Figure 6.41 shows that after reaching the highest peaks in these two measures at T1, both measures seem to decrease simultaneously. The descriptions of the language produced provided below will attempt to shed some light into the relationship between complexity and accuracy. Table 6.144, for example, displays the answers provided by student A in the interview task. At T0 the learner produces five error-free units out of the six units shown in the table. The only unit which contains errors is 'Who is the job of your family?', in which the learner fails to use the right wh-question word. Despite the high percentage of error-free units produced by student A at T0, the language employed by the learner is quite simple, which may prevent the learner from making mistakes. In addition to the use of simple language, the use of formulaic language is also decisive in the production of correct output. At T1, the tendency to produce simple language is still visible, for example, 'Eleven' as the answer to question 1 or 'One sister, one father and one mother' as the answer to question 2. Despite this, some traces of complex language use, which affect accuracy, are found at T1. For instance, at this time, the learner constructs the following question: 'How years have he?', in which, despite not using an auxiliary verb, he uses the verb 'have' and subject-verb inversion, which has been proved to be difficult to master by young language learners. At T2, the first instance of self-correction is

detected in the utterance 'I have two brothers [//] *bueno* a brother and a sister' when the learner corrects his use of 'brothers'. At T3, the highest number of errors is found along with the most complex language structures. To start with, morphological errors related to verb tenses suggest the evolution of the learner's L2 morphological system. These are found in the following utterances: 'I have my birthday last month', 'He had I think so he had fifty [//] forty-nine years old'. In other words, as the language becomes more complex, more errors are found. Besides morphological errors, student A also uses the verb 'have' to indicate his age 'I have twelve years' at T3, when this same message was 'I'm eleven years old' at T2. Again, this seems to be another example of the increase in errors as the learner's L2 system becomes more complex. At this point, it would be relevant to comment on the lack of errors as well. For instance, student A succeeds in marking the verb 'work' inflectionally for third person singular present tense subject agreement on two occasions, namely 'My father works in a factory' and 'My mother works in a school'. Likewise, student A succeeds in forming questions which require subject-verb inversion in the following utterances 'Is it because you like teachers work or because…?'. As mentioned earlier, traces of subject-verb inversion in questions were already detected at T1, which might also be an example of formulaic language use, as learners memorise chunks like 'Is it' to start questions which use the verb 'be'. Finally, at T3 student A also uses the auxiliary verb 'do' in the production of questions for the first time in 'Do you have some other teachers in your family?', which clearly shows increased complexity and accuracy in the learner's interlanguage system.

As shown in Table 6.145, which displays the language produced by student B in the interview task, the errors found at T0 are related to the lack of predicates as in the utterances 'I ten years old', which from T1 onwards is produced correctly as 'I am ten years old', or 'It one sister', which is also produced correctly from T1 onwards as 'I

have one sister'. The correction of errors at T1 may explain the increase in the percentage of error-free units found from T0 to T1 shown in Figure 6.41. Besides the lack of predicates, errors associated with the use of formulaic language are also common, especially, in questions. At T0, for example, student B asks the researcher 'How old are she?' to refer to the age of the researcher's mother. The learner is forming a question using the unanalysed chunk 'how old are' and adding 'she' at the end. Similarly, the use of formulaic language is also found in questions such as 'What is your favourite sport/colour/subject?', 'How are you?' or 'How old are you?', which are used by the learner all through the four times of data collection. In this case, the use of formulaic language increases the percentage of error-free units at T1 and accounts for it. In addition, wrong verb selection is also present at T1 when student B talks about his parents' and sister's age using the verb 'have' in the utterances 'He have forty five years old' or 'He [//] she have seven years old'. However, despite the use of the wrong verb, the learner corrects himself in the use of the subject pronoun in the second utterance. As for the production of questions, new and more elaborate attempts to make questions are found at T1 when the learner asks 'Today you go to other schools?'. Even though student B does not include the auxiliary verb *do,* he is trying to go beyond the formulaic questions asked at T0 and create his own. Despite his attempts, no instances of correct yes-no questions are found at any of the times. As for wh-questions, the learner does show mastery of wh-word questions in terms of meaning, even though most wh-questions seem to contain unanalysed chunks of language. Another common error found at all four times of data collection is the use of the masculine possessive determiner instead of its feminine counterpart, as in 'I have a sister. His name is…', 'My mother is a dentist, his name is …'. At T2, however, the learner seems to be aware that there needs to be gender agreement when he (incorrectly) corrects himself 'I have one sister

that his name is [//] she's name is…'. At T2, errors related to verb tense are also present: 'Normally we went to the house of …'. This error has already been detected at T1 with the same verb, as in 'I go to ride a horse and we went to…'. Finally, the absence of errors is also worth commenting on. For example, the learner marks the verb *work* for third person singular subject verb agreement in the answer 'He works in the hospital' at T3, which could be explained in terms of formulaic language as well, as the learner might have learnt 'He works' as a chunk of language. Additionally, student B also succeeds in using a comparative adjective correctly: 'I have a small sister younger than me'.

In regard to fluency, both learners improve from T0 to T3. Student A shows improvement throughout the whole study. However, the period of time when he improves the most is from T1 to T3. Student B, on the other hand, shows a gradual increase from T0 to T2 followed by a slight decrease from T2 to T3. Interestingly, Student A's fluency and complexity scores display simultaneous improvement from T1 to T3, while student B's fluency and complexity scores show a simultaneous and slight decrease from T2 to T3. Despite this, the language samples displayed in Table 6.144 and 6.145 clearly show an increase in the amount of language produced by both students throughout the four data collection times. Student A's answers in Table 6.144 are quite simple and short at T0. The researcher has to ask one of the questions twice in order for the learner to provide an answer, which means that the total task time of the task increases while the language produced by the learner is kept the same, which in turn results in lower scores in speech rates in words. At T0, one instance of L1 use can also be seen: '*no sé com es diu*' (I don't know what this is called). The function of the L1 in this case is to indicate that the learner does not know a word in English and hence L1 use generates metatalk. At T1, the utterances are longer and more language is

produced by the learner. As for L1 use, no instances of L1 words are found. At T2, the length and number of units continue increasing substantially. Interestingly, four instances of L1 use are found. The first one, '*bueno*' (well), is placed right before a self-correction and denotes a false start. The following two, namely '*parents*' (relatives) and '*tiet*' (uncle) are used in the answer to the question related to the learner's family. What is relevant here is that the learner uses the expression 'I don't know the name of *tiet*' (meaning I don't know what *tiet* is called) instead of saying the whole utterance in the L1, as he did at T0 '*no sé com es diu*' (I do not know what it is called). This time the learner uses English to a greater extent and limits the use of the L1 to isolated words he does not know in English, as in the last example of L1 use, '*empresa*' (company) found at T2. Again, in this case L1 use does not prevent the learner from continuing his speech in English as the following utterance shows: 'My father is a boss of a *empresa* and my mother is a teacher of maths'. At T3, the only instance of L1 use is found right before a self-correction as well: 'My mother works in a school is a maths *prof* [//] sorry teacher of maths'. In this example the learner is about to say the word *professora* (teacher), however, he corrects himself and says it in English. As for the amount of language produced, T3 is the time when student A produces the highest amount of speech.

The language sample displayed in Table 6.145 also indicates that the amount of language produced by student B increases over the four data collection times. At T0, the answers are short and simple. Furthermore, the learner does not answer one of the questions. At T1, the utterances are visibly longer. In addition, the learner answers all the questions. At T2 and T3, the tendency to produce long utterances is maintained. At T2, however, the researcher has to ask a question twice in order to get an answer from the student. As for the use of the learner's L1, no instances of L1 words are found in the sample provided.

To sum up, the data displayed in Figures 6.40 and 6.41 and Tables 6.144 and 6.145 show an increase in complexity and fluency from T0 to T3 for both the CLIL and control learners. Additionally, the first attempts to produce subordination are observed at T1. The highly fluent CLIL learner, however, shows increased complexity from T2 to T3, which is observed in the use of syntactically complex structures in questions, while the highly fluent control learner shows a decrease in the percentage of coordinate units from T1 onwards, probably due to the emergence of subordination. According to the samples, there seems to be a negative relationship between complexity and accuracy, as accuracy levels decrease at the same time as complexity increases. In addition, the use of formulaic language chunks at T1 and T2 also seems to affect the learners' accuracy levels. Regarding lexical diversity, both learners show an increase in the number of different verbs throughout the four data collection times, while the use of adjectives is quite limited. As for fluency, both learners show improvement in terms of the number of words used and decrease in L1 words.

6.2.1.1.2 Fluent learners from the CLIL and control groups

Figures 6.42 and 6.43 display the evolution of CAF of the two fluent learners selected. As seen from the graphs, fluency and complexity scores develop similarly for the two learners. However, one of the differences in terms of complexity between the CLIL and control fluent learners is that student C (fluent CLIL learner) did not produce coordination until T1, while student D (fluent control learner) already produced coordination at T0. Despite this difference, both students achieve similar scores in complexity at T3. The development of accuracy differs between the two students. Both learners display similar scores in the percentage of error-free units at T0, however, while student C improves slightly from T0 to T1, student D shows a sharp decrease.

From T1 to T3, student C decreases considerably, whereas student D obtains similar scores in accuracy at T1, T2 and T3. Eventually, the two students have similar levels in accuracy at T3. In regard to the development of fluency, both learners show improvement in general terms. At T3, student D's fluency score is slightly above student C's score. Tables 6.146 and 6.147 display the output produced by the two stuents in the interview task.

Figure 6.42
Evolution of CAF in the interview task. Fluent CLIL learner (student C).

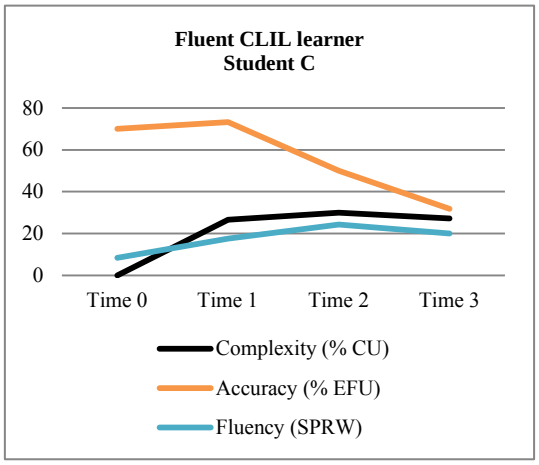


Figure 6.43
Evolution of CAF in the interview task. Fluent control learner (student D).

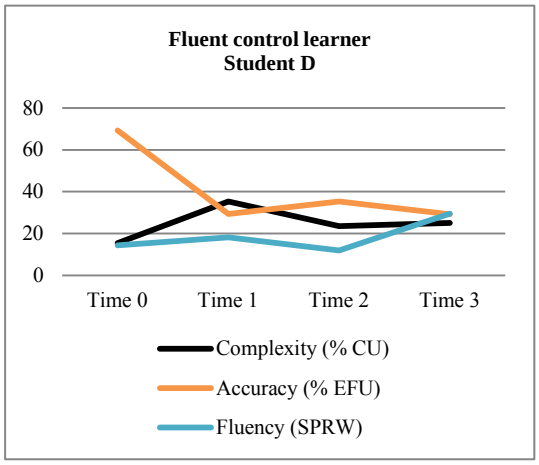


Table 6.146
Fluent CLIL learner. Student C's answers in the interview task.

Questions from the interview	**T0**	**T1**	**T2**	**T3**
1.How old are you?	I ten years old.	Eleven.	I'm eleven years.	I'm twelve years old.
2.Tell me about your family? **Any brothers or sisters?**	[2nd attempt] My father is thirteen years old.	[2nd attempt] The name of my mum is … and my dad …	The name of my dad is … the name of my mum is … My brother is … and my dog is …	My family is good. Yes, one brother. Has got [//] is fifteen years old.

3.What are your parents' jobs?	No an swer from the participant.	Work i n t he ferreteria@s:c	Has a ferreteria@s:c an d my dad també@s:c [//] h e is a ferreter@s:c.	My parents j ob i s iron monger.
4.What do you usually do at the weekends	No an swer from the participant.	[5th attempt] I'm p laying t he video consola@s:c in my house.	I'm play the +…	On S aturdays I 'm doing t he homework a nd I'm p laying t he playstation.
5.Now it's your turn ask me a few questions.	How years old? Where ar e y ou live? Where ar e y our profession?	Where [//] What is your name? What's ye ars yo u have? What's t he name of your mum? And your dad?	Has got years? What i s t he name of your dad? And your mum? You has go t a brother? What's your name? (Intended form: his name) Has got a dog? You p ractise sport?	What doing in the weekend? What w ork y our mother a nd your dad? You have a sister?

Table 6.147
Fluent control learner. Student D's answers in the interview task.

Questions from the interview	**T0**	**T1**	**T2**	**T3**
1.How old are you?	Ten years old.	I am eleven years.	Eleven.	Twelve years old.
2.Tell me about your family? **Any brothers or sisters?**	Three.	Yes, I t hree brothers.	I ha ve t hree brothers.	My mother he rs name i s … a nd work i n t he restaurant an d my father in the restaurant. Yes, t hree brothers.
3.What are your parents' jobs?	Secretary a nd cook.	My mother secretary and my father cooking.	My father co oker and my mo ther secretary.	
4.What do you usually do at the weekends	[2nd attempt] I go to skating.	On S aturday I a m going o f s kating and house.	[2nd attempt] Sunday I go i n skating.	Saturday in morning I go t o skating a nd afternoon go t o a beach and water. Sunday family a house.
5.Now it's your turn ask me a few questions.	How are you? What i s your favourite colour? You I ha ve animals? What i s your name? (Intended form: your dog 's name)	What old are you? What is the j ob of you? And t he j ob of your parents?	You have a brothers? What i s your parents' job?	How old are you? Your parents were jobs? And any brothers? And sisters? Where i s your job?

Table 6.146 displays the answers given by student C in the interview task. In terms of complexity, no instances of coordination are found at T0 while at T1 two coordinate units are shown in the sentence 'The name of my mum is…and my dad…'. Even though the second coordinate unit does not contain a verb (most probably due to the use of ellipsis in oral interactions), student C does use the coordinating conjunction 'and'. At T2, two more instances of coordination are used by the learner. The first one, 'The name of my mum is …and the name of my dad is…', is a more elaborate version of the coordinate units found at T1, as the verb is present in both units. The second instance is 'My brother is…and my dog is…'. At T3, the coordinate units produced by the learner are of a different type. So far, coordination has been used to talk about his family's names. At T3, in contrast, coordination is used to describe the activities student C does at the weekend: 'On Saturdays I'm doing the homework and I'm playing the playstation'. At T3 coordinate units no longer contain the verb *be*.

Student D, unlike student C, already uses coordination according to Figure 6.43. at T0. Table 6.147, however, does not show the instances of coordination produced by this learner at this time, as this qualitative analysis does not include all the questions asked by the interviewer. Instead, it includes those questions which generally generate more relevant language features. The first instance of coordination produced displayed in the table is ' My mother secretary and my father cooking' at T1. In spite of the absence of predicates in these coordinate units, the learner shows an increase in complexity and an attempt to reproduce SVO order[6] when this utterance is compared to the answer provided at T0: 'secretary and cook'. At T1, the learner not only includes the possessive determiner but also the coordinating conjunction in an attempt to produce a complete sentence. Additionally, an attempt to produce a subordinate unit is also found

[6] The term SVO refers to the structure of a sentence (subject-verb-object) in this part of the study, not to transitive verb types.

at T1: 'On Saturday I am going of skating'. This utterance evolves into 'I go in skating' at T2 and 'I go to skating' at T3. As seen from the utterances, the learner is trying to produce subordination in spite of the fact that he is struggling with the use of prepositions, the to-infinitive marker and non-finite verb forms. At T3, two instances of coordination are found: 'My mother hers name is …and work in the restaurant and my father in the restaurant'. This time, the first two coordinate units are complete in terms of SVO structure. Additionally, student D also says 'Saturday in morning I go to skating and afternoon I go to a beach and water'. Even though the same verb is used in the two coordinate units, the learner displays more sophisticated L2 syntactic knowledge through the use of coordination.

In regard to lexical diversity, Table 6.146 shows that the only verbs used by student C at T0 are 'be' and 'live' in utterances like 'My father is thirteen years old' and 'Where are you live?'. At T1, the learner also uses 'work', 'play' and 'have' in the following sentences: 'Work in the *ferreteria*' (ironmonger's), 'I'm playing the video *consola*' (console) and 'What's years you have?'. Interestingly, the learner uses the verb 'have' to ask about the researcher's age, when at T0 the same question was formulated as 'How years old?' with neither 'be' nor 'have', which suggests that the question 'What's years have you?' might no longer be formulaic. At T2, the verb 'practise' in 'You practise sports?' is also used in addition to the verbs used at previous times. Furthermore, the learner uses the verb 'be' for the first time to answer the question about his age. Likewise, at T3 the learner uses the verbs 'be', 'play', 'do', 'work' and 'have (got)'. Clearly, there is an increase in the number of verbs throughout the four times of data collection. As for the use of adjectives, 'old' is the most commonly used adjective in utterances like 'I'm ten years old'. At T3, the adjective 'good' is also used to describe the learner's family: 'My family is good'.

As for student D use of verbs according to Table 6.147, at T0 the learner uses the verb 'be' and 'have' in questions like 'How are you?', 'You I have animals?' or 'What is your name?'. At T1, the learner also uses 'go' and 'skate' in the utterance 'On Saturday I go to skating'. These verbs are also used at T2 and T3. Finally, at T3, student D uses 'work' in the utterance 'My mother hers name is …and work in the restaurant'. Concerning the use of adjectives, the only two adjectives used by the learner in the language samples provided in Table 6.147 are 'old' at T0 and T3 in sentences like 'Twelve years old' and 'favourite' at T0 in the question 'What is your favourite colour?'.

In terms of accuracy, as shown in Figures 6.42 and 6.43, both learners obtained similar scores at T0. At T1, however, student C increases the number of error-free units produced, while student D displays a sharp decrease. The rise in accuracy shown in Figure 6.42 might be explained by the fact that, as Table 6.146 indicates, student C produces a higher number of correct questions at T1 than at T0. At T0, the three questions asked by the learner are incorrect: 'How years old?', 'Where are you live?' or 'Where are your professions'. At T1, on the other hand, the learner produces three correct questions out of four 'Where [//] what is your name?', 'What's the name of your mum?', 'And your dad?'. Interestingly, the learner corrects himself in the first question. From T1 onwards, student C starts to decrease the percentage of error-free units. The errors found at T2 are related to the lack of vocabulary, as in the utterance 'Has a *ferreteria*' or to the wrong use of verb tenses, as in the sentence 'I'm playing the house' to describe a typical weekend activity. Incorrect verb forms are also present at T3, when the learner says 'I'm doing the homework and I'm playing playstation', instead of using the present simple tense. Additionally, another reason that may account for the decrease in accuracy at T2 is the increase in the number of questions the learner asks, most of

which contain errors. As previously seen, an increase in the amount of language produced affects accuracy levels. Some of the questions student C asks at T2 are 'Has got years?', 'You has a brother?', 'You practise sports?'. The first question seems to be an attempt to use subject-verb inversion with the verb 'have got', while in the second and third questions the learner did not use the auxiliary verb *do*. Besides, the learner uses the third person singular inflection in the verbs 'have got' and 'have'. At T3, other examples in regard to the lack of auxiliary verbs in questions can still be seen in the following sentences: 'What work your mother and your father?' 'You have a sister'? and 'What doing in the weekend?'. As for the lack of mistakes, it is worth noting that at T2 and T3 the learner uses the verb *be* to talk about his age: 'I'm eleven years old' and 'I'm twelve years old'.

With regard to the accuracy levels displayed in Table 6.147 by student D, it is worth commenting that the high percentage of error-free units at T0 seems to be explained by the little amount of language produced by the learner. For instance, student D answers question 2 (Any brothers or sisters?) by saying 'Three'. Even though the answer may seem incomplete, it is grammatically correct. Likewise, the learner answers the question about his parents' jobs with the utterance 'secretary and cook', which is also correct. Looking at the evolution of these two answers over the three remaining data collection times, it becomes clear that development involves errors. At T1, the learner answers the question about his family with the utterance 'Yes, I three brothers' while at T2, the learner adds the verb in 'I have three brothers'. As for the question related to his parents' jobs, at T1 the learner answers 'My mother secretary and my father cooking' and at T2 'My father cooker and my mother secretary'. These examples show that even though the utterances are still incorrect the learner is attempting to reproduce SVO order, while at T0 the learner did not use the subjects explicitly. At T3,

this utterance becomes even more elaborate when the learner uses the verb 'work': 'My mother hers name is … and work in the restaurant and my father in the restaurant'. Another utterance which exemplifies the relationship between complexity and accuracy is 'I am going of skating' at T1, 'I go in skating' at T2 and 'I go to skating' at T3. As seen from this example, the learner's attempt to produce a subordinate unit derives in the occurrence of errors. Finally, the type of errors found in the questions asked by the learner over the four data collection times are related to the use of formulaic language. As seen from the language sample provided, formulaic language might help learners obtain higher scores in accuracy. However, the use of unanalysed language chunks may also result in an increase in errors as in 'You I have animals?', in which the learner uses 'I have' as formulaic language at T0. At T2, the learner's use of the verb 'have' in the question 'You have brothers?', suggesting that the learner no longer views 'I have' as a chunk. Additionally, errors related to the wrong use of wh-question words are also found from T1 onwards in questions like 'What old are you?' or 'Where is your job?'. At T0, most questions are formulaic, which might be why no errors of this sort are found. Finally, at T2 and T3 errors concerned with the lack of the auxiliary verb in questions are also identified in the following examples: 'You have a brothers?' or 'And any brothers?'.

As for fluency, both graphs show an increase in speech rate in words from T0 to T3. As Tables 6.146 and 6.147 display, an increase in the amount of language produced is seen for both student C and D. In the case of student C, the increase is considerable in terms of the number of questions asked over the four data collection times. Student D, on the other hand, asks a similar number of questions from T0 to T3, however, he does increase the number of words in the answers provided to the researcher. Additionally, it is important to point out that student C is unable to provide an answer to questions 3 and

4 at T0. No instances of unanswered questions are found for student D. Another aspect which affects the learners' fluency levels relates to the number of times the researcher has to ask a question to obtain an answer from the learner. In the case of student C, the researcher has to ask question number 2 twice at T0 and T1 to obtain an answer from the learner. Moreover, the learner provides an answer to question 4 at T1 after the researcher has asked it five times. This is also found in Table 6.147. At T0, Student D answers question 4 after the researcher has formulated it twice. Similarly, at T2, the learner answers this same question in the second attempt. Finally, as for the use of the L1, student C uses the L1 at T1 and T2 to refer to his parents' jobs. At T3, the learner overcomes that by providing the English translation. No instances of L1 use are found in Table 6.147 for student D.

To sum up, the description of fluent CLIL and control learners' oral output at the four times of data collection suggests that language development at this level of language proficiency seems to take place at a slower rate when compared to highly fluent learners. Nevertheless, examples of the effects of syntactic complexity levels on accuracy were also found for average achievers. As for lexical diversity, both learners show an increase in the number of verbs, however, the fluent CLIL learner seems to produce a higher number of different verbs than the control fluent learner. In addition, it was observed that one of the explanations for the high number of errors was the misuse of formulaic language. In regard to fluency, the description of the learners' output showed improvement from T0 to T3. However, it is worth noting that the CLIL learner shows greater improvement in fluency at T2 and T3, while the control learner at T3.

6.2.1.1.3 Dysfluent learners from the CLIL and control groups

With regard to dysfluent learners, Figures 6.44 and 6.45 show their development in the CAF measures. As seen from Figure 6.44, student E (dysfluent CLIL learner) did not produce coordination until T1. From T1 to T2 there is a decrease in the production of coordinate units, while from T2 onwards the graph shows a sharp increase. As for his accuracy scores, Figure 6.44 shows a slight decrease from T0 to T1 followed by a sharp increase from T1 to T2. From T2 to T3, in contrast, the graph shows a considerable decrease in the accuracy measure. Interestingly, student's complexity scores are at the highest level when his accuracy levels are the lowest. As for fluency, student E shows a slight increase from T0 to T3. The data displayed in Figure 6.45, indicates that student F (dysfluent control learner) produced coordination at T1 as well. From T1 to T2 there is a gradual decrease in the percentage of coordinate units, while from T2 to T3 the percentage starts to increase slightly. In terms of accuracy, student F's percentage of error-free units is pretty high at T0. At T1, however, the percentage of error-free units decreases significantly. As seen in Figure 6.44, Figure 6.45 also shows that the lowest score in accuracy coincides in time with the highest score in complexity. From T1 onwards, accuracy shows a gradual but slight rise. Finally, fluency scores show a modest increase from T0 to T3. Tables 6.148 and 6.149 display the output produced by the two lerarners.

Figure 6.44
Evolution of CAF in the interview task. Dysfluent CLIL learner (student E).

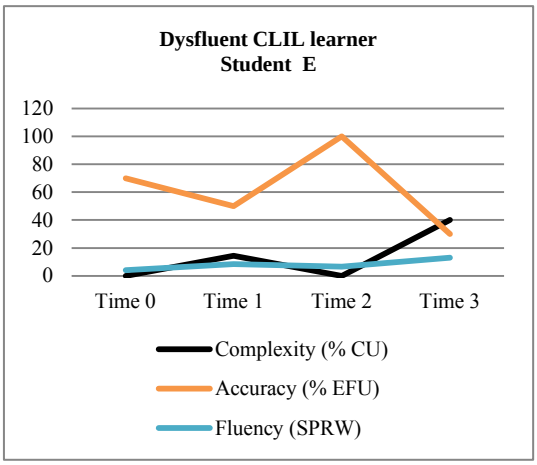


Figure 6.45
Evolution of CAF in the interview task. Dysfluent control learner (student F).

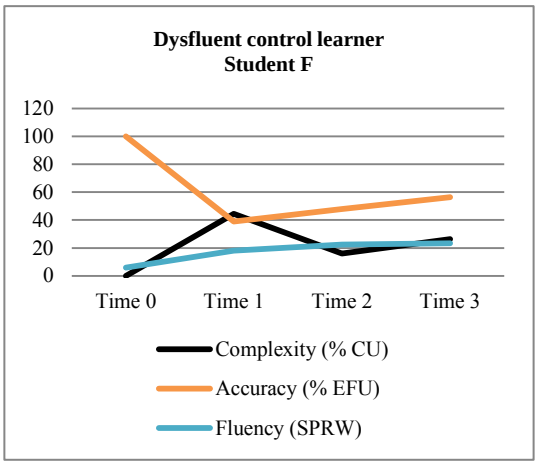


Table 6.148
Dysfluent CLIL learner. Student E's answers in the interview task.

Questions from the interview	**T0**	**T1**	**T2**	**T3**
1.How old are you?	[2nd attempt] Ten years old.	quants@s:c anys@ tinc@s:c? Ten years old.	Eleven years old.	Eleven years old.
2.Tell me about your family? **Any brothers or sisters?**	No an swer from the participant. Five.	One b rother a nd three sister.	Two br others a nd three sisters.	Two br others a nd three sisters.
3.What are your parents' jobs?	Play football.	No an swer from the participant.	[2nd attempt] This football.	No.
4.What do you usually do at the weekends	Play football.	què@s:c faig@s:c? Sunday I pl ay football.	Play football.	Sunday pl ay football a nd v isit my family.
5.Now it's your turn ask me a few questions.	What your name? Is years old?	What i s your name?	What's your name?	What's your name? How are you? The next Saturday you p lay t he sports? You have a boyfriend? You like pizza? Where [//] w hat is your o ffice? (Intended f orm: what is your job?)

Table 6.149
Dysfluent control learner. Student F's answers in the interview task.

Questions from the interview	**T0**	**T1**	**T2**	**T3**
1.How old are you?	OK.	Ten years.	Eleven years.	Eleven.
2.Tell me about your family? **Any brothers or sisters?**	A sister.	My family is my mum his name is … and my dad his name is … And my sister his name is … and my dog …	My family is great. My sister have eighty years. And the name of my mother is…and my dad is…And I have a dog. My dog is old.	I have a dog. I have a sister and two parents. My dog's name is … My parents' names are …and…my sister's name is…
3.What are your parents' jobs?	No answer from the participant.	[3rd attempt] Police.	My mum stay in the office and my dad engineer.	*This question wasn't asked by the researcher.*
4.What do you usually do at the weekends	Camping.	I go to the camping and play of my friends.	Usually go to the camping and play.	Well a tth e weekends usually I go to the camping but this weekend I go to the party.
5.Now it's your turn ask me a few questions.	What is your name? How old are you? It is teachers? (Intended form: are you a teacher?)	How old are you ? Have a sister? You go to the camping?	You like macaroni? When you was young you love school? What's your favourite colour? You have a brother? How old are you?	You have a brother? When you was young you like going to school? You like chocolate? What is your favourite animal?

As seen from Table 6.148, Student E's evolution in terms of complexity starts to take place at T1, with utterances like 'One brother and three sisters'. At T0, the learner only provides the total number of siblings 'Five' without indicating whether they are brothers or sisters. At T2 and T3, the utterance becomes even more accurate as the learner says 'Two brothers and three sisters'. Another relevant change in regard to complexity is seen in the comparison of the answers given to question 4 at T0 and T1. At T0, the learner says 'Play football' while the same question at T1 is answered with 'Sunday I play football'. The comparison of these two answers shows that the learner

has improved in terms of the use of the SVO structure and the addition of the adjunct 'Sunday'. This same utterance at T3 becomes a coordinate unit 'Sunday play football and visit my family', showing further syntactic development at the expense of accuracy.

Table 6.149 displays the language produced by student F in the interview task. As with student E, the more visible changes in complexity start at T1 with utterances like 'My family is my mum his name is…and my dad his name is…and my sister his name is…and my dog…'. At T0, the same question was answered with a mere 'A sister'. Likewise, at T1 the learner produces the following coordinate units: 'I go to the camping and play of my friends'. At T0, this question was answered with a one-word unit: ' camping'. At T2, student F's evolution continues. The answer related to his family is more elaborate in the sense that more ideas are transmitted: 'My family is great. My sister have eighty years. And the name of my mother is... and my dad is…. And I have a dog. My dog is old'. In addition to the amount of information conveyed in this utterance, it is worth highlighting that all the sentences produced by the learner have an SVO structure. The evolution of this utterance at T3 undergoes further transformations: 'I have a dog. I have a sister and two parents. My dog's name is…My parents' names are…and…and my sister's name is….' This utterance presents the pieces of information more clearly than at previous times and shows a correct use of coordination and the Saxon genitive. Additionally, two more instances of development in syntactic complexity displayed in Table 6.149 should be mentioned. The first one takes place at T2, when the learner uses a subordinate unit in one of the questions he asks the researcher ' When you was young you loves chool?'. Interestingly, this utterance is used at T2 when the percentages of coordination decrease. It appears that low percentages in coordination at certain times of the study do not involve a decrease in syntactic complexity. At T3, this question becomes even more complex: 'When you

was young you like going to school?'. This time, the learner has added another subordinate unit 'going to school' to the utterance in addition to the adverbial clause. The second instance of syntactic evolution is seen at T3 in the following sentence: 'Well at the weekends usually I go to the camping but this weekend I go to the party'. Interestingly, the same idea at T0 was transmitted with one word: 'camping'. From T0 to T3, the learner has succeeded in producing SVO structures, adding temporal expressions and using coordination.

Concerning lexical diversity, the biggest growth in student E's lexicon is seen at T3. At T0 the only verbs the learner uses are 'play' and 'be' in the utterances 'Play football' and 'Is years old?'. At T1, the verbs used are the same, the only difference is that this time the verb 'be' is used correctly in the formulaic question 'What is your name?'. Exactly the same occurs at T2. Finally, at T3 the learner uses a wider variety of verbs in the following sentences: 'Sunday I play football and visit my family', 'The next Saturday you play sports?', 'You have a boyfriend?' and 'You like pizza?'. As for the use of adjectives, the only adjective used throughout the study is 'old' in the utterances 'Ten years old', 'Eleven years old' and 'Is years old?'. As for student F, the strongest growth in terms of the number of verbs and adjectives used takes place at T2 and T3. At T0 the only verb used is 'be' in the formulaic questions 'What is your name?' and 'How old are you?'. At T1, the learner also uses the verbs 'go', 'play' and 'have' in the utterances 'I go to the camping and play of friends' and 'Have you a sister?'. Student F uses 'stay', 'love' and 'like' in addition to the already mentioned verbs at T2 in 'My mum stay in the office', 'When you was young you love school?' and 'You like macaroni?'. It is worth noting that the learner uses the past tense of the verb *be,* which indicates further development in his L2 grammar system. At T3, the learner uses the verbs 'be', 'have', 'go' and 'like'. No instances of new verbs are found at this time. In

regard to the use of adjectives, at T0 the only adjective used is 'old' in the question 'How old are you?'. This same question is repeated at T1. At T2, in contrast, the adjective 'old' is used in a different context, 'My dog is old', in addition to the formulaic question 'How old are you?'. At T2, the learner also uses the adjective 'great' in the utterance 'My family is great' and 'favourite' in the question 'What's your favourite colour?'. Finally, at T3 the learner uses 'young' and 'favourite' in the questions 'When you was young you like going to school?' and 'What is your favourite animal?'.

Regarding accuracy, the language samples displayed in Table 6.148 show that the amount of language produced by student E at T0 is quite limited, which might explain the high scores in accuracy shown in Figure 6.44. In spite of this, errors are also found at T0. Some of the mistakes displayed in Table 6.148 are related to the absence of the verb *be*, as in the question 'What your name?', the lack of a wh-question word and the wrong verb selection as in 'Is years old?'. At T1, the verb *be* is inserted into the question 'What is your name?', which is also used correctly at T2 and T3. At T2, no instances of mistakes are found in the table, apart from the non-context dependent utterance produced to answer question 3. At T3, two main errors are identified. The first one concerns the absence of the subject in the sentence 'Sunday play football and visit my family' and the second one relates to the lack of the auxiliary verb *do* in the questions 'The next Sunday you play sports?', 'You have a boyfriend?' and 'You like pizza?'. In regard to self-corrections, student E corrects himself at T3 in the formulation of the following question: 'Where [//] what is your office?'. The intended question here is *What is your job*?, however, the student seems to have anglicised the L1 word for job (*ofici*) and uses 'office'.

As for student F's accuracy development, one of the most relevant changes concerns the genitive case. As seen from Table 6.149, at T1, the learner indicates possession using the possessive determiner 'his' in masculine and feminine contexts as in 'My mum his name is…and my dad his name is…'. At T2, the possessive is expressed by means of the preposition *of* in the same answer 'And the name of my mother is…'. Interestingly, at T3, the learner uses the genitive 's in the answers 'My dog's name is…My parents' names are…and… and my sister's name is…' In addition to the production of the genitive case, errors are also found concerning the lack of subjects. At T1, for instance, the learner uses the subject in the sentence 'I go to the camping'. However, at T2, the learner expresses the same without using the subject: 'Usually go to the camping'. Eventually, at T3 the learner uses the subject again 'I go to the camping'. As for question formation, the most common errors found at T1, T2 and T3 are the absence of the auxiliary verb *do* in yes-no questions. For example, at T1 the learner uses the following questions: 'Have a sister?', and 'You go to the camping?'. At T2, three instances of questions lacking the auxiliary verb are found 'You like macaroni?', 'You have a brother?' and 'When you was young you love school?'. At T3 this error still persists in utterances like 'You like chocolate?'. Additionally, an instance of the use of *is it* at the beginning of the question as formulaic language in the formulation of questions is also found at T0 'Is it teachers?'. Another relevant error in the formation of questions is the lack of subjects. At T1 student F produces the question 'Have a sister?'. At T3, however, the learner does use the subject explicitly in the question 'You have a brother?', showing improvement. Finally, it is worth mentioning, that even though the learner uses the past tense form of the verb *be* at T2 and T3, the learner fails to achieve subject-verb agreement in the question 'When you was young…?'.

Concerning fluency, student E's improvement is rather modest according to Figure 6.44. In spite of this, two fluency-related aspects do seem to undergo considerable improvement according to Table 6.148. To start with, the learner's use of the L1 decreases from T0 to T3. As seen from the table, the only two instances of L1 use are found at T1: '*quants anys tinc*?' (How old am I?) and '*què faig*?' (What do I do?). As seen from the translations provided, the function of the L1 in these two cases is to check comprehension by echoing the question asked by the researcher in Catalan. The second aspect of fluency relates to the fact that amount of L2 language produced undergoes a substantial increase from T0 to T3, particularly when the learner is asked to formulate questions. The learner asks two questions at T0, one question at T1 and T2 and five questions at T3. In addition, another example of improvement in fluency is shown by the gradual disappearance of problems to understand question 3 (What are your parents' job?). At T0, the answer provided by the learner 'Play football' indicates that he does not understand the meaning of the question. In addition, the learner does not answer the question at T1. At T2 the learner needs to listen to the question twice to provide the answer 'This football', which does not fulfil the communicative purpose of the question, either. Finally, at T3, the learner answers 'No', indicating that he still does not understand the question. The fact that student E shows comprehension problems has a direct impact on his fluency scores, as the total number of words in English decreases as a result of not understanding what he is being asked. Besides, pause time increases, as well.

As for student F, his fluency development is basically seen in the length of the utterances he produces. As seen from Table 6.149 there is a big change from T0 to T1 in terms of amount of language used. At T0, most utterances are one-word units. At T1, on the other hand, the sentences are much longer and more information is conveyed in the

answers provided. This tendency continues through T2 and T3 when the learner starts producing subordination. As for the use of L1 words, no instances of L1 were found in the samples provided in Table 6.149. In terms of comprehension, it is worth pointing out that student F also shows problems in understanding question 3. At T0, the learner does not answer the question. At T1, the learner answers by saying 'Police' after the researcher has asked the question three times. Finally, at T3, the learner provides a complete and coherent answer, 'My mum stay in the office and my dad engineer', which seems to suggest that the reason why the learner does not answer at T0 or answers 'Police' at T1 relates to the fact that he does not know the English translation for the words he needs.

All in all, one of the most distinctive features of dysfluent learners' development is that most instances of improvement in CAF are seen between T2 and T3 in the production of coordination, use of SVO word order and vocabulary growth. The first two data collection times, on the other hand, show an overuse of one-word units and comprehension problems. Additionally, the description of the output indicates that the dysfluent control learner seems to display greater development than the dysfluent CLIL learner in terms of complexity and fluency mainly.

The most relevant aspects in regard to the differences in the output produced by highly fluent, fluent and dysfluent CLIL and control learners will be addressed in the discussion section. In the following section a description of the oral output produced by these same learners in the narrative task will be provided.

6.2.1.2 Narrative task

CLIL and control groups' evolution of CAF in the narrative task is shown in Figures 6.46 and 6.47. As displayed in the graphs, CLIL and control learners'

development in CAF measures is quite similar throughout the two years of the study. Syntactic complexity, measured by the percentage of coordinate units, shows similar tendencies in both groups. Very few instances of coordination are found at T0 while the amount of coordination produced in the two groups increases considerably at T1. From T1 to T2 the graphs show a slight decrease. At T3, the percentage of coordinate units increases again for the two groups. As for the percentage of error-free units, both graphs display similar achievement scores at the four data collection times. Finally, the learners' fluency scores show a gradual and constant increase from T0 to T3 in the two groups.

Figure 6.46
Evolution of CAF in the CLIL group. Narrative task.

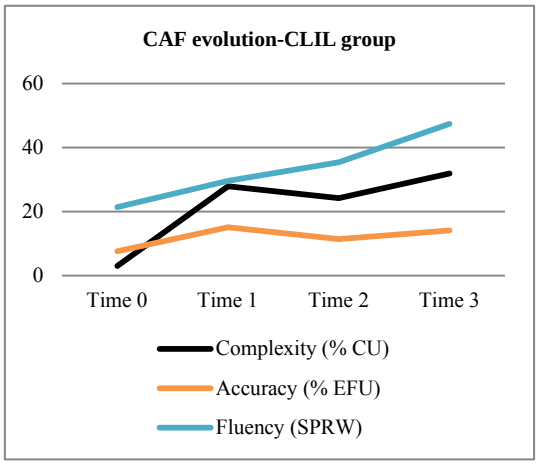
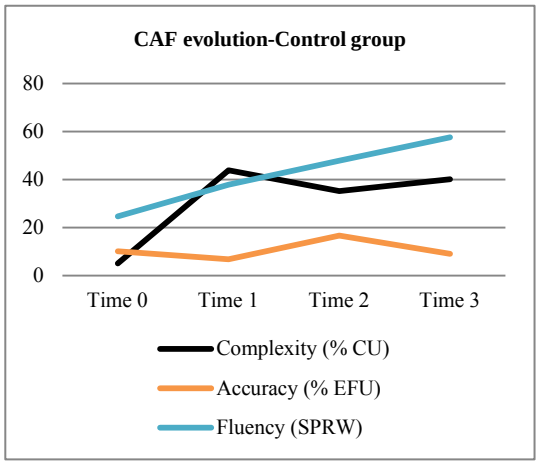


Figure 6.47
Evolution of CAF in the control group. Narrative task.



6.2.1.2.1 Highly fluent learners from the CLIL and control groups

The evolution of CAF measures of the two highly fluent learners in the narrative task is displayed in Figures 6.48 and 6.49. Additionally, Tables 6.150 and 5.151 show their output during the task at the four times of data collection (see Appendix C for the

complete transcipts). The coding symbols used in the tables are described in Appendix D.

Figure 6.48
Evolution of CAF in the narrative task. Highly fluent CLIL learner (student A).

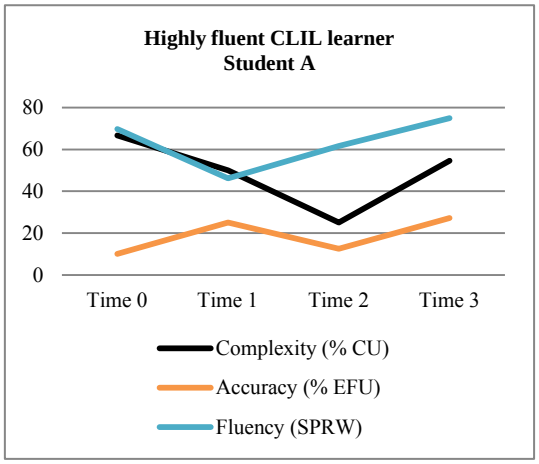
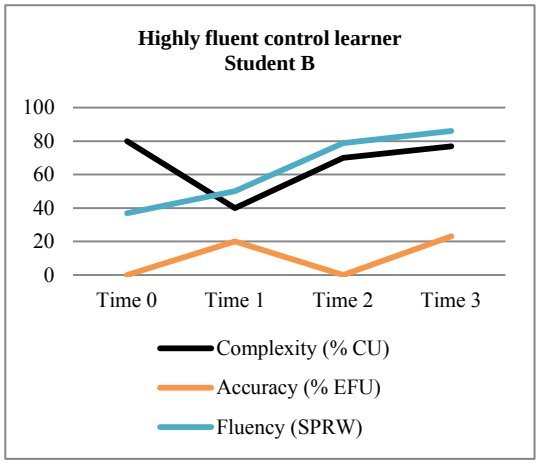


Figure 6.49
Evolution of CAF in the narrative task. Highly fluent control learner (student B).



As seen from Figures 6.48 and 6.49, both student A (highly fluent CLIL learner) and student B (highly fluent control learner) show a decrease in the percentage of coordinate units produced from T0 to T1. Student A's decrease is gradual and constant up to T2. From T2 to T3, student A displays a substantial increase in the percentage of coordinate units. Student B, on the other hand, shows a sharp increase from T1 to T2 and slight improvement from T2 to T3. In regard to accuracy, both learners display quite irregular tendencies. Both students show an increase from T0 to T1. From T1 to T2, in contrast, the Figures show a decrease, which in the case of student B goes down to zero. Eventually, from T2 to T3, the learners' accuracy levels increase again. Finally, the learners' fluency levels improve from T0 to T3 in general terms. Despite that, they display different trends. Student A shows a decrease from T0 to T1, which is followed by a constant increase from T1 to T3, whereas student B shows constant improvement from T0 to T3.

Table 6.150
Highly fluent CLIL learner. Student A's output in the narrative task.

Time	Learner's output
Time 0	This boys are cooking a cake. In the picture two the dog is looking in the basket and the mum is [//] and the mum and the boys are looking one map. The picture three the boys go out the home and his mothers says bye. In the picture four the boys are in the mountains in the camp@s:c two cows and xxx. And the boy and the girl open the basket and the dog is go out. The girl are surprise. The boys are looking in the basket and in the basket no have [/] no have the cake. And the dog is go out.
Time 1	The boy and the girl preparate a picnic and the dog go [//] goes in [//] goes on the +["] basket. In picture three the boy is going camp@s:c and his mother says bye. In the camp@s:c the boys open the basket and the dog goes out. And the food que@s:c estava@s:c in the basket the dogs [//] the dog eats.
Time 2	Picture number the boy has to prepare a picnic and the dog look the bueno@s:c the boys. In picture number two when mum no [//] when the boys look into the map the dog goes into the [/] into the cest@s:c. The picture number three the boys go out of the home. In four he are in the picnic. In five he open the bueno@s:c [//] in five the dogs go out of the picnic cest@s:c and in six no are [/] no are more food in the cest@s:c. The dog eat the food.
Time 3	First picture the boys are preparing the breakfast and the dog looks about it. And in picture number two mum tells him to go to the park to eat the breakfast and the dog is jumping in the entering the cest@s:c. And in picture three the boys are saying bye to mum and they are going to the park. In picture number four the boys are looking for [/] for seat. In number five the boys are looking into [/] into the cest@s:c and the dog goes out. And in number six in the cest@s:c aren't any [/] any food and the dog runs out.

Table 6.151
Highly fluent control learner. Student B's output in the narrative task.

Time	Learner's output
Time 0	It is one boy and one girl and a dog. His mum explain the route of this map and the dog is in the castel@s:c. The two boys go to the [/] to the forest. And his dog eat the food and run in the forest.
Time 1	In the picture number one the boy and the girl is doing the breakfast. In picture number two his mum is [//] bueno@s:c he have a map and he point the [/] the bueno@s:c the mountain the forest. In the number three the two boys are in the street and he says byebye for him mum. In the number four they are arrive in the forest and this. In number five they look that in [/] in her nest there are a dog and + ... ++ In the end in number six he see [//] they see that in the nest do not have the breakfast the dog eats.
Time 2	In the firs pic [/] first pic [/] first picture there are a boy and a girl and his mum and they prepare sandwich for a picnic. In the second picture her [/] her [//] his mum [//] they [/] they [/] they give him a map but the dog he's put in the cist [/] cist [//] castle@s:c bueno@s:c cistell@s:c. In the other picture the boy and the girl they go to the forest and in the forest they [/] they [/] they [/] they they app [//] app [//] they appear [//] they look the dog but it [//] they are in the +...["] in the basket. The dog go to run. The boys look in the cistell@s:c and they are [//] they aren't the sandwich. The dog eat the sandwich.
Time 3	In the picture number one we can look the boys preparing the breakfast and his dog looking. In the number two his da [//] her mum are showing the map of [/] of the [/] of the [/] yes of the world and his dog is looking to the [//] to the basket. In the number three the boys walking in the street and say goodbye to his mum. In number four the boys are climbing a [//] one small cliff. In number five they open the basket and they see her [//] his dog In [/] in this [//] in number six they look another time the basket and they see that the breakfast isn't [/] isn't there. The dog eat [/] the dog eat the breakfast.

As displayed in Table 6.150, at T0, student A produces eight coordinate units: 'In the picture two the dog is looking in the basket and the mum is [//] and the mum and the boys are looking one map', 'The picture three the boys go out the home and his mothers says bye', 'And the boy and the girl open the basket and the dog is go out' and

‘The boys are looking in the basket and in the basket no have [/] no have the cake’. It is important to note that all these coordinate units, as well as the other units produced by the learner at this time, contain complete SVO structures, indicating the learner’s high degree of syntax development and complexity. Additionally, the learner uses the prepositional phrase ‘In the picture…’ or the noun phrase ‘Picture…’ to organise his narrative discourse. No instances of subordination are found at T0. At T1, the learner produces the following coordinate units: ‘The girl and the boy preparate a picnic and the dog go [/] goes in [//] goes on the [“] basket’, ‘In picture three the boy is going *camp* (countryside) and his mother says bye’ and ‘In the *camp* the boys open the basket and the dog goes out’. As seen at T0, no instances of subordination are found at T1 either, however, the learner does attempt, unsuccessfully, to produce a relative clause: ‘The food *que estava* (that was) in the basket the dogs [//] the dog eats’. In spite of the fact that the learner is struggling with word order, the attempt to produce a relative clause suggests that the learner is trying to use his already learnt knowledge to produce syntactically more complex language structures. In addition, the learner continues using prepositional phrases like ‘In picture three’ to organise his narrative discourse. The comparison of the output produced by student A at T0 and T1 seems to indicate that the learner has not improved in terms of syntactic complexity, however, the attempts to produce subordination and the attempts to become more accurate in his productions (as explained in the following paragraphs) might account for the decrease in complexity measured by the percentage of coordinate units. At T2, the percentage of coordinate units still shows a decreasing tendency. This time the learner produces four coordinate units: ‘Picture number the boy has to prepare a picnic and the dog look the *bueno* (well) the boys’ and ‘In five he open the *bueno* [//] in five the dogs go out of the picnic *cest* (basket) and in six no are [/] no are more food in the *cest*’. At this time, the learner

seems to decrease the number of coordinate units in favour of subordination in the following utterances: 'when mum no [//] when the boys look into the map the dog goes into the [/] into the *cest*', which includes an adverbial clause. At T2, the learner's output shows an increase in the use of discourse markers like: 'Picture number', 'In picture number two' or 'In five'. Finally, at T3 the learner seems to be using all his linguistic resources to produce syntactically complex language through extensive use of coordination (as at T0) in addition to subordination (which emerged at T1 with an attempt to produce a relative clause). As a result, all the utterances he produces contain either coordination, subordination or both. This increase in complexity is also illustrated in the graph. The instances of coordination are: 'First picture the boys are preparing the breakfast and the dog looks about it', 'And in picture number two mum tells him to go to the park to eat the breakfast and the dog is jumping in the entering the *cest*', 'And in picture three the boys are saying bye to mum and they are going to the park', 'In number five the boys are looking into [/] into the *cest* and the dog goes out' and 'And in number six in the *cest* aren't any [/] any food and the dog runs out'. In regard to subordination, student A produces a nominal to-infinitive clause and an adverbial clause in the utterance 'mum tells him to go to the park to eat the breakfast' which at the same time is a coordinate unit. In addition, there is an attempt to produce another adverbial clause in: 'and the dog is jumping in the entering the *cest*' (meaning to get into the basket). Besides, as at previous times, the learner uses discourse markers to organise the narrative discourse. The only two differences in comparison to previous times is that at T3 the learner uses the ordinal determiner 'first' and the phrase 'In picture number…' instead of 'in four' or 'in the picture three' to introduce his descriptions of narrative events.

In regard to student B, the highly fluent control learner, Figure 6.49 also shows a decrease in the percentage of coordinate units produced from T0 to T1. As seen from Table 6.151, the proportion of coordinate units in relation to the total number of units is higher at T0 than at T1. At T0, the learner produces three utterances, two of which contain coordination: ‘His mum explain the route of this map and the dog is in the *castel’* (meaning basket) and ‘His dog eat the food and run in the forest’. As for the structure of the utterances produced at T0, all of them contain SVO structures. In regard to discourse narrative markers, no use of expressions to refer to the pictures in the story is found at T0. At T1, despite the fact that the percentage of coordinate units in relation to the total number of units decreases, the first subordinate unit is produced: ‘In number five they look that in [/] in her nest there are a dog and + ...’. Even though the utterance contains lexical and morphological errors, the learner attempts to produce a nominal that-clause. Additionally, at the same time, the learner produces another nominal clause very similar in structure and meaning to the one just presented in which some errors are corrected: ‘In the end in number six he see [//] they see that in the nest do not have the breakfast the dog eats’. As seen from the utterance, the nominal clause also includes an attempt to produce another embedded clause ‘the dog eats’, by either introducing a relative pronoun ‘that the dog has eaten’ or the conjunction ‘because the dog has eaten it’. In addition to the emergence of subordination, the learner also uses coordination at T1 in the utterances: ‘In picture number two his mum is [//] *bueno* (well) he have a map and he point the [/] the *bueno* the mountain the forest’ and ‘In the number three the two boys are in the street and he says bye bye for him mum’. As seen from Figure 6.49 and the sample provided in Table 6.151, at T1 the learner uses less coordination than at T0 in proportion to the total amount of language produced, however, subordinate clauses start to emerge as an indicator of syntactic development. Besides, it is worth

highlighting that as seen at T0, the learner succeeds in using SVO structures during the task. At T1, the learner also uses discourse markers like 'In the picture number one' 'in the number three' or 'in the end'. At T2, the amount of coordination starts to increase as seen in the utterances: 'In the first pic [/] first pic [/] first picture there are a boy and a girl and his mum and they prepare sandwich for a picnic', 'In the other picture the boy and the girl they go to the forest and in the forest they [/] they [/] they [/] they [/] they app [//] app [//] they appear [//] they look the dog' and 'The boys look in the *cistell* (basket) and they are [//] they aren't the sandwich'. Two more attempts to produce coordination are found using the conjunction *but* in the utterances: 'but the dog he's put in the cist [/] cist [//] *castle bueno cistell*' and 'but it [//] they are in the +...["] in the basket'. As far as subordination is concerned, an instance of subordinate unit is found in the utterance 'The dog go to run', which includes a to-infinitive clause. As for discourse markers, student B uses the ordinal determiner for the first time in the following phrases: 'In the first picture' or 'In the second picture'. At T3, coordination is used in almost all the utterances student B produces: 'In the picture one we can look the boys preparing the breakfast and his dog looking', 'In the number two his da [//] her mum are showing the map of [/] of the [/] of the [/] yes of the world and his dog is looking to the to the basket', 'In the number three the boys walking in the street and say goodbye to his mum', 'In number five they open the basket and they see her [//] his dog' and 'In [/] in this [//] in number six they look another time the basket and they see that the breakfast isn't [/] isn't there'. Besides the use of coordination, the learner also uses two subordinate nominal clauses which are included in coordinate clauses: 'In the picture one we can look the boys preparing the breakfast' and 'and they see that the breakfast isn't [/] isn't there'. As seen from these examples, the learners' linguistic system already counts on subordination and coordination to express syntactically complex messages.

Finally, the learner also uses discourse markers to refer to the events in the picture: 'In the picture number one', 'In number four', or 'In number five'.

As for lexical diversity, as seen from the samples provided in Table 6.150, Student A increases the number of verbs and verb types he produces during the task throughout the four data collection times. At T0, the verbs used by the learner are 'cook', 'look', 'go out', 'say', 'be', 'open' and 'have'. At T1, the learner changes 'cook' for 'preparate' (prepare) and uses the verb 'go' with the prepositions 'in', 'on' and 'out' in the following utterances: 'The dog goes in [//] goes on the ["] basket' and 'the dog goes out'. Even though the learner does not succeed in the completion of the first utterance and struggles with the use of prepositions, the fact that the learner is aware of the meaning of the prepositions in his discourse reveals the learner's growth in vocabulary at T1. The use of prepositions to modify verbs will also be seen at the following times. At T2, the learner uses 'prepare', 'look', 'look into', 'go into', 'go out', 'open' and 'eat'. This time, the learner uses 'go into' in the utterance he was struggling with at the previous data collection time: 'The dog goes into the *cest*'. Additionally, he also uses a similar structure in 'The dog go out of the picnic *cest*', indicating that he is starting to handle the use of prepositions to complement the meaning of the verb 'go'. At T3, the learner uses the phrasal verb 'look for' in the utterance 'the boys are looking for seat' and uses the preposition 'out' and 'into' to complement verbs other than *go* in the utterances 'the dog runs out' and 'the boys are looking into the *cest*'. In addition, it is worth highlighting the use of the verb 'tell' at T3 in the sentence 'Mum tells him to go to the park to eat breakfast'. Regarding the use of adjectives, the only attempt to produce an adjective is found at T0 to describe one of the characters' feelings: 'The girl are surprise'. As displayed in Table 6.151, student B uses similar verbs. At T0, the learner uses 'explain', 'be', 'go', 'eat' and 'run'. At T1, the verbs produced are 'do',

'be', 'have', 'point', 'say', 'arrive', 'look', 'see' and 'eat'. At T2, the number of verbs continues increasing. This time the learner also uses the verbs 'give' and 'appear'. The verb 'give' is used in the utterance 'They give him a map'. As for the verb 'appear', the learner attempts to say that the dog appears, however, he fails to do so and changes the utterance to 'They look the dog' (meaning they see the dog). Finally at T3, the new verbs used by the learner are 'walk', 'show' and 'climb'. No instances of phrasal verbs or prepositions to complement the meaning of verbs are found in student B's output. In regard to adjectives, student B does not use any adjective either.

Regarding the percentage of error-free units, as seen from Table 6.150, student A produced one error-free unit at T0: 'And the girl and the boy open the basket'. As for the types of mistakes, Table 6.150 shows that at T0 errors are lexical, as in 'cooking a cake' or 'the boys' to refer to children, morphological as in the utterance 'the girl are surprise' and syntactic as in the use of negation 'In the basket no have the cake'. Despite these mistakes, the learner also unfolds correct use of the inflectional system of the verb *be* in: 'The boys are in the mountains' and 'The dog is looking in the basket'. Additionally, student A also succeeds in marking the verb 'say' for third person singular subject verb agreement in the sentence: 'His mothers says bye'. At T0, mistakes are also found in relation to the insertion of the verb *be* before main verbs. As exemplified in the utterance 'The dog is go out', student A uses the verb 'be' as an auxiliary verb using the present simple tense. This error is no longer found at T1, when the learner succeeds in using the correct verb form of 'go' without the verb 'be': 'The dog goes out'. As regards third person verb forms, it is worth noting that the learner corrects himself in: 'The dog go [//] goes in', suggesting development in his morphological system. No instances of incorrect verb forms are found at T1, which might explain the increase in the percentage of error-free units produced at T1: 'the dog eats' and 'the dog goes out'.

Additionally, the samples from the table seem to suggest that the learner is making attempts to become lexically more accurate in his productions. An example would be the use of 'the boy and the girl' instead of 'the boys', as he did at T0. At T2, the learner makes mistakes related to verb forms again. In particular, the learner does not inflect verbs for third person singular as the following utterances exemplify: 'The dog look the boys' or ' The dog eat the food'. As previously described, the learner's syntactic complexity seems to increase through the use of subordination, which might explain the decrease in accuracy. Finally, another of the mistakes found at T2 worth commenting on is found in the utterance ' In six no are more food', which derives to ' In number six aren't any food' at T3. Even though the learner is still making mistakes, traces of evolution towards an accurate use of the sentence *there isn't any food* can be observed through the use of negation in the verb. At T3, the learner uses most verbs correctly again: 'The dog runs out', 'The dog goes out', 'They are going to the park', 'The dog is jumping', at the same time as he produces accurate subordinate clauses, such as 'Mum tells him to go to the park to eat breakfast'. As a result, both accuracy and complexity levels increase. Nevertheless, the child has not succeeded in overcoming previous mistakes such as using 'the boys' to refer to children (as he did at T1) or the use of prepositions as in 'The dog looks about it', yet. Regarding student B's accuracy levels, no instances of error-free units are found at T0. The most relevant mistakes are related to the use of formulaic language at T0 in the utterance 'It is one boy and one girl and one dog '. As explained previously, the chunk ' it is' is usually learnt and used as formulaic language among YLLs. Additionally, the learner only uses bare infinitive forms: 'His mum explain the route', 'The two boys go to the forest' and 'And his dog eat'. At T1, two error-free units are found: ' They arrive in the forest' and ' They see that...'. At this time, student B uses the third person singular form of the verb correctly

once for the first time: 'He says bye bye for him mum'. At T1, the learner also makes mistakes when using the verb 'look' instead of 'see' in the utterance: 'They look that in the nest…'. Additionally, other lexical errors like the use of 'net' instead of *basket* or the use of 'boys' instead of *children* are observed. The former lexical error is not corrected until T3, when the learner uses 'basket'. The latter is corrected at T2, when the learner uses 'the boy and the girl'. At T2, however, no instances of error-free units are found, probably due to the increase in syntactic complexity. At this time, the most common mistakes are related to the use of possessive adjectives as in: 'His mum and they', meaning their mum, or the use of subject and object pronouns: 'They give him a map' instead of *She gives them a map*. Another mistake found at T2 refers to the use of 'They are' instead of *There is* as exemplified in 'They aren't the sandwich'. Finally, the learner fails to use the simple past tense or present perfect tense to express a past event: 'The dog eat the sandwich'. At T3, the misuse of possessive adjectives seems to become generalised: 'Her mum are showing…'or 'His dog looking'. A piece of evidence that accounts for the problems student B has with the use of possessive adjectives is exemplified by the presence of self-corrections and retractions in utterances where possessive adjectives are used: 'And they see her [//] his dog' or 'In number two his da+…[//] her mum. At T3, the learner uses the phrase 'the boys' to refer to children again. Likewise, as seen at T2, the learner fails to use the past tense of the verb in the utterance 'The dog eat the breakfast'. Despite the number of errors found at T3, the learner succeeds in producing some error-free units: 'They open the basket' and 'They see that the breakfast isn't there'.

In regard to fluency, student A shows a decrease in the amount of language produced from T0 to T1, along with an increase in the number of L1 words used. At T2, in contrast, the learner produces longer utterances. Some L1 use is still present at T2 to

refer to the word *basket* or to indicate a false start: 'In five he open the *bueno* (well) [//] in five the dogs go out…'. At T3 the number of words in English continues increasing at the same time as false starts and L1 words decrease. This time the only L1 word used is '*cest*' to refer to basket. Student B's fluency development shows a more regular tendency inasmuch as the learner increases the amount of language produced from one data collection time to the other. The biggest growth in fluency is seen from T0 to T1, when the learner increases the number of words produced. As for the use of L1 words, the only instances of L1 use found is the word '*castel*' or '*cistella*' (basket) at T0 and T2 and the word '*bueno*' (well) at T2 and T3 to indicate false starts or problems with L2 vocabulary knowledge.

To sum up, the data displayed in Figures 6.48 and 6.49 and Table 6.150 and 6.151 show an increase in complexity and fluency for both student A and B. In addition, instances of the relationship between coordination and subordination as well as syntactic complexity and accuracy have been described. As shown in the descriptions, both highly fluent learners improve significantly in complexity at T3 through extensive use of coordination and subordination. However, the highly fluent CLIL learner seems to develop further in lexical complexity thereby the use of prepositions to complement the meaning of verbs. In regard to accuracy, both students display irregular tendencies.

6.2.1.2.2 Fluent learners from the CLIL and control groups

As seen in Figures 6.50 and 6.51, Student C, the fluent CLIL learner, does not produce any coordinate units at T0. At T1, the percentage of coordinate units produced increases considerably. At T2, on the other hand, the graph shows another decrease in coordination, which is followed by an increase at T3. Student D, the fluent control learner, in contrast, displays a more regular trend. At T1, the learner produces the

highest percentage of coordinate units while at T2 the percentage decreases slightly and remains constant until T3. As for accuracy, student C's tendency is quite irregular. At T0 and T2, no instances of error-free units are found, while at T1 and T3 the percentages increase. In regard to student D, no error-free units are found at any of the four data collection times. Concerning fluency, both learners show an increase in speech rate in words from T0 to T3. In the case of student C, the sharpest increase is from T2 to T3. As for Student D, a slight decreasing tendency is observed from T0 to T2, while from T2 to T3 the learner shows a considerable increase. Tables 6.152 and 6.153 display the output produced by the two learners in the narrative task.

Figure 6.50
Evolution of CAF in the narrative task. Fluent CLIL learner (student C).

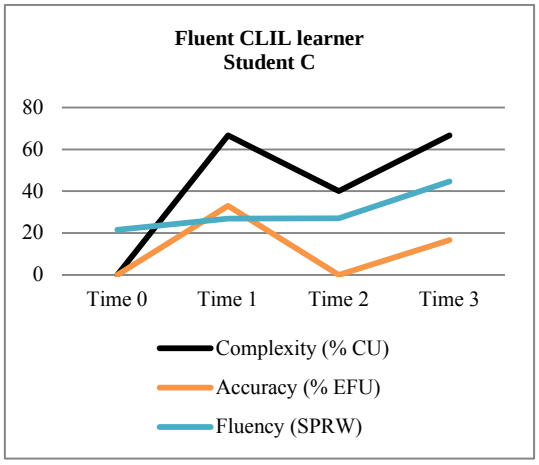


Figure 6.51
Evolution of CAF in the narrative task. Fluent control learner (student D).

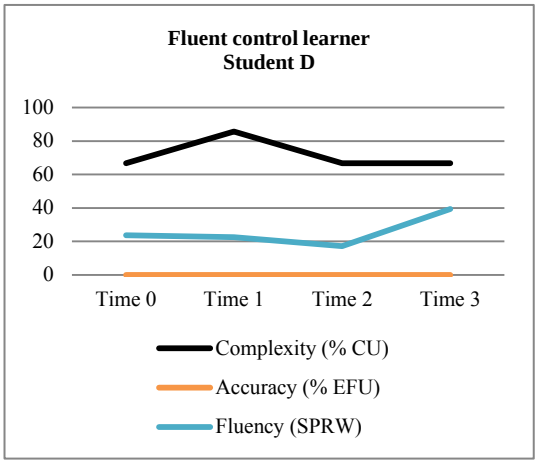


Table 6.152
Fluent CLIL learner. Student C's output in the narrative task.

Time	Learner's output
Time 0	A girl [/] a girl and a boy a picnic [//] doing a picnic. The dog putting a picnic. The mother. Bye bye. In the mountain. In the picnic have got a dog and a xxx
Time 1	The boy and the girl prepare a picnic and the dog eat the picnic. The boy and the girl ++ they say the man. And this they go [/] they go the xxx boy The dog eat the food.
Time 2	The boy and girl preparing the picnic. And your mum [/] mum [/] mum has a [/] a map and the dog has a picnic. The boy and the girl going the [/] the picnic and the picnic the is very happy for the dog has eat the picnic. The [/] the dog is food the picnic.
Time 3	The boy and girl prepare the picnic. His mum he's got the map but the dog he's eat [/] eat [//] ate the picnic. The boy and the girl going to the mountain. In the mountain he's with the dog and the dog eat [//] is ate the picnic.

Table 6.153
Fluent control learner. Student D's output in the narrative task.

Time	Learner's output
Time 0	The mum and the boys. I go to. The mum of the boys looking the map and the dog looking the food. The boys [/] the boys we say bye bye at the mum and I go to the. In here the dog eating of the food of the boys. The boys it is sad and the basket no [/] no of food.
Time 1	The boy and girl I going to [//] I cooking and [/] and the mother a map. A dog look inside the basket and yours mother say bye bye. A boy and girl going a mountain and look at the dog in the [/] in the food. The dog is [//] the dog eat the [/] the food and the cheese.
Time 2	Two boys preparing a [/] a food and her [/] her mum take a map A dog he's put in the basket. Boys xxx forest. And hers dog eating a food of the boys and boys [//] they look at the basket. They find the dog then eating dog.
Time 3	A boy and girl go to the excursion and prepare the breakfast Hers mother give a map. A boys [//] boy and girl go to the mountain and # find a dog in a basket. A breakfast the dog eating [//] the dog eating the breakfast.

As shown in Table 6.152, no instances of coordination are found in the fluent CLIL learner's output at T0 due to the fact that the utterances produced are very simple. The utterances consist of simple noun phrases as in 'A girl and a boy', 'The mother', 'The dog' or prepositional phrases like 'In the mountain' or 'In the picnic'. The only unit observed is: 'A girl [/] a girl and a boy a picnic [//] doing a picnic'. This utterance is the only one that is meaningful in the context of the task and which contains the SVO structure. At T1, student C produces two coordinate units: 'The boy and the girl prepare the picnic and the dog eat the picnic'. In addition to these coordinate units, the learner also produces the following unit: 'The dog eat the food'. The rest of the utterances are incomplete units containing noun phrases like 'The boy and the girl' or a combination of a subject pronoun and verb: 'They go'. At T2, student C's syntactic complexity increases substantially. First of all, the learner produces two coordinate units: 'your mum has a map and the dog has a picnic'. In addition, the learner unsuccessfully attempts to produce two more coordinate units: 'The boy and the girl going a picnic and the picnic the is very happy'. Despite the failure to produce a meaningful coordinate unit, the learner attempts to produce subordination for the first time: 'for the dog has eat the picnic'. Even though this unit contains mistakes, the learner succeeds, in general terms, in producing an adverbial clause. Again, the emergence of subordination seems to lead to a decrease in the percentage of coordinate units produced by the learner as Figure 6.50 shows. At T3, the learner produces six units, four of which are coordinated: 'His mum he's got the map but the dog he's eat [/] eat [//] ate the picnic' and 'In the mountain he's with the dog and the dog eat [//] is ate the picnic'. No instances of subordination are found at this time. At T3, all the utterances in the learners' output are meaningful and contain SVO structures. As for the use of discourse markers, the only instance found was the use of 'this' to refer to one of the pictures at T1. In regard to

student D's syntactic complexity, as seen from Table 6.153, the learner already produces four coordinate units at T0: 'The mum of the boys looking the map and the dog looking the food' and 'The boys it is sad and the basket no [/] no food'. In addition to that, the learner also produces two complete units containing SVO order: 'The dog eating of the food of the boys' and 'The boys we say bye bye at the mum'. The learner makes an attempt to produce another coordinate unit 'and I go to the', but he fails to complete it. At T1, student D uses coordination in three utterances: 'The boy and girl I going to [//] I cooking and [/] and the mother a map', 'A boy and girl going a mountain and look at the dog in the [/] in the food' and 'A dog look inside the basket and yours mother say bye bye'. In addition to these six units, the learner also produces the following unit containing SVO structure: 'The dog is [//] the dog eat the [/] the food and the cheese'. At T2, the instances of coordination found are: 'Two boys preparing a [/] a food and her [/] her mum take a map' and 'And hers dog eating a food of the boys and boys [//] they look at the basket'. Additionally, the first attempt to produce a subordinate nominal clause is found at this time: 'They find the dog then eating dog' (meaning they find the dog eating). Finally, at T3, student D produces the following coordinate units: 'A boy and girl go to the excursion and prepare the breakfast' and 'A boys [//] boy and girl go to the mountain and find a dog in a basket'. As seen from the table, no attempts to produce subordination are found at T3. Finally, the only use of discourse markers is observed at T0 with the phrase 'In here'.

As for lexical diversity, student C seldom uses verbs at T0. The only verbs he uses are: 'do', 'put' and 'have got'. As for nouns, the learner mentions almost all the important features in the story already at T0: 'girl', 'boy', 'mother', 'dog', 'picnic' and 'mountain'. At T1, the learner also produces the verbs 'prepare', 'say', 'eat' and 'go'. Regarding nouns, he uses the word 'food' this time as well. At T2, the learner uses the

verb ‘be’ for the first time in the utterance ‘The picnic is very happy’, in which he also uses the only adjective found throughout the four data collection times. At this time the word ‘map’ is also introduced. At T3, no new verbs are found. As for student D, he uses the verbs ‘look’, ‘say’, ‘go’, ‘eat’ and ‘be’ already at T0. At this time the only adjective found is also produced ‘The boys it is sad’. Regarding nouns, Table 6.153 shows that the learner already mentions the key elements from the story at T0: ‘mum’, ‘boys’, ‘map’, ‘dog’, ‘food’ and ‘basket’. At T1, the learner adds the following verbs and nouns: ‘cook’, ‘eat’, ‘girl’, ‘mountain’ and ‘cheese’. At T2, the learner also employs more sophisticated verbs: ‘prepare’, ‘take’, ‘put’ and ‘find’. Finally, at T3, the new words used by the learner are ‘excursion’, ‘breakfast’ and ‘give’.

Regarding accuracy, student C does not produce any error-free units at T0. At this time, the output produced by the learner is fragmented and the utterances are characterised by consisting of isolated content words. The most common mistakes found at this time relate to the lack of either the auxiliary verb or the main inflected verb: ‘A girl and a boy doing a picnic’. Interestingly, the learner corrects himself at this time when realizing that he has not used a verb: A girl and a boy a picnic [//] doing a picnic’. At T1, one error-free unit is found: ‘The boy and the girl prepare a picnic’. At this time, the learner uses more verbs, however, he fails to complete the utterances by providing (correct) verb complements: ‘And this they go [/] they go the xxx boy’ and ‘They say the man’. In addition, the learner does not succeed in marking the verb for subject verb agreement: ‘The dog eat the food’. At T2, no error-free units are found either. At this time, even though the learner marks the verb for subject verb agreement, he makes lexical errors by selecting wrong verbs: ‘The dog is food the picnic’ and ‘Your mum has a map and the dog has a picnic’. Errors related to verb forms are more complex at this time, as the learner is attempting to use more sophisticated verb tenses:

‘for the dog has eat the picnic’. At T3, the learner produces one error-free unit, which was the same as at T1: ‘The boy and girl prepare a picnic’. Again, at this time the learner makes mistakes related to the lack of auxiliary verbs: ‘The girl and the boy going to the mountain’. Additionally, instances of self-corrections are found in relation to the use of verbs: ‘but the dog he’s eat [/] eat [//] ate the picnic’ or ‘and the dog eat [//] is ate the picnic’. As for student D, no error-free units are found throughout the four data collection times. At T0, errors due to the use of formulaic language are found: ‘The boys we say bye bye’, ‘and I go to the’ and ‘The boys it is sad’. As seen from these examples, the learner uses a subject pronoun before the verbs. In the first example, it is likely that the learner learnt the expression ‘we say bye bye’ within the context of the classroom as a chunk. Likewise, the second example contains the subject pronoun ‘I’ before ‘go’ which suggests the use of ‘I go to’ as formulaic language. Finally, the structure ‘it is’ found in the last example is also used as an unanalysed chunk of language, as seen in previous descriptions. In addition to the use of formulaic language, errors regarding the lack of auxiliary verbs are also common: ‘The mum of the boys looking the map and the dog looking the food’, ‘A boy and a girl going a mountain’ and ‘In here the dog eating of the food of the boys. At T1, the use of formulaic language is still present: ‘The boy and the girl I going to I cooking’. In these utterances, the learner uses the subject pronoun ‘I’ before the verbs. Mistakes when using the present progressive or marking the verb for subject verb agreement are also common: ‘A boy and a girl going a mountain, ‘A dog look inside’ ‘Yours mother say bye’. At T2, the same errors are still found: ‘Two boys preparing a food’ and ‘And hers dog eating a food’. This time, however, the verb ‘look’ is used correctly for the first time: ‘They look at the basket’. Finally at T3, no use of formulaic language is found. Besides, errors related to verb forms are not so common, although some are still present: ‘Hers mother

give a map' or 'The dog eating the breakfast'. Additionally, an instance of self-correction is found related to the production of SVO structures, which indicates that the learner has interiorised English word order, as most of the utterances at this time show: 'And a breakfast a dog eating [//] the dog eating a breakfast'.

Finally, in regard to the learners' fluency levels, student C's evolution displayed in Figures 6.50 shows that speech rate in words remains constant from T0 to T2, while it increases from T2 to T3. As seen in Table 6.152, T3 is the time when the learner expresses his ideas more clearly and without hesitation, which might result in a reduction of total task time (TTT) needed by the learner. Thus, even though the learner uses the same amount of language at T2 and T3, the learner's output seems to be organised with fewer repetitions, which might have a positive effect on speech rate in words. The same tendency is found for student D. From T0 to T2, his fluency scores do not show much variability, whereas from T2 to T3 the learner improves considerably. As seen in student C's output, student D also seems to produce a more coherent and organised description of the events, which in combination with the decrease of TTT may account for the increase in speech rate in words. As for L1 use, no instances of L1 words are found in either of the samples.

In brief, the two fluent learners' evolution of CAF is similar to that found in the output produced by the two highly fluent learners, inasmuch as improvement is seen in regard to the use of coordination and subordination. Despite this, the two fluent learners' improvement takes place at a slower rate. As for accuracy, it is worth highlighting that no instances of mistakes related to the use of formulaic language are found in the fluent CLIL learner's output, while part of the mistakes produced by the control learner at T0 and T1 are associated to the use of unanalysed chunks of language. As for fluency, both learners show a considerable increase from T2 to T3.

6.2.1.2.3 Dysfluent learners from the CLIL and control groups

Figures 6.52 and 6.53 display the evolution of CAf of the two dysfluent learners. As seen from Figure 6.52, student E, the dysfluent CLIL learner, does not produce any coordinate units until T3. As for accuracy, no error-free units are found at any of the four data collection times. His fluency levels, on the other hand, do show improvement from T0 to T3. Regarding student F, the dysfluent control learner, no instances of coordinate or error-free units are found at T0. From T0 to T3, however, these two measures improve. The percentage of coordinate units improves considerably from T0 to T1 and continues increasing gradually until T3. The percentage of error-free units also undergoes a slight but constant increase up to T3. In regard to his fluency levels, student F improves gradually from T0 to T3. Tables 6.154 and 6.155 show the output produced by the two learners in the narrative task.

Figure 6.52
Evolution of CAF in the narrative task. Dysfluent CLIL learner (student E).

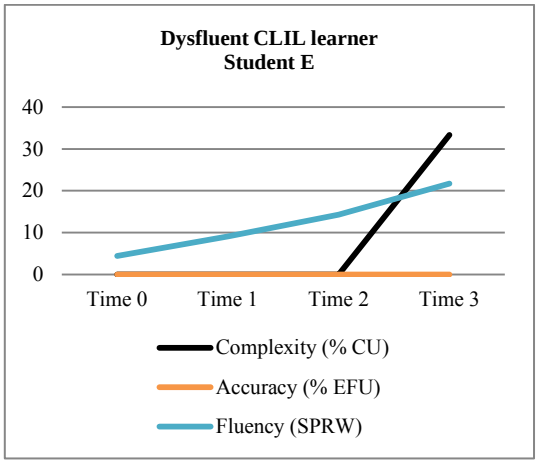


Figure 6.53
Evolution of CAF in the narrative task. Dysfluent control learner (student F).

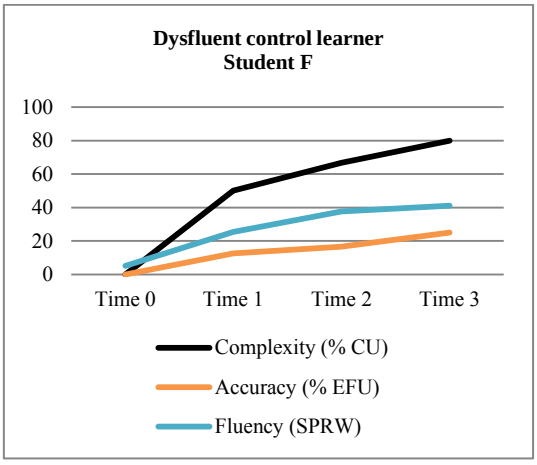


Table 6.154
Dysfluent CLIL learner. Student E's output in the narrative task.

Time	Learner's output
Time 0	Mum and brother [/] mum a brother a picture xxx brother. La@s:c cinc@s:c? Brother is [/] is +... brother and dog [/] and dog. brother dog i@s:c +...
Time 1	One at xxx and aliments@s:c in cest@s:c. Mother xxx. Map. Dog. Dog is the xxx. girl and boy goodbye a [/] a mother and +... ++ and [/] and farmer. Five dog is +... Dog is amb@s:c +... Cake. En@s:c el@s:c gos@s:c.
Time 2	Sister and brother. In cake of the sister no xxx. And two bueno@s:c [/] two mum look at the map and dog is +... Mum no [//] sister and brother. This camp@s:c and mum goodbye. Sister and brother is in the camp@s:c. Five dog this running in the camp@s:c and sister and brother in surprended@s:c. Six dog is xxx to the [/] to the picnic.
Time 3	One picture cut the [//] sister and brother sister cut the cake and brother pos@s:c the cake in the cistell@s:c . A two picture [/] two picture dog into the [/] the cistell@s:c ai@s:c Mum sister and brother look the map. Picture three mum[/] mum goodbye the sister and brother and [/] and y@s:c esta@s:c. Picture cuatro@s:c [//] four sister and brother com@s:c era@s:c ? Picture five ? Dog [/] dog escape the cistell@s:c and [/] and sister and brother xxx . Picture six sister and brother look a no cake and dog is running for the camp@s:c.

Table 6.155
Dysfluent control learner. Student F's output in the narrative task.

Time	Learner's output
Time 0	The boy and sister van@s:c de@s:c picnic mon [/] mon [//] mountain. And the dog s'ho@s:c menja@s:c. It tristos@s:c.
Time 1	The boy and the girl go to the picnic and the two the dog eat the [/] the bread when his mum look the map and his boy and girl. Boy and girl walk in the forest and go to the mountain and his dog stay in the [//] in the packet. And boy and girl look to the it is no [//] boy and girl look so the dog eat the bread.
Time 2	there is a [/] a boy and girl and he his mum and his dog. His dog put in the cistell@s:c and go to the [/] the [/] the park. The boy and the girl stay in the [/] in the park and [/]and the dog eat the sandwich . His dog is happy but the girl and the boy is unhappy.
Time 3	The boy and the girl go to the picnic and the dog eat the [/] eat the [/] eat the sandwich when the mum els@s:c explica@s:c el@s:c map. And then the girl and the boy go at the mountain and the girl and the boy look at the dog is in the [/] is in the [/] the cesta@s:c and look that the dog ate the sandwich.

Table 6.154 shows that student E, the dysfluent CLIL learner, does not produce any coordinate units at T0 due to the fact that his speech is fragmented and made up of content words mainly, as seen in the following utterances: 'Mum and brother' and 'Brother and dog'. Additionally, he does not produce verbs within complete sentences. There is only one utterance in which the learner attempts to produce the SVO structure: 'Brother is +…'. However, as seen from the sample, he fails to complete the sentence. At T1, the learner's output does not seem to progress in regard to syntactic complexity, as the type of output produced resembles the one produced at T0: 'Mother', 'Map', 'Dog is the xxx', 'Girl and boy goodbye'. The most significant changes in student E's output are visible at T2 with the use of verbs, the emergence of SVO structures and attempts to produce coordination. The following utterances exemplify the use of verbs: 'Two mum look at the map', 'Sister and brother is in the *camp* (countryside)' and 'The dog this running'. Additionally, the learner attempts to coordinate units, however, his limited vocabulary prevents him from completing the utterance: 'And two [/] two mum look at the map and the dog is +…' and 'Five dog this running in the *camp* and sister and brother in *surprended*' (surprised). At T3, the learner shows further syntactic development in the consolidation of SVO word order and the attempts to produce coordination and subordination. Examples of utterances containing SVO structures are: 'Sister and brother cut the cake' and 'mum, sister and brother look the map'. As for coordination, he uses coordination in 'Picture six sister and brother look a no cake and dog is running for the *camp*'. Interestingly, attempts are also found to use subordination in the utterance 'Sister and brother look a no cake'. In this utterance the learner seems to be attempting to produce a nominal clause (sister and brother see there is no cake). However, his linguistic resources are still too limited. Concerning the use of discourse markers, the learner uses words like 'one' and 'five' to refer to the number of the

pictures at T1 and T2, while at T3 the learner uses phrases like 'one picture', 'picture three' or 'picture six'. Student F's (dysfluent control learner) syntactic complexity undergoes surprising improvement from T0 to T1. At T0, his utterances consist of content words accompanied by definite articles mainly: 'The boy and sister picnic mountain' and 'And the dog'. At T1, in contrast, the learner manages to produce coordinate units containing SVO structures: 'The boy and the girl go to the picnic and the two the dog eat the [/] the bread [/] the bread' and 'Boy and girl walk in the forest and go to the mountain and his dog stay in the [//] in the packet'. In this last utterance the learner succeeds in coordinating two verb phrases and two clauses. Additionally, student F attempts to produce a nominal subordinate clause: 'boy and girl look so the dog eat the bread' meaning that the boy and the girl find out that the dog has eaten the bread. At T2, greater syntactic development is observed through the use of the coordinating conjunction *but* in addition to *and*: 'His dog is happy but the girl and the boy is unhappy' and 'The boy and the girl stay in the [/] in the park and [/] the dog eat the sandwich'. Finally, extensive use of coordination is observed along with attempts to produce subordination at T3: 'The boy and the girl go to the picnic and the dog eat the [/] eat the [/] eat the sandwich' and 'And then the girl and the boy go at the mountain and the girl and the boy look at the dog is in the [/] is in the *cesta* and look that the dog ate the sandwich'. This last utterance contains another attempt to produce a nominal clause: 'The girl and the boy look at the dog in the [/] is in the *cesta* and look that the dog ate the sandwich' (meaning the girl and the boy see the dog in the basket and see that the dog ate the sandwich). Likewise, the learner also attempts to produce an adverbial clause in 'when the mum *explica el map*' (meaning when the mum 'explains' the map). In regard to his use of discourse markers, 'And then' is the only instance found at T3 indicating considerable discourse development.

As for lexical diversity, student E produces content words to refer to key elements in the story in addition to the verb *be* at T0*:* 'mum', 'brother', 'picture', 'dog'. At T1, the tendency is similar. However the learner produces new words like 'goodbye' and 'farmer'. At T2, the first verbs other than 'be' appear: 'look' and 'run'. Additionally, at this time the learner attempts to produce the adjective 'surprised' by saying: 'and sister and brother in *surprended*'. At T3, new verbs are introduced: 'cut' and 'escape'. Student F produces the same type of words at T0 as student E: 'boy', 'sister', 'picnic', 'mountain' and 'dog'. At T1, however, the learners' vocabulary undergoes considerable growth, as seen from the number and types of verbs he uses for the first time: 'go', 'eat', 'look', 'walk' and 'stay'. At T2, the learner uses the same verbs in addition to the verb 'be' and 'put'. Finally, no new verbs are found at T3. In regard to the use of adjectives, the learner uses two adjectives at T2 in the utterance 'His dog is happy but the girl and the boy is unhappy'.

Concerning student E's accuracy levels, at T0 and T1 the learner does not produce any units, as his speech is fragmented and made up of content words mainly. At T2, the first units appear. However, these contain errors related to the lack of subject verb agreement as in 'Mum look at the map' and 'Sister and brother is in the *camp*'. In addition, an error related to the lack of the verb *be* in progressive forms is also seen: 'Dog this running'. At T3, the errors are of the same sort as at T2: 'Dog [/] dog escape'. Additionally, as seen in the samples provided up to now, the learner shows problems with the use of articles. As for the use of the verb 'be' with progressive verbs forms, the learner succeeds in using it in the following unit: 'Dog is running'. Student F does not produce any units at T0 either. At T1, the number of units increases and so does the number of error-free units. The learner produces one error-free unit at T1: 'The boy and the girl go to the picnic'. The most common mistakes found at this time are associated

with the lack of subject verb agreement as in 'The dog eat the bread', 'when his mum look the map', 'his dog stay in the packet' and 'dog eat the bread'. In addition, lexical errors are also found when the learner uses the word 'packet' to refer to *basket*. Finally, another common mistake is the absence of the definite article in utterances like 'Boy and girl walk in the forest'. At this time, the learner also makes errors related to possessive adjectives, as he uses 'his' when *their* should be used. This mistake is still present at T2: 'His dog is happy' or 'There is a boy and girl and he his mum and his dog'. Another common mistake at T2 is, as seen at previous times, the lack of subject verb agreement: 'The dog eat the sandwich' or 'The girl and the boy is unhappy'. At this time, one error-free unit is produced: 'The girl and the boy stay in the park'. At T3, mistakes related to the lack of subject verb agreement are still found. A lexical error is also detected when the learner uses the verb 'look' instead of *see* in the utterance: 'and look that the dog ate the sandwich'. The learner also makes mistakes with the choice of preposition: 'The girl and the boy go at the mountain'. Besides, an attempt is made to use the right form of the verb *eat* in: 'and look that the dog ate the sandwich', in which the present perfect was the target verb form. Finally, it is worth noting that the learner produced one-error free unit at this time: 'The boy and the girl go to the picnic', which was also produced at T1.

Regarding fluency, the sample provided in Table 6.15 4 shows a constant increase in the number of words in English used by student E from T0 to T3. At T0 the learner produces isolated words. At T1, the learner continues producing isolated words, but this time the amount of language produced is greater. At T2, the first units are seen, although the learner's speech is still quite fragmented. Finally, at T3 the units become longer, which has an impact on the total number of English words produced. As for L1 use, it is important to highlight that the number of L1 words the learner produces from

T0 to T3 increases. At T0, three L1 words are found: '*La cinc*? (number five?) and '*i*' (and). At T1, the L1 words found relate to words or sentences needed to tell the story: '*cest*' (basket) and '*en el gos*' (in the dog). At T2, three instances of the word countryside are found in the L1 ('*camp*') in addition to '*bueno*' (well), which is used after a false start. Finally, at T3, the learner uses the L1 to refer to words he does not know in English. However, that does not prevent him from continuing the sentence in the L2, as seen in this example: ' Brother *posa* (puts) the cake in the *castel*'. Additionally, he also uses the L1 to generate metatalk: '*com era*?' (what was it?). Finally, the L1 is also used as a discourse marker '*y ésta*' (and this). Regarding student F's fluency levels, the learner shows a constant increase in the number of English words from T0 to T2 while from T2 to T3 the amount of language is very similar. At T0, the learner's speech contains isolated words while at T1, in contrast, the learner produces units, which indicates that the amount of language is greater. Likewise, the number of English words continues increasing at T2 and T3. Concerning L1 use at T0, the learner uses the words '*van de*' (go for) and '*s'ho menja*' (it eats it) within an utterance in English: 'The boy and the sister van de picnic' or 'And the dog *s'ho menja*'. The same is found at T3: 'when the mum *els explica el* (explains the) map'. Finally, the words *cistell* (basket in Catalan) and *cesta* (basket in Spanish) are used at T2 and T3.

To summarise, the two dysfluent learners display improvement in the three CAF measures selected when undertaking a qualitative analysis of the output produced. Despite this, student F's (dysfluent control learner) development is more visible throughout the two years of the study, especially in syntactic complexity. Regarding accuracy, improvement is also observed in some areas for both students. In terms of fluency, both learners show progress, although student F seems to display greater improvement. In the next section, the output produced at the four times of data

collection in the narrativetask by the six learners selected will be categorised in a series of stages of oral narrative discourse (Álvarez, 2006) in an attempt to further examine their development.

6.2.2 Stages of oral narrative discourse

This section aims at analysing the progress displayed by young learners in the development of oral narrative performance using the stages of narrative discourse identified by Álvarez (2006) for the narrative task used in this piece of research. In her study, the author analysed the morphological, syntactic and discourse development in the narrative performance of school-learners who started learning English at different ages and two groups of adults who had received two and three years of English instruction with the aim of identifying a set of stages learners progress through. In this dissertation, the output produced by the six learners selected will be classified according to the stages proposed by Ávarez (2006) shown in Table 6.156.

Table 6.156
Summary of Álvarez's (2006) stages of oral narrative discourse.

Stage	Linguistic and discourse features	Language samples from Álvarez (2006)
Stage 1	The narrative is completed in the L1.	
Stage 2	Emergence of bare nominal content words. Use of an L1 determiner.	*Child, boy, dog* *El [the] picnic, un [a] dog*
Stage 3	Use of the L2 definite article or an ordinal number. Emergence of plural –s. Emergence of the first prepositions, usually accompanied by an L1 word. Emergence of first prepositional phrases.	*The child, the boy, the mother* *The boys, two boys, two brothers* *In the cesta (basket)* *At basket*
Stage 4	Further development of prepositional phrases. Use of one or several verbs. Emergence of SVO and SVA structures. Emergence of –ing forms. Emergence of the coordinating conjunction *and*. Use of possessive adjectives.	*In the room* *In the mountain* *Is a mountain* *The dog look* *Prepare sandwich* *The brother and sister no breakfast* *The cows eating* *Look a basket and play.* *Your mother*

Stage 5	SVA structures appear in their complete form. Emergence of the present progressive. Bare lexical verbs lacking inflectional markers. Use of a broader variety of prepositions: *into, on, under* or *for.* Incipient discourse organisation.	*The children is at the mountain* *In the basket is a dog* *The family is preparing the sandwich* *In the second (picture)*
Stage 6	Emergence of S V, S VC, S VOO and S VOA structures. The conjunction *but* may occur. The third person present singular *–s* and future *will* may occur. Use of L2 indefinite article.	*The mum helps the children* *The mum read the map* *They will going to the mountain* *She look a dog but don't have a any dinner*
Stage 7	Use of post-nominal modification in the form of prepositional phrases. Use of temporal discourse markers such as *then* or *after.* Emergence of adverbial and nominal clauses.	*A place for a picnic* *When the children open the basket, they see that the dog eat the food.*
Stage 8	Use of modals and infinitive constructions. Emergence of *have* as a perfect auxiliary occurs. Use of the conjunction *while*. Subordination becomes more complex.	*The children must go* *They want to go to breakfast.* *The dog has eat the sandwiches* *While the children are watching the plan, the little dog is watching something.* *After, her mother teaches him where is the road that he have to take you go to the mountain.*
Stage 9	Complex embedding, which may include the use of relative clauses. Grammatical aspect alternations.	*There is the dog who is searching for some sandwiches.* *The get out of the house and they began to walk.*

As seen from the table, Álvarez's (2006) stages of narrative discourse development trace the progress learners undergo in terms of morphosyntactic aspects mainly. The stages the author proposes move from the use of the L1 to the use of L2 complex embedding structures like relativisation, moving through stages in which different verb type structures and coordination emerge and consolidate.

In the next tables, the output of the learners selected for the qualitative analysis will be classified using the linguistic structures and features proposed by Álvarez (2006). Tables 6.157 and 6.158 present the main stages identified in the language produced by highly fluent CLIL and control learners in the narrative task at the four data collection times.

Table 6.157
Oral narrative stages in the highly fluent CLIL learner's output.

	Stage	Language features	Language samples
Time 0	5	Complete SVA structures Present progressive Prepositions Incipient discourse organisation	*The boys are in the mountains* *The dog is looking in the basket* *This boys are cooking* *The dog is go out* *In the picture four*
Time 1	5	Present progressive Incipient discourse organisation	*The boy is going camp@s:c* *In picture three*
	6	Third person singular -s	*The dog go [//] goes in [//] goes on +...* *The dog goes out*
Time 2	6	Third person singular -s	The boy has to prepare a picnic
	7	Emergence of adverbial clauses	*When the boys look into the map, the dog goes into the cest@s:c*
Time 3	8	Infinitive constructions	*Mum tells him to go to the park to eat breakfast*

Table 6.158
Oral narrative stages in the highly fluent control learner's output.

	Stage	Language features	Language samples
Time 0	4	SVO and SVA structures Conjunction *and* Use of possessive adjectives	*His mum explain the route* *The two boys go to the forest and his dog eat the food* *His dog eat the food and run in the forest*
Time 1	5	Present progressive Incipient discourse organisation Bare lexical verbs lacking inflectional markers Complete SVA structures	*The girl and the boy is doing the breakfast* *In the picture number one* *He have a map* *They arrive in the forest.*
	6	Emergence of SVOO structures Third person singular –s Use of L2 indefinite article	*He says bye bye for him mum* *He have a map*
	7	Emergence of nominal clauses	*They see that in the nest do not have the breakfast.*
Time 2	6	Emergence of SVOO structures	*They give him a map*
	7	Use of post-nominal modification in the form of prepositional phrases Emergence of nominal clauses	*They prepare sandwich for a picnic.* *The dog go to run.*
Time 3	7	Use of post-nominal modification in the form of prepositional phrases Emergence of nominal clauses	*A map of the world* *They see that the breakfast isn't there*

As seen from Tables 6.157 and 6.158, both learners progress through three stages in the course of two academic years (from T0 to T3). However, the learners' initial stage differs, as the highly fluent CLIL learner starts from a higher stage than the highly fluent control learner. At T0, the output produced by the highly fluent CLIL learner contains most of the linguistics features within stage 5: complete SVA

structures, the emergence of the present progressive, use of prepositions other than *in* and traces of incipient discourse organisation. At T1, the amount of output produced by this learner decreases when compared to T0 (see Table 6.150), however, language features from stages 5 and 6 are found. The use of discourse markers and the present progressive belong to stage 5, while the third person present singular –s is used for the first time: 'The dog goes out'. This is the only feature from stage 6 found in the output at T1. The use of the third person present singular -s is also used at T2 along with the emergence of adverbial clauses, one of the language features from stage 7. Finally, at T3, the learner uses infinitive constructions, linguistic structure which belongs to stage 8, along with features that belong to previous stages (present progressive, third person present singular –s and adverbial clauses among others). No other linguistic features from stage 8 are found in the learner's output at T3.

The oral output produced by the highly fluent control learner over the four times provides a higher number of features used to characterise the stages. At T0, for instance, the learner's output contains features from stage 4 such as the emergence of SVO and SVA structures, the use of *and* and the use of possessive adjectives. At T1, the learner's output contains linguistic features from stages 5, 6 and 7. At this time, the learner seems to have achieved stage 5 as most language features that belong to this stage are present in the learner's output: the use of present progressive forms, discourse markers, SVA structures and bare lexical verb lacking inflectional markers. Despite this, features from stage 6 such as the emergence of SVOO structures, the use of the third person present singular -s and the L2 indefinite article are also found: 'He says bye bye for him mum' and 'He have a map'. Additionally, the learner also produces his first nominal clause, indicating that the learner is starting to produce features from stage 7. At T2, the learner uses SVOO structures, which are typical from stage 6, as in 'They give him a map' as

well as features from stage 7 like the use of post-nominal modification and the emergence of nominal clauses. At T3, the learner produces the same linguistic structures from stage 7 as at T2: the use of post-nominal modification ('A map of the world') and the use of nominal clauses ('They see that the breakfast isn't there').

To sum up, it is important to note that both learners attain stage 7, which is characterised by the emergence of nominal and adverbial clauses. As for the comparison of the development displayed by the CLIL and control learners, one of the main differences is that the control learner seems develop faster through stages from T0 to T1, producing features from stages 5, 6 and 7, whereas the CLIL learner's development takes place at a faster rate from T2 to T3, time period in which features from stages 6, 7 and 8 are found. In tables 6.159 and 6.160, the stages identified in the output of the two fluent learners are displayed.

Table 6.159
Oral narrative stages in the fluent CLIL learner's output.

	Stage	Language features	Language samples
Time 0	3	L2 definite article First prepositions and prepositional phrases	*The dog* *In the mountain*
	4	Emergence of -ing forms	*The dog putting a picnic*
	6	Indefinite L2 article	*A picnic*
Time 1	4	Emergence SVO structures	*The boy and the girl prepare a picnic* *The dog eat the food*
Time 2	4	-ing forms Conjunction *and*	*The boy and the girl preparing the picnic* *And your mum has a map and the dog has a picnic*
	6	Third person present singular –s	*And your mum has a map*
	8	Emergence of *have* as a perfect auxiliary occurs.	*The dog has eat the picnic.*

Time 3	4	-ing forms Conjunction *and*	*The boy and the girl going to the mountain and in the mountain he's with the dog.*
	5	Complete SVA structures	

Table 6.160
Oral narrative stages in the fluent control learner's output.

	Stage	Language features	Language samples
Time 0	4	Emergence of -ing forms Emergence S VO s tructures Conjunction *and*	*The mum of the boys looking the map and the dog looking the food.*
	6	Emergence of SVOO structures	*We say bye bye at the mum.*
Time 1	4	-ing forms SVO structures	*A boy and a girl going a mountain.* *The dog eat the food and cheese.*
	5	Bare l exical v erbs l acking inflectional markers Complete SVA structures	*A dog look inside the basket* *A boy and girl going a mountain*
	6	Use of L2 indefinite article	
Time 2	4	-ing forms SVO structures Use possessive adjectives	*Two boys preparing a food* *Hers dog eating a food.*
	5	Complete SVA structures	*A dog he's put in the basket.* (Intended f orm: T he d og g ets i n t he basket)
	6	Use of L2 indefinite article	*And her mum take a map.*
Time 3	4	Conjunction *and* SVO and SVA structures	*A boy and girl go to a excursion and prepare the breakfast.*
	6	Use of L2 indefinite article	*Hers mother give a map.*

As seen from Table 6.159, the fluent CLIL learner produces features which belong to stages 3, 4 and 6 at T0. The features from stage 3 are the use of the definite article and the emergence of prepositional phrases. In addition, the learner also uses the –ing form the verb 'put', which is a typical feature from stage 4. Finally, the learner uses the indefinite article to mark first mention of inanimate nouns ('a picnic'), which according to Álvarez (2006) characterises stage 6. At T1, the emergence of SVO structures are found, which is a typical feature from stage 4. At T2, the learner uses linguistic features from stage 4 (-ing forms and the conjunction *and)*, stage 6 (third person present singular –s) and, surprisingly, from stage 8 (the emergence of *have* as perfect auxiliary in 'The dog has eat the picnic). Eventually, at T3, the output produced by the learner contains features from stage 4 and 5.

In regard to the fluent control learner (see Table 6.160), the output produced at T0 contains features from stage 4 and 6. The linguistic features from stage 4 are the emergence of SVO structures, the production of -ing forms and the use of the conjunction *and.* Additionally, the emergence of SVOO verb structures, typical of stage 6, is also found in the utterance 'We say bye bye at the mum'. At T1, the learner produces linguistic features from stages 4 (as at T0), from stage 5 (the use of bare lexical verbs lacking inflectional markers and complete SVA structures) and from stage 6 (the use of the indefinite article for first mention of inanimate nouns: 'a mountain'). At T2, a new feature from stage 4 is introduced: the use of possessive adjectives. Besides, -ing forms, (typical from stage 4) and complete SVA structures (typical from stage 5) are also found. Finally, the learner continues using the indefinite article, feature which belongs to stage 6. At T3, features belonging to stages 4 and 6 are found again. No new linguistic features or structures are found at this time in comparison to previous times.

Altogether, as seen from Table 6.159, the fluent CLIL learner does not seem to develop through stages in an orderly way, as he produces language features which are one or two stages above his current level. Furthermore, considering the language features produced at T3, it looks as if the learner displays backward development when compared to the language produced at T2. As for the fluent control learner, no development seems to be found from T0 to T3, according to Álvarez's stages of narrative development, as his output contains features from the same stages (4 and 6) at the four data collection times. As for the development of dysfluent learners, Tables 6.161 and 6.162 present their progress through the stages from T0 to T3.

Table 6.161
Oral narrative stages in the dysfluent CLIL learner's output.

	Stage	Language features	Language samples
Time 0	2	Emergence of bare nominal content words	*Mum and brother* *Brother and dog*
Time 1	2	Bare nominal content words	*Mother* *Map* *Dog*
	3	Emergence of the first prepositions, usually accompanied by an L1 word	*In cesta*
Time 2	4	Emergence of SVA structures Coordinating conjunctions *and*	*Mum look at the map and dog is in the+…*
Time 3	5	Bare lexical verbs without inflectional markers Broader variety of prepositions Emergence of the present progressive	*Dog escape* *Dog into the cistell.* *Dog is running*

Table 6.162
Oral narrative stages in the dysfluent control learner's output.

	Stage	Language features	Language samples
Time 0	3	Use of the L2 definite article	*The boy and sister* *And the dog*
Time 1	4	Emergence SVO structures Use possessive adjectives Conjunction *and*	*The dog eat the bread* *His dog* *The boy and the girl go to the picnic and the two the dog eat the bread.*
	5	Bare lexical verbs lacking inflectional markers	*His dog stay in the packet* *The dog eat the bread*
	6	Emergence of SV structures	*Boy and girl walk in the forest*
	7	Emergence of adverbial clauses	*The dog eat the bread when his mum look the map*
Time 2	4	SVO structures Use possessive adjectives	*The dog eat the sandwich* *His dog is happy*
	5	Complete SVA structures	*The boy and the girl stay in the park*
Time 3	4	Conjunction *and*	*The boy and the girl go to the picnic and the dog eat the bread*
	5	Complete SVA structures	
	7	Adverbial clauses Use of temporal discourse markers	*The girl and the boy look that the dog ate the sandwich* *And then the girl and the boy go to at the mountain*

As Table 6.161 displays, the output produced by the dysfluent CLIL learner at T0 contains the main linguistic feature from stage 2, which is the emergence of bare nominal content words. At T1, the learner still produces this feature, but this time he produces the first prepositions as well, which characterises stage 3. At T2, the first SVA structures and the coordinating conjunction *and* are used for the first time, which are

typical from stage 4. Finally at T3, features from stage 5 such as the use of verbs without inflectional markers, the use of a wider variety of prepositions and the emergence of the present progressive are found in the output produced by the learner.

Regarding the dysfluent control learner, the learner's output is categorised at stage 3 at T0, as the learner uses the L2 definite article. At T1, the learner produces features from stage 4 (the emergence of SVO structures, the use of possessive adjectives and the use of the conjunction *and*), from stage 5 (the use of verbs lacking inflectional markers), from stage 6 (the emergence of SV structures: 'Boy and girl walk in the forest') and from stage 7 (the emergence of adverbial clauses: 'The dog eat the bread when his mum look the map'). At T2, the output contains SVO structures and possessive adjectives, which characterise stage 4, and complete SVA structures, typical from stage 5. Finally, at T3, features from stage 4 and 5 which the learner has already produced at previous times are also found. Besides, the learner produces features from stage 7: the use of adverbial clauses along with the use of temporal discourse markers in the utterance 'And then the girl and the boy go at the mountain'.

In the case of the two dsyfluent learners selected, the CLIL learner progresses through 3 stages from T0 to T3 while the control learner moves through 4 stages. Additionally, their initial stage differs, as the CLIL learner starts from stage 2 whereas the control learner's output has features from stage 3 at T0. As for their progress through stages, the CLIL learner's output seems to develop in a more organised manner by producing features from stage 2, 3, 4 and 5 in the course of the two academic years. The control learner, in contrast, seems to skip stages by producing features from stage 7 without having attained most of the features from stage 6.

In general terms, the learners selected for this qualitative analysis develop through Álvarez's stages of narrative development. However, as seen in the description

provided here, learners might progress through one stage without having fully achieved the features from the previous stage. Furthermore, learners may produce features from different stages for the first time at the same data collection, indicating that the move from one stage to the other does not always take place in a linear way.

Altogether, the qualitative analyses presented in this chapter have attempted to provide a more detailed picture of the characteristics and development of YLLs' output in CLIL+EFL and EFL instructional settings over a period of two academic years. The most relevant findings derived from these descriptions will be further developed in the next chapter in an attempt to complement and interpret the results obtained from the quantitative analyses.

CHAPTER 7 Discussion

The data analyses conducted in this study aimed at exploring the effects of CLIL on young learners' oral production skills in terms of complexity, accuracy and fluency over a period of two academic years. Thus, in the light of the results obtained from both quantitative and qualitative data analyses, a discussion of the hypotheses proposed in the Introduction will be presented.

This chapter consists of four main sections. Section 7.1 presents a summary and discussion of the achievement and development results obtained from the interview and narrative tasks in relation to Hypotheses 1 and 2. In section 7.2, the results of the inter and intragroup comparisons in the two tasks will be summarised and dealt with in relation to the effects of proficiency level to explore Hypotheses 3 and 4. In section 7.3, the results obtained from the correlation tests will also be considered in regard to Hypothesis 5. Finally, a discussion of the description of the output analysed in the qualitative part of this study will be presented in section 7.4 to investigate Hypothesis 6.

7.1 Achievement and development results

The most striking finding in this study, when compared to previous CLIL research, is the lack of significant differences in fluency in favour of CLIL learners after two years of CLIL instruction. The only significant differences in favour of CLIL learners were in the percentage of nouns produced at T0, T1, T2 and T3 in the interview task and at T0 and T1 in the narrative task. However, the fact that CLIL learners already obtained significant differences at T0 suggests that CLIL instruction had little impact on the differences found at T3 in this measure.

One of the reasons that could account for the lack of significant results in favour of CLIL learners relates to the type of CLIL programme implemented in the schools

selected. As described in the Method chapter, CLIL learners were exposed to 1 hour of CLIL a week, which makes a total of 54 hours of CLIL instruction from T0 to T3 in addition to 191 hours of EFL instruction. This type of CLIL exposure is widely spread in Catalan schools which attempt to introduce CLIL into their academic curriculum. Some of the reasons why schools choose minimal CLIL exposure programmes over more intense CLIL instruction relate to the lack of qualified teachers, the lack of institutional support or parent' concerns. As a consequence, the amount of CLIL input received and accumulated by CLIL learners in these programmes in such a short term (two years) might not have been enough to make a difference between CLIL and non-CLIL groups in terms of oral production skills. Thus, taking these findings into consideration, educational institutions should foster the implementation of more intense CLIL programmes in which learners are exposed to CLIL instruction at least two hours a week in an attempt to further develop their language skills, particularly their speaking skills. As Edelenbos and Kubanek state (2009) factors like 'time' and 'intensity' affect language gains among young learners exposed to bilingual education.

The non-CLIL group, on the other hand, obtained statistically significant results in propositional and syntactic complexity, accuracy and fluency at least once over the four data collection times, despite having received the same amount of English instruction as the CLIL group. Nevertheless, the majority of these differences gradually disappeared at T3 owing to the development CLIL learners underwent. For instance, CLIL learners showed significant development from T0 to T3 in propositional complexity, syntactic complexity and fluency in either the interview or narrative tasks or both. Additionally, CLIL learners also obtained significant improvement in fluency from T2 to T3 in both the narrative and interview tasks. Consequently, CLIL learner's development in these areas led to the disappearance of significant differences in favour

of non-CLIL learners. These results suggest that CLIL does enhance language learning, however, its effects are not to be expected in the short term, as Pladevall-Ballester (in press) and Vallbona (2014) have already pointed out. Furthermore, the high degrees of development found within the CLIL group also reveal that one of the areas which seems to benefit the most from CLIL instruction is fluency, as Dalton-Puffer (2011) and Ruiz de Zarobe (2011) indicate. These results are in line with previous CLIL studies reporting positive effects of CLIL on oral fluency (Järvinen, 2005; Mewald, 2007; Juan-Garau, 2010; Bret, 2011). Interestingly, the fluency measures are the only ones which show significant development during the second year of the study (T2-T3) as well, which provides further evidence to support some of the claims in this study.

The first claim in this study is that language-related outcomes in CLIL are not immediate. As shown in this study, CLIL learners rarely outperformed their counterparts in the non-CLIL group significantly. However, after 54 hours of CLIL instruction (T3), CLIL learners attained similar levels in CAF to those obtained by non-CLIL learners, which suggests that CLIL learners may, in the long run, surpass their peers in the non-CLIL groups. This prediction leads to the second claim, which defends the idea that language-related gains in minimal CLIL exposure will take place after learners have accumulated a reasonable amount of CLIL instruction. As seen in this study, 54 hours of CLIL instruction did not generate significant results for the CLIL group. Thus, greater amounts of CLIL input need to be provided for CLIL to have a significant impact, especially on oral production skills among young language learners. Thirdly, this study also provides evidence to claim that the strongest effects of CLIL are on fluency, since this is the area where learners exposed to CLIL showed the greatest improvement when compared to the non-CLIL group.

Positive effects of CLIL were also observed in syntactic complexity, measured as the percentage of coordinate units, as CLIL learners' development helped them overcome significant differences with the control group. Other pieces of research have also reported a positive impact of CLIL on syntactic complexity (Järvinen, 2005; Mewald, 2007; Bret 2011; Lázaro and García-Mayo, 2012). The fact that CLIL exposure is likely to provide learners with a wider variety of syntactic structures than EFL instruction due to the nature of content matter (science) may account for the development of syntactic complexity in CLIL settings. In this study, CLIL learners showed the highest percentages of improvement in this measure, which could derive from the type of input received, characterised by containing complex language structures to express high-order thinking processes and abstract knowledge (i.e. CALP).

As for lexical complexity, inconclusive results were obtained. As previously presented, the CLIL group produced a higher number of nouns at the four data collection times in the interview task and at T0 and T1 in the narrative task. Additionally, the CLIL group displayed a significant decrease in the use of nouns from T0 to T1 and from T0 to T3. The fact that CLIL learners produced a significantly higher number of nouns does not seem to be a positive result, as according to Broeder, Extra and Van Hout (1993) and Muñoz (2006), a high number of nouns in the learners' speech is an indicator of low vocabulary levels. As for the use of verbs and adjectives, no significant differences were found, which is in line with Jiménez, Ruiz de Zarobe and Cenoz's (2006) findings which reveal that even though CLIL learners unfold more lexical richness and sophistication, CLIL does not seem to have a strong effect on vocabulary production, at least in the production of types and tokens. Altogether, it is also important to note that, as Vermeer (2001) points out, high levels of content-specific vocabulary do not necessarily relate to high levels of lexical complexity.

Additionally, and contrary to what would be expected in meaning-based instruction, CLIL instruction also seems to improve learners' accuracy levels. In this study, CLIL learners usually presented the highest percentages of improvement in the production of error-free units and correct verb forms (in spite of not showing significant development periods in these measures). Similarly, Järvinen (2005), Mewald (2007) Hüttener and Rieder-Bünemann (2007, 2010) and Bret (2011) also reported beneficial effects of CLIL on accuracy. A possible explanation for the improvement of accuracy could relate to the type of input CLIL learners receive. In CLIL lessons, learners have to deal with cognitively demanding knowledge and language in the L2, which may result in higher degrees of attention to language on the part of the learners in order to be able to understand and follow the explanations. As a consequence of the degree of attention given to the input, learners' linguistic awareness might be enhanced, which may affect the learners' accuracy levels in turn. Besides, the fact that CLIL learners were also exposed to EFL instruction, where accuracy is a key element, may also account for their improvement in accuracy. In addition, another possible reason behind their improvement in accuracy might stem from the emphasis given to accuracy (at least in the context of this research project), which seems to prevent young language learners from adopting risk-taking attitudes when producing language in order to avoid errors.

Another relevant factor which could explain the lack of statistically significant results in favour of CLIL learners is the limited amount of oral practice offered to learners in CLIL settings. According to previous CLIL research (Stotz and Meuter, 2003; Mewald, 2007; Wannagat, 2007; Lim Falk, 2008), CLIL lessons rarely focus on the development of oral production skills. The only instances of L2 output on the part of the learners in these studies are found when answering teachers' questions.

Furthermore, the answers provided by the learners in these studies were usually made up of short utterances containing only a few words in the L2.

According to the classroom observation sessions conducted for this study, little use of the L2 was observed on the part of the learners, even though CLIL teachers used English during the whole lesson. Despite that, it is true that CLIL teachers promoted the use of the L2 in oral interactions through questions to assure comprehension. However, the linguistic support provided to the learners to help them develop their oral production skills was quite limited or inexistent. Consequently, learners rarely use the L2 to answer the teachers' questions or to interact with their peers, as was also corroborated in the students' interviews and questionnaires analysed in Pladevall-Ballester (2014). Most activities and tasks in these CLIL lessons consisted in lecturing about a topic, asking some questions to the students, doing tasks or activities to practice vocabulary, writing or reading skills and experimenting. No explicit use of oral tasks to practice content and language (i.e. information gap activity, questionnaires, role-plays) was found in any of the lessons observed, which may explain why CLIL learners did not significantly outperform their non-CLIL counterparts in terms of speaking skills. Additionally, if CLIL learners' oral production skills are to develop, CLIL teachers need to encourage the use of formulaic language, which was not the case in the lessons observed. As previously stated in the section on YLLs' language, young learners need to count on formulaic structures to develop their oral skills and their L2 language system. The inclusion of focus on form tasks in CLIL settings may help CLIL teachers adopt a more integrative approach which actually develops both language and content objectives to the same degree. As described in previous chapters, many researchers (Long and Robinson, 1998; Swain, 1998, Lightbown and Spada, 2000; Ellis, 2005; de Graaff et al., 2007; Lyster, 2007, García-Mayo, 2012; Basterreachea et al., 2014) suggest the

integration of focus on form into meaning-based instruction in order to promote a better and deeper leaning of the target language. In the context of this research, the adoption of a focus on form approach might have helped teachers enhance the development and teaching of formulaic language and speaking skills.

According to the classroom observation sessions conducted, the lack of oral practice seems to be one of the main flaws of CLIL and other forms of bilingual education. Authors like Swain (1985, 1993) and Genesee (1994) already highlighted the need to focus on productive skills to develop the learners' interlanguage system. Furthermore, as de Graaff et al. (2007) pointed out, one of the indicators of successful CLIL programmes relates to the creation of output opportunities for learners. Additionally, according to Ellis (2005), successful instructed language learning requires opportunities for output and interaction through the use of tasks in which learners have to fulfil a communicative objective. The use of this type of tasks is more common in regular EFL instruction than in CLIL settings. CLIL learners in this study have also received EFL exposure (75% of the instruction came from EFL lessons, while the remaining 25% derived from CLIL lessons) which implies that CLIL learners also developed their speaking skills through tasks in their EFL lessons. However, the non-CLIL group, who was exclusively exposed to EFL instruction, received a higher number of opportunities to develop their oral skills by means of interactive tasks and that might be why the non-CLIL group achieved statistically higher results in some of the measures at certain times. The CLIL group, on the other hand, did not attain significant results.

According to the results in this study, CLIL programmes seem to emphasise comprehension over production. As described in Chapter 3, this was precisely one of the shortcomings of content-based approaches (Genesee, 1994; Lyster, 2007). Thus, one

of the reasons behind the lack of significant differences in favour of CLIL learners as far as their oral production skills are concerned could be related to the fact that productive skills are generally overlooked in CLIL settings. This claim seems to be supported by Vallbona's findings (2014). In her study, no significant differences were found between CLIL science learners and control learners in any of the CAF measures used to analyse the learners' writings skills at T3. However, the CLIL science group did obtain significantly higher scores in listening, in comparison to control learners, at T3, which backs up the idea that, as observed in immersion and content-based instruction, productive skills in programmes which use the L2 to teach content matter are underdeveloped, when compared to the development of receptive skills. In brief, these findings support Dalton-Puffer's (2007: 11) claim that "CLIL students are listeners most of the time".

In addition to the type of CLIL programme implemented and the lack of opportunities for oral output, the type of learner analysed in this study is another of the factors that could account for the lack of significant results in favour of CLIL learners. As Housen and Pierrard (2005) and the de Graaff and Housen (2009) indicate, one of the factors that mediates the effects of instruction is the type of learner. As previously presented (Chapter 2), young learners are characterised by being in constant cognitive development and by presenting difficulties when dealing with complex abstract concepts. CLIL learners in this study have been exposed to CLIL science lessons in which abstract concepts and ideas have been dealt with. The fact that these lessons are in a foreign language increases the degree of difficulty, which may explain the modest language gains, in terms of oral skills, in the CLIL group. Moreover, according to Bruton (2011), cognitively demanding content may discourage oral interactions, especially if the content is new. Several studies on the effects of CLIL in secondary

school (Mewald, 2007; Wannagat, 2007) highlight the limited amount of L2 output produced by teenagers, L2 foreign language learners with greater cognitive maturity than young language learners. Taking this into account, it is not surprising to find out that CLIL does not enhance, at least in a significant way and in the short term, the development of the learners' oral production skills. Additionally, it is also important to highlight that their L2 knowledge is much more limited than that of teenagers. Because of this, young language learners rely on their exemplar-based system (Skehan, 1998; Lyster and Sato, 2013) to communicate in the L2. Therefore, it is crucial that instructional settings provide learners with plenty of linguistic support in the form of language chunks or formulaic language to develop their oral production skills, especially in CLIL settings, in which the focus on language objectives seems to be generally overlooked.

The rigorous design of the study may also explain the lack of significant results obtained from the intergroup comparisons in favour of CLIL learners. As thoroughly presented in Chapter 3, relevant instances of CLIL research conducted up to the present day in the European context (Jarvinen, 2005; Admiraal et al., 2006; Lasagabaster, 2008; Lorenzo, 2010; Navés and Victori, 2010; Vallbona, 2009 and 2011; Juan-Garau, 2010; Várkuti, 2010; Bret, 2011; Lázaro and García-Mayo, 2012) reveal that CLIL learners significantly outperform non-CLIL learners in many language-related areas. However, CLIL groups in these studies were exposed to a higher number of hours of English instruction than non-CLIL groups, which might have affected the results obtained in favour of CLIL learners. In this study, both groups have been exposed to exactly the same hours of English instruction, which may explain why CLIL does not seem to have a significant impact on young learners' oral production skills, at least when CLIL and non-CLIL learners' scores in the CAF measures are compared at different data

collection times. Nevertheless, the fact that significant differences in favour of non-CLIL groups disappear at T3 suggests that CLIL instruction does have the potential to enhance young learners' speaking skills to a greater extent than EFL instruction, at least under the right conditions (i.e. intense CLIL instruction, the use of focus on form tasks and the practice of oral skills).

Altogether, the results conducted in this study partially confirm Hypothesis 1, which predicted that statistically significant differences in favour of CLIL learners would be found in fluency mainly, while no differences would be detected in the other areas of analysis. The results obtained from the analyses show no significant differences in favour of CLIL learners in propositional and syntactic complexity, accuracy or fluency. The only significant differences in favour of CLIL learners were in the percentage of nouns produced at T0, T1, T2 and T3 in the interview task and at T0 and T1 in the narrative task. As for Hypothesis 2, it predicted that CLIL learners would display significant development in all the areas analysed in the two tasks. According to the results, significant development was found in propositional and syntactic complexity and fluency in the interview task and in propositional complexity, the percentage of nouns produced and fluency in the narrative task. On the basis of these findings, Hypothesis 2 is only partly confirmed.

7.2 Achievement and development results according to proficiency level

As opposed to research conducted on the effects of CLIL on low proficient learners' receptive skills (Pladevall-Ballester, in press; Aguilar and Muñoz, 2014 and Vallbona 2014) and according to the results from this study, statistically significant differences between CLIL and non-CLIL low achievers are marginal in regard to their achievement scores in the CAF measures employed to investigate the development of

their oral production skills. The only significant differences in favour of CLIL low achievers are in the percentage of nouns in the interview task and in the percentage of error-free units in the narrative task at T3. As for control low achievers, the analyses yield significant differences in the percentage of coordinate units at T2 in the interview task only. Likewise, CLIL high achievers significantly outperformed their non-CLIL counterparts in the percentage of nouns produced at T0, T1, T2 and T3 and in the percentage of L1 words at T0, T1 and T2 in the interview task only.

As seen from these results, the tendency in this study is not to find significant differences between CLIL low and control low achievers in terms of their oral production skills, which contradicts the idea that CLIL low achievers might be in disadvantage, when compared to their counterparts in the EFL group, due to the cognitively challenging nature of the input received in CLIL contexts (Halbach, 2009). On the contrary, one of the findings of this study is that CLIL does not have a negative impact on low proficient learners, as no significant differences were found between CLIL and control low achievers in the total number of units, percentage of coordinate units, correct verb forms and speech rate in words at T1, T2 or T3. Likewise, Vallbona's (2014) study on the effects of CLIL on receptive and writing skills reported no significant differences between low achievers in the CLIL science and control groups in any of the CAF measures used to explore the learners' writing skills. The listening results, on the contrary, do show significant differences in favour of CLIL low achievers after two years of CLIL instruction.

The achievement results obtained by high achievers do not indicate that CLIL benefits high proficient learners' oral skills, either, since no differences were detected in favour of CLIL high achievers (except for the percentage of nouns produced and L1 use), which is in line with Vallbona's (2014) research study as well. Despite the lack of

significant differences in favour CLIL learners according to their proficiency level, it is important to highlight the fact that no significant differences were found between CLIL low and high achievers at any of the four data collections, while significant differences were observed in favour of control high achievers, when compared to control low achievers, in propositional complexity and fluency. These results suggest that differences in the learners' oral skills are much more noticeable in non-CLIL settings than in CLIL settings. Non-CLIL high achievers, for instance, showed greater improvement than non-CLIL low achievers in fluency, despite the fact of having started with similar fluency levels at T0. Differences between low and high achievers in the CLIL group, in contrast, are not visible at any of the four data collection times. This finding suggests that the development of oral skills is better fostered in EFL settings, thus, the idea that oral skills need to be further supported in CLIL settings by means of teaching practices that enhance the development of the learners' speaking skills seems to be backed up by the results obtained in this study again.

With regard to the intragroup analyses of the data obtained in the interview task, all four groups (CLIL high achievers, CLIL low achievers, control high achievers and control low achievers) tended to obtain significant improvement in the same measures and time periods. The only difference worth pointing out was found in speech rate in words, where CLIL learners, both high and low, displayed significant development during the last time period (T2-T3) and from T0 to T3, while control learners showed significant improvement during the first year (T0-T1) and from T0 to T3. The fact that CLIL learners showed greater improvement in fluency during the second year of CLIL instruction (T2-T3) backs up the claim that the effects of CLIL will take place once learners have accumulated enough CLIL input. Additionally, CLIL high achievers

displayed higher percentages of improvement than CLIL low achievers in propositional complexity, the percentage of adjectives and speech rate in words.

The results obtained from the narrative task, on the contrary, do seem to show an advantage of CLIL learners over control learners. First of all, there is a higher number of instances of significant development in favour of CLIL learners, both high and low achievers. Secondly, according to the results obtained, CLIL low achievers display greater improvement than CLIL high achievers in propositional and syntactic complexity and fluency from T0 to T3 or from T2 to T3. As for the effects of proficiency on the degree of development in the control group, the analyses conducted report that control high achievers displayed higher percentages of improvement than control low achievers in propositional and lexical complexity, accuracy and fluency.

These results seem to provide some evidence to support the claim that low achievers benefit more from CLIL instruction than high achievers. Nevertheless, it is important to note that the positive effects of CLIL on low achievers are seen in the output produced in the narrative task, mainly. The data obtained from the interview task, on the other hand, indicate that CLIL high achievers benefit more from CLIL instruction. All in all, the results obtained in this study seem to point to the direction that CLIL instruction is not detrimental to low achievers' language development. In addition, data has also been provided to show that both CLIL high and low achievers undergo development. However, they do so in different communicative tasks. CLIL high achievers perform better than CLIL low achievers in the interview task. Interactive tasks might create more difficulties to low achievers, as comprehension problems may arise as a result of their low proficiency level. Consequently, if low achievers display more difficulties in understanding the questions in an interview task, it is likely that their oral production skills, measured as CAF measures, will also be affected, especially

fluency, since the learners will present more difficulties in producing output when they do not understand what they are being asked. Low achievers, on the other hand, performed better in the narrative task, where they did not depend on external linguistic input. The reason why CLIL low achievers generally outperformed CLIL high achievers in the narrative task while the situation in the control group is exactly the opposite (i.e. high achievers obtained higher percentages of improvement than low achievers), might be accounted in terms of the input CLIL learners receive. As Pladevall-Ballester (2014) suggests, low achievers might be pushed to make greater efforts in CLIL settings and that explains why their improvement becomes more noticeable.

To conclude, Hypothesis 3, which stated that more significant differences would be found in favour of CLIL low achievers than CLIL high achievers when compared to their respective counterparts in the non-CLIL group, cannot be confirmed in relation to speaking skills and according to the results obtained. Nevertheless, evidence has been provided to partially confirm Hypothesis 4, which predicted that CLIL low achievers would display greater development than CLIL high achievers.

7.3 Relationship of CAF elements

This study also attempted to explore the evolution of CAF in CLIL+EFL and EFL settings in the interview and narrative tasks with the aim of gaining insight into the development of young learners' oral production skills in two different instructional settings and in two tasks. According to the mean scores obtained by CLIL and non-CLIL learners in the interview task, both groups displayed constant and almost simultaneous improvement of syntactic complexity, measured by the percentage of coordinate units, and fluency, measured as speech rate in words, from T0 to T3. This tendency was confirmed by the correlation tests conducted, which reported a strong

positive correlation between syntactic complexity and fluency from T0 to T3 for both CLIL and non-CLIL groups. The simultaneous increase of syntactic complexity and fluency was also observed in the qualitative individual analyses of the learners' output, which indicated that the number of English words increased at the same time as syntactically complex structures started to be used. The accuracy measure employed to study the development of CAF, percentage of error-free units, on the contrary, displayed a gradual decrease throughout the four data collection times for the two groups. However, despite the decreasing tendency found in CLIL and non-CLIL learners' mean scores in accuracy, a strong negative correlation between syntactic complexity and accuracy was only found for the non-CLIL group.

These findings confirm Skehan and Foster's (Extended) Trade-off Hypothesis (2012) which claims that greater fluency will be accompanied by greater accuracy or complexity (but not both) due to the fact that our attentional resources are limited and, thus, high levels of complexity, accuracy and fluency are unlikely to take place simultaneously. Even in the case of the CLIL group, where the correlation between complexity and accuracy is weak, the mean scores still showed the existence of tradeoffs between accuracy and complexity. According to the Trade-off Hypothesis, learners can rarely focus on both complexity and accuracy at the same time, which explains why an increase in the use of coordinate or subordinate units may lead to a decrease in accuracy. This has also been illustrated in the qualitative analyses carried out at the individual level which showed that the learners' attempts to produce syntactically complex structures (i.e. coordination or subordination) triggered an increase in the number of errors. Concerning the relationship between accuracy and fluency, the correlation tests reported a medium negative correlation for both CLIL and non-CLIL groups. The correlation tests showed that learners with high accuracy levels

tended to obtain low scores in fluency. Again, the idea that CAF develops simultaneously is unsupported by the results from the correlation tests.

The mean scores obtained by CLIL and non-CLIL learners in CAF in the narrative task showed more irregular tendencies. To start with, the learners' scores in this task could not be submitted to correlational analyses due to the low scores in error-free units and coordinate units throughout the four data collection times in both groups. Additionally, even though complexity and fluency improved from T 0 to T 3, their development was not as simultaneous as observed in the interview task. As for accuracy, the mean scores underwent little variation from time to time. Altogether, it can be concluded that complexity and fluency showed the greatest improvement in this task, while accuracy seemed to remain unaffected by time and instruction.

The differences found in CAF scores between the interview and narrative tasks can be explained in terms of the effects of task type on L2 performance. As regards complexity, the mean scores obtained in the narrative task were generally much higher than those obtained in the interview task. According to Skehan (1998) tasks requiring interpretations often lead to the use of more complex language. In the narrative task, young learners had to interpret the images of the narrative to produce a coherent story. Additionally, they were expected to verbalise complex meaning relationships such as cause-effect or temporal relationships between different events in the story. This prediction is also found in Robinson (2001) for monologic complex tasks and in Ferrari (2012) for non-interactive tasks. As for accuracy, CLIL and non-CLIL learners' mean scores were quite high in the interview, as according to Skehan (1998), structured tasks dealing with familiar information, such as the interview in this study, lead to high scores in accuracy. According to Robinson (2001), on the contrary, simple interactive tasks are likely to trigger low accuracy levels. In this study, the interview is considered to be

relatively easier than the narrative task, as most of the questions asked are extensively dealt with in regular EFL lessons. Finally, Skehan (2003) predicts that monologic tasks will produce greater fluency than interactive tasks, which seems to be supported by the data in this study as fluency scores were much higher in the narrative task than in the interview task. Robinson (2001), on the other hand, defends the idea that complex tasks (the narrative task is seen as a complex task in this study) produce less fluency, which cannot be confirmed by the data obtained from the narrative task.

In the light of these findings, enough data has been provided to confirm Hypothesis 5, which claimed that complexity, accuracy and fluency would not develop simultaneously in the interview and narrative tasks from T0 to T3. In the case of the interview, the results are robust enough to claim that tradeoff effects were observed between complexity and accuracy, particularly in the non-CLIL group. In regard to the narrative task, the quantitative analyses show that fluency and complexity improve from T0 to T3, while accuracy remains unaffected in both CLIL+EFL and EFL contexts, which seems to yield further evidence to suggest that CAF measures do not improve simultaneously.

7.4 YLLs' oral output at the end of primary school

This section of the chapter aims at characterising the oral output produced by the young language learners selected for the qualitative analyses in the interview and narrative tasks in their last two years of primary school (from 9/10 years old to 11/12 years old) to explore Hypothesis 6. In addition, this section will also discuss young learners' oral development in relation to L2 syntactic developmental stages and L2 processes such as simplification and restructuring. Firstly, an interpretation of the language features produced in the interview and narrative tasks associated with

syntactic and lexical complexity, accuracy and fluency will be provided for the first and second year of this investigation (T0-T1 and T2-T3, respectively), in an attempt to present realistic goals for the assessment of young language learners' oral output in instructed SLA in their final stages of primary school. Additionally, a comparison between CLIL and non-CLIL learners will be presented using CAF measures. Next, the learners' oral output produced during the narrative task will be considered using Álvarez's (2006) stages of narrative discourse. It should be acknowledged that even though the characterisation of young language learners' oral output presented in this section will be based on the description of 6 learners only, several generalisations will be made in regard to the type of oral output that can be expected from young language learners with different fluency levels. To do so, the results obtained from the statistical analyses will also be employed.

To begin with, syntactic complexity is one of the dimensions which undergoes the greatest changes in the last two years of primary education according to the qualitative analyses conducted in this study. Dysfluent learners' oral output at T0 and T1, for instance, was featured by the production of content words with little syntax or morphology. This lack of complexity (of any kind), also known as simplification, is an L2 process which entails the transmission of a message with little language. According to Ortega (2009), massive simplification takes place at early stages of language development as a result of scarce L2 knowledge. In the case of young language learners, though, the use of simplification strategies is also likely to stem from their limited cognitive and processing capacity, which prevents them from carrying out cognitively demanding processes to produce language at phrasal or clausal levels. The use of simplification strategies was also acknowledged in Pienemann's Processability Theory (1998) applied to English (Ellis: 2008 : 98), which proposed that early stages of

language development are featured by the use of invariant forms composed of single constituents or formulaic language. At T2 and T3, the output produced by dysfluent learners achieved a stage in which canonical English order was stabilised with the production of SVO structures first and the emergence of coordination later. Fluent learners' output was also characterised by the use of simplification strategies at T0. At T1, however, instances of canonical SVO order were found. Consequently, T1 also provided instances of coordinate units. As stated above, the mastery of SVO word order and the emergence of coordination seem to be closely related. At T2 and T3, the use of coordination increased. As for subordination, some attempts to produce subordinate clauses were sporadically found. Highly fluent learners, on the other hand, already produced coordination, or at least attempted to do so, in the first year of the study (T0-T1). Additionally, some attempts to produce subordination occurred. In the second year of the study (T2-T3), successful attempts to produce subordination emerged at the same time as coordination was used extensively. The types of subordination found were adverbial and nominal clauses (mainly to-infinitive clauses of purpose and that-clauses). The development found in the learners' syntactic system seems to confirm Pienemann's acquisition hierarchy (1998). According to the author (ibid), the use of embedding characterises stage 6 of morphosyntactic development, which means that learners might have previously learnt and produced noun phrase agreement, inversion and subject-verb agreement. According to the samples described in the previous chapter, highly fluent learners generally produced these features prior to the production of subordination.

Regarding subordination, it was also observed that, in some cases, unsuccessful attempts to produce subordinate clauses triggered a decrease in coordination. These tradeoff effects between subordination and coordination might explain the decrease in the percentage of coordinate units shown in the quantitative analysis for CLIL and non-

CLIL low achievers in the narrative task from T1 to T2. A possible explanation for the decrease in coordination may relate to the learners' willingness to convey messages which contain higher degrees of precision and complexity through the use of subordination, which in turn might have caused a decrease in coordination. These tradeoff effects, however, seem to be temporary, as learners increased the number of coordinate units again from T2 to T3. In addition, according to the quantitative analyses, the non-CLIL group displayed a slight decrease in coordination from T2 to T3, while the CLIL group continued increasing. This slight decrease in coordination in the non-CLIL group might be related to the use of subordination as well.

The development of questions and negation are also worth discussing in terms of developmental stages. According to Pienemann, Johnstone and Brindley (1988), the first stage of question formation consists of isolated words with rising intonation. No instances of questions at this stage were found in the samples analysed. However, examples of utterances containing canonical English word order with rising information (stage 2) were seen in the output produced by dysfluent and fluent learners, mainly. Instances of questions at stage 2 were also found in the output of the non-CLIL highly fluent learner. These kinds of questions were mainly produced by dysfluent learners along with accurate wh-questions formed by formulaic language during the second year of the study (T2 and T3). Fluent learners, on the other hand, produced questions which contained a question element in the first position (stage 3) at T2 and T3. Finally, the highly fluent CLIL learner seemed to produce more complex questions than the highly fluent control learner. At T3, the CLIL learner produced questions from stage 4, featured by the use of inversion in wh-questions and in yes/no questions, while the highly fluent control learner produced questions from stage 3 (fronting of a question

element). Altogether, the development found in the output analysed in the qualitative analyse seems to confirm Pienemann et al.'s (1988) question developmental stages.

Regarding the development of negation, the only two learners who used negation and showed development through the stages in the sequence of acquisition in L2 English summarised in Ellis (2008) were the two highly fluent learners. According to the sequence of negation acquisition, learners start using external negation (i.e. *no have the cake)*. Instances of external negation were found at T0 and T2 in the output of the highly fluent CLIL learner. The non-CLIL learner, on the other hand, already produced internal negation (i.e. *in the nest do not have the breakfast)* at T1. Despite these differences, both learners attached negation to the verb *be* at T3 which indicates that the CLIL learner underwent greater progress than the non-CLIL learner during the second half of the study. According to Ellis, negative attachment to modal verbs is the stage that follows internal negation. No instances of modal verbs were found in the output of young language learners. However, learners did succeed in attaching the negative particle to the verb *be.*

Even though CLIL and non-CLIL learners shared some similarities in terms of syntactic complexity, several differences were observed. The highly fluent non-CLIL learner, for example, displayed a more advanced degree of syntactic complexity from T0 to T1, while the highly fluent CLIL learner showed greater improvement in the second year of the study, which is line with the results obtained from the statistical analyses indicating that CLIL learners displayed greater development during the second year of the study. In contrast, fluent learners displayed more differences. For example, the fluent CLIL learner produced coordination at T1, while the fluent control learner already did so at T0. Despite this difference, both learners produced subordination at some point during the second year in one of the two tasks. Regarding dysfluent learners,

the non-CLIL learner already produced coordination at T1, while the CLIL learner did not attempt to do so until T3. This is also supported by the quantitative results, since significant differences in favour of non-CLIL low achievers were found at T2. At T3, the differences disappeared. Yet, non-CLIL low achievers still outstripped CLIL low achievers.

In the light of these findings, several generalisations can be drawn concerning the oral output that can be produced by young language learners at the end of primary school who have received similar amounts and types of English instruction as the participants in this study. Firstly, young language learners with low fluency levels can be expected to attain canonical English order through the production of complete SVO utterances by the end of primary school. Additionally, and as consequence of the stabilisation of this syntactic structure, some attempts to use coordination are likely to take place. Fluent learners can make extensive use of coordination as well as attempts to produce subordination. Finally, highly fluent learners can be expected to use a considerably high number of coordinate and subordinate units. In regard to the use of questions, it is worth noting that only those learners with high fluency levels can be expected to produce inversion. Learners with lower fluency levels may only be able to use formulaic questions and questions containing a question element (i.e. wh-word or do) in the first position. According to the samples in this study, the correct formulation of questions might only take place at higher levels of cognitive development due to young learners' limited processing capacity. Likewise, the formation of negation may also present some difficulties among young language learners. As previously seen, learners at the end of primary school do not show absolute command of this language feature. Thus, learners entering the secondary school system cannot be expected to produce accurate use of negation structures in all contexts, although some instances of

negative particles attached to the verb *be* might occur. No instances of negative attachment to the auxiliary verb were found. Despite that, they might be produced in certain contexts by learners with high fluency levels.

With regard to lexical complexity, the descriptions provided in Chapter 6 along with the quantitative analyses obtained in the interview task mainly, provide evidence to support Broeder et al. (1993) and Muñoz's (2006) claim that the use of verbs increases as the learners' interlanguage develops. Highly fluent learners displayed a considerable increase in the number of verbs from T1 onwards. The same situation was observed in the fluent learners' oral output. Dysfluent learners, on the other hand, showed an increase in the number of verbs some time later (T2 or T3). At T0, their speech is characterised by the use of isolated nouns and, on some occasions, the verb *be*. As previously explained, massive simplification is quite common at incipient levels of language development. As for the use of adjectives, the most common adjectives found in the learners' speech were *old* and *favourite*. The use of adjectives was scarce among the young learners selected. According to the quantitative analyses, CLIL and non-CLIL learners' percentages of adjectives in the interview task ranged from 3% to 5% over the two years of the study. Additionally, the number of adjectives produced in the narrative task was not high enough to conduct the statistical analyses. The use of adjectives was generally found in formulaic language at T0 and T1, while at T2 and T3 instances of adjectives in a wider variety of contexts were observed. This finding may provide partial evidence to support Myles' (2012) view that formulaic language triggers creative language use. The teaching of formulaic language has been widely acknowledged among researchers (Skehan, 1998; Wood, 2002; McKay, 2006; Muñoz et al., 2010; Myles, 2012) on the basis of their role to help beginner language learners communicate effectively and develop their interlanguage. The functions and effects of formulaic

sequences on language development will be further dealt with in more depth later on in the section on accuracy.

Concerning the differences between CLIL and non-CLIL learners' lexical complexity, it is worth pointing out that the highly fluent CLIL learner used prepositions to complement the meaning of some verbs along with phrasal verbs. No instances of prepositional verbs were found in the speech of the highly fluent control learner. The reason behind this difference could be related to the type of input received. In CLIL lessons, the type of input received is more varied in terms of lexical and syntactic complexity (Muñoz, 2007). As previously pointed out, CLIL learners were exposed to mini-lectures in their CLIL lessons. According to Dalton-Puffer (2007) the use of lectures in CLIL settings provides learners with great amounts of input in which facts, concepts as well as semantic relations between them are established and integrated into coherent discourse structures containing certain degrees of syntactic, lexical and textual complexity. The fact that CLIL learners were exposed to this type of input may account for the lexical differences encountered between CLIL and non-CLIL highly fluent learners in the qualitative analyses.

In the light of these results, young learners' oral output at the end of primary education can be expected to contain a wide use of verbs such as *play, live, work, go, be, have, like, walk, look, eat* and *say*, among other verbs related to their everyday actions. In regard to adjectives, their use is expected to be restricted to very concrete contexts. Spontaneous use adjectives in a wide variety of contexts is not likely to take place due to the fact that learners tend to learn the majority of adjectives in chunks of language that are used in specific contexts.

As for accuracy, it is important to note that the young learners selected for the qualitative analyses all showed high levels of accuracy at the beginning of this study

(T0) in the interview task mainly, independently of their fluency levels. The quantitative analysis also showed high levels of accuracy at T0, especially in error-free units. Two reasons may account for the low number of errors found at T0. First of all, young learners tended to answer the questions with very short utterances containing just a few words (i.e. massive simplification strategy), which diminished the risk of making mistakes. In addition to that, another factor that influenced the low number of mistakes at T0 was the use of formulaic language. The majority of the questions learners formulated, for instance, were unanalysed chunks of language, which allowed them to produce a high number of error-free units. As Palloti (2009) and Myles (2012) point out, high levels of complexity, accuracy and fluency in early language cannot be interpreted as indicators of language development. On the contrary, as Ellis (2008) claims "errors are viewed as indicators of learners' interlanguage development and also of the mental processes involved" (p. 548). Thus, as pointed out by previous researchers, errors are closely related to development, which explains why the number of error-free units decreased considerably from T1 onwards. As seen from the qualitative analyses, learners started to restructure and analyse already learnt chunks of formulaic language to create new utterances at T1. When this restructuring process took place, learners' accuracy levels started to decrease while their amount of language and their ability to produce creative utterances increased. The quantitative analyses of the percentage of error-free units and the total number of units also confirm that. These results provide further evidence to back up Myles' (2012) claim that "formulaic sequences play a crucial role in feeding increasingly complex structures into the construction of productive grammar" (p. 91). In addition, instances of how learners restructure already learnt chunks of language to be able to communicate in certain situations emphasises, as previously stated, the need to foster the teaching of formulaic language in both CLIL

and EFL settings. In so doing, learners will be provided with sufficient linguistic resources to communicate effectively at low stages of language development. Furthermore, the storage of formulaic sequences will enable them to develop their linguistic system as well as to produce their own utterances.

According to the qualitative analyses, the development of syntactic complexity also affected the learners' accuracy levels. That is to say, as learners attempted to produce longer utterances containing syntactically complex structures like coordination and subordination (or even SVO structures), the learners' accuracy levels decreased, which provides further evidence to support Skehan's Trade-off Hypothesis (1998). This negative relationship between syntactic complexity and accuracy backs up Long's (2009) view that the occurrence of errors may precede, or even trigger, the acquisition of a new linguistic feature, in this case syntactic structures. The errors found in questions and utterances containing negation are a clear example of the relationship between the emergence of errors and syntactic development.

In regard to the typology of errors found, research has provided a great number of classifications. The classification used here will be the one proposed by Pica (1983), summarised in Long (2009). This choice was made on the basis of the participants used in Pica (1983), who were native speakers of Spanish learning English in naturalistic, instructed and mixed contexts. According to Pica's findings, all learners made the same type of errors: overgenerlisation errors, overuse errors, omission errors and substitution errors. In this study, the most common errors were those which involved overuse of L2 linguistic features in non-obligatory contexts (i.e. the use of the –ing morpheme), omission errors (i.e. lack of auxiliary verbs in questions, lack of verb inflections or lack of verbs) and substitution errors (i.e. using the past tense when the present tense should be used). Highly fluent learners, for instance, produced instances of omission errors

throughout the four data collection times in the two tasks (i.e. lack of auxiliary verbs in questions and the lack of verb inflections). Substitution errors were also found in instances in which learners used inappropriate verb tenses. The same kinds of errors were found in the speech produced by the two fluent learners and dysfluent learners. In addition, instances of predicate omissions and overuse of –ing forms were also observed at some point during the first year of the study in the output produced by fluent and dysfluent learners.

As for the comparison of CLIL and non-CLIL learners, the samples presented in the qualitative analyses display almost no differences in the number of mistakes produced by learners with similar fluency levels according to the type of instruction received. No differences in error-free units were found between CLIL and non-CLIL learners in the quantitative analyses either, which suggests that the development of accuracy and the typology of errors might remain unaffected by instruction type, as Pica (1983) and Long (2009) defend. Despite that, it is worth noting that CLIL learners generally presented the highest percentages of improvement in accuracy, according to the quantitative analysis.

Taking all this into consideration, it is expected that the output produced by young learners at the end of primary will contain mistakes concerning the restructuring of formulaic sequences along with errors related to the overuse, omission and substitution of certain linguistic items. The most common mistakes may relate to use of certain linguistic items in non-obligatory contexts, the lack of verb inflections, the lack of auxiliary verbs in questions, the choice of verb tense and the choice of words (i.e. lexical errors). Errors related to the lack of predicates may also be observed in the output produced by those learners at lower stages of language development.

Finally, fluency is, along with syntactic complexity, one of the areas which undergoes the greatest improvement and reveals the highest number of differences among learners according to their fluency level. According to the statistical analyses in this study, both CLIL and non-CLIL learners showed significant development from T0 to T3 in speech rate in words. Additionally, differences between high and low non-CLIL achievers were also seen while differences according to proficiency level were not so noticeable among CLIL learners.

According to the qualitative analyses, highly fluent learners are the learners who displayed the greatest development in fluency from T0 to T3. The highly fluent CLIL learner improved the most during the second year of the study, producing the longest sentences. The highly fluent control learner, on the other hand, showed the biggest change in fluency from T0 to T1. Development in the fluent CLIL learner's speech also took place during the second year of the study, since the highest amount of language produced was observed at the end of the study (T3). A similar situation was observed among dysfluent learners. The CLIL learner seemed to maintain the amount of language produced constant at T0 and T1, while the amount of language increased considerably during the second year of the study. The control learner, in contrast, shows the greatest improvement at T1. As previously seen in the quantitative analyses, control learners tend to show the greatest improvement during the first year of the study, whereas CLIL learners need time to adapt to the new methodology. Once they are adapted and have accumulated enough CLIL input, the benefits of CLIL instruction become visible.

In addition to the amount of language, other fluency-related aspects were observed among young learners. Firstly, in the interview task, the majority of students showed, at some point in the study, problems in understanding some of the questions, which prevented them from providing an immediate answer. Additionally, all the

learners, except the fluent control learner, produced L1 words in one of the two tasks. L1 use may be related to the degree of precision learners want to show when performing a task. Thus, when learners did not know the English word for what they wanted to express, they used Catalan. Besides, L1 utterances were also used to generate metatalk or to check comprehension. The use of the L1 was quite scarce, according to the descriptive analyses, which seems to be in line with Lázaro and García-Mayo's (2012) study of L1 use in CLIL classrooms in secondary school and Serra's (2007) results on L1 use in CLIL settings in primary education. In addition, L1 use showed a decrease from T0 to T3, which is corroborated by the statistical analyses conducted and confirmed by Lázaro and García-Mayo's (2012) findings as well.

As previously stated, the main difference between CLIL and non-CLIL learners in terms of fluency is the time when the greatest development took place. In the case of CLIL learners the most important changes were observed during the second half of the study (T2-T3) while control learners developed to a greater extent at the end of the first year of the study (T1). This provides evidence to support the claim that the effects of CLIL are not immediate, particularly in minimal CLIL exposure programmes. The fact that CLIL learners were exposed to one hour of CLIL instruction a week only might account for the low degree of improvement displayed in the first year of the study, since according to her results, learners need to accumulate greater amounts of CLIL instruction in order to start benefiting from it as well as time to adapt to methodological changes.

On the basis of these interpretations, young language learners' output is expected to be fluent at the end of primary education, meaning that learners will be able to produce complete sentences rather than with simple and short incomplete utterances. In addition, L1 use and some comprehension problems may also take place (despite

their decreasing tendency), regardless of the learners' familiarity with the vocabulary being dealt with.

As for narrative discourse (Álvarez, 2006), several aspects regarding narrative and language development will be discussed in the following paragraphs. First of all, a word needs to be added in regard to the narrative discourse displayed by the learners selected for the qualitative analyses. As Álvarez (2006: 151) found out, the output produced by the learners in her study "cannot be said to constitute true narratives until stage 7 with the incipient emergence of discourse features" (i.e. subordination). Similarly, the learners selected for the qualitative analyses did not unfold narrative discourse features until stage 7. Instead, they described the different pictures in the story without establishing clear links between them. According to the author and the qualitative descriptions in this study, the use of subordination at stage 7 was only used to mark the most important event in the story (i.e. when the children realise that their dog has eaten the food), which is why the author suggests that the production of subordination is solely linked to a specific discourse function. As seen in the qualitative analyses, three out of the six learners produced subordination (the two highly fluent learners and the dysfluent control learner) to mark the climax of the story during the second year of the study mainly (T2 and T3).

Due to lack of narrative discourse features in the learners' productions, the author suggests that narrative tasks might not be adequate elicitation instruments for incipient stages of L2 language development and proposes the use of other elicitation instruments which lessen the cognitive and linguistic demands on the part of the learners, particularly for young language learners. Despite that, it must be acknowledged that the use of this narrative task did fulfil the purpose of examining the

emergence of subordination among incipient levels of language learning, which was one of the aims of this study.

Regarding the results obtained by Álvarez (2006), it is relevant to note that the only group of school learners which obtained stage 7 was the group of late starters after 416 hours English instruction at the second data collection time at the age of 14. The group of early starters (aged 12 years old), on the other hand, did not achieve stage 7 after the same amount of English instruction. These findings contrast the ones obtained in this research study, since instances of young learners producing features from stage 7 were found around T2 and T3 (after 164.5 and 244 hours of English instruction since the onset of the investigation). Two reasons might account for the difference in the findings. Firstly, the amount of previous English exposure students in this study received prior to the start of this investigation (around 420 hours) could explain why learners in this study produced features from stage 7. In addition, the procedure used to assign stages may also explain the differences in the results obtained. In Álvarez's study, learners were assigned to stages using a stage average, which was validated using the analyses of the output produced by the learners in an interview task and a cloze test. In this study, stages were assigned on the basis of the features produced while telling the narrative. No other task was used to carry out the classification of the learners output into stages. Taking this into account, the data treatment employed in Álvarez's study would explain why young language learners were not assigned to stage 7.

Regarding the development of linguistic features, the results obtained from this study back up the developmental patterns proposed by Álvarez. The author claims that the developmental patterns found in her study coincide with the route of development summarised in Ellis (1994): silent period, syntactic development and morphological development. According to Álvarez, learners start developing their syntactic system at

the phrasal level with the production of nominal and prepositional phrases. Once these features stabilise, syntactic development at the clausal level is further developed through the use of complete SVO and SVA structures. At this point, the learners' morphological system starts developing through incipient verb morphology (i.e. present participle and third person singular –s). The results obtained in this study confirm such development, as the classification of the learners' output in the qualitative analyses shows that syntactic development at the phrasal and clausal level takes place earlier than morphological development. To be more precise, syntactic development is usually seen at T0 and T1 while morphological development has been observed to take place at T1 for highly fluent learners and at T2 for fluent learners. As for dysfluent learners, little morphological development was found, although the non-CLIL learner produced linguistic features from stage 7 (i.e. subordination) at T3. These findings back up the view that morphological development takes place later than syntactic development.

As for the progress learners unfolded from T0 to T3, highly fluent learners, for instance, progressed through three stages and both of them attained stage 7, which is characterised by the use of subordination to mark the highest point in the story. The highly fluent CLIL learner produced linguistic features from stage 8 as well (i.e. infinitive constructions), while the control learner progressed from stage 4 to 7. The differences between CLIL and non-CLIL highly fluent learners may relate to the fact that the CLIL learner started from a higher stage than the non-CLIL learner at T0. No instances of subordinate units were found in the output produced by fluent learners. However, the CLIL learner developed through five stages, while the control learner produced features from stages 4 and 6 at T0 and T3. In this case, the fluent CLIL learner displayed faster development than the control learner, suggesting that CLIL instruction may benefit the rate of development of learners with average fluency levels. Finally,

dysfluent learners showed quite different degrees of development. The CLIL learner progressed through three stages while the control learner moved through four stages, attaining stage 7 at T3 through the production of an adverbial clause to express the most important event in the story. The reasons behind the differences between dysfluent learners might be explained in terms of individual differences, as the quantitative analyses do not generally show such big differences among low achievers according to the type of instruction received.

According to the description of the oral output produced by learners with different fluency levels in the interview and narrative task, Hypothesis 6 is partially confirmed. This hypothesis claimed that all the learners would improve in syntactic complexity, especially CLIL learners, and that the amount of formulaic language would vary according to the learners' fluency levels. To start with, all the learners showed substantial improvement in complexity through the consolidation of canonical word order and coordination mainly, being the two highly fluent learners (who made extensive use of subordination at T3) the group of learners who developed the most. However, little evidence was gathered to claim that CLIL learners displayed greater syntactic development than non-CLIL learners. As for the use of formulaic language, all the learners showed instances of formulaic sequences throughout the four data collection times. The difference found among learners with different fluency levels was that those learners with the lowest degrees of fluency tended to analyse and restructure formulaic sequences later in time than fluent and highly fluent learners. With regard to accuracy, the use of formulaic language as well as simplified language was generalised, which is why learners obtained high levels in accuracy at T0. As learners started to analyse and restructure formulaic sequences, their accuracy levels decreased. In addition, accuracy was also affected by syntactic development. Thus, learners' attempts

to produce syntactically complex structures such as subordinate clauses or questions triggered an increase in the number of errors. At T3, learners' accuracy levels had decreased considerably when compared to T0. Finally, great improvement was observed in fluency as well. At T3, learners produced longer utterances and greater amounts of language. Highly fluent learners showed considerable improvement, while fluent and dysfluent learners' improvement was more modest. The main difference between CLIL and non-CLIL learners was that CLIL learners' tended to show greater improvement from T2 to T3.

Hypothesis 6.1 predicted that highly fluent and fluent learners would attain stage 7, featured by the use of subordination and the emergence of narrative discourse. In addition, it was also expected that CLIL learners would develop faster through stages than non-CLIL learners. The data collected provides partial evidence to confirm this hypothesis, since the only learners who attained stage 7 were the two highly fluent learners and the dysfluent control learner. The dysfluent control learner produced subordination already at T1. The fact that this learner showed greater progress than his counterpart in the CLIL group might be accounted for in terms of individual differences. As for rate of development, only the fluent CLIL learner showed faster development than their counterpart in the non-CLIL group. Additionally, the oral data analysed in the qualitative analyses shows that learners do progress through the stages proposed by Álvarez (2006). The first development takes place at the syntactic level with the production of noun and prepositional phrases. Next, greater syntactic development takes place at the clausal level with the emergence of SVO and SVA structures. Finally, young learners in this study started to develop their morphological verb system by using the present participle and the third person singular –s.

To conclude, this chapter has provided a discussion of the results and the hypotheses posed in this study. The next chapter will present the answers to the research questions along with a summary of the most relevant concluding remarks and pedagogical implications derived from the findings in this dissertation. Additionally, some of the limitations found in this study as well as future research directions will be provided.

CHAPTER 8 Conclusion

This chapter presents the main conclusions derived from the results obtained in the data analyses and acknowledges the limitations of the study. In addition, several implications related to the teaching and development of oral production skills in CLIL settings, as well as the effects of minimal CLIL exposure will also be highlighted.

This two-year longitudinal study attempted to examine the effects of minimal CLIL exposure on the learning of young learners' oral production skills in an interview and narrative task by means of CAF measures at three data collection times. In so doing, the development of CAF in CLIL and non-CLIL learners' oral output was traced and analysed with the objective of identifying the areas that are mostly affected by CLIL instruction and comparing CLIL and non-CLIL learners' language-related gains. Additionally, this study also explored the effects of the learners' initial proficiency level (high versus low achievers) on the oral development in CLIL+EFL and EFL settings with the aim of determining to what extent the effects of CLIL vary according to proficiency level. Even though many studies have analysed the effects of CLIL on learners' oral skills, this piece of research is one of the first attempts in which exposure to English was kept constant between CLIL and non-CLIL groups. Besides, quantitative along with qualitative data analyses have been conducted in order to thoroughly investigate the effects of CLIL.

Hence, in order to explore the linguistic gains in young learners' oral production skills in CLIL+EFL and EFL settings, the following research questions were posited:

RQ 1 Are there statistically significant differences in propositional, syntactic and lexical complexity, accuracy and fluency achievement scores between young learners in the

CLIL and non-CLIL groups keeping amount of English input (hours of instruction) constant between the two groups at T1, T2 and T3 in the interview and narrative tasks?

According to the results, several significant differences in favour of non-CLIL learners were obtained in propositional complexity at T1 in the interview and narrative tasks and in the percentage of coordinate units at T1 and T2 in the interview task. No differences in propositional or syntactic complexity were observed at T3 between CLIL and non-CLIL groups, though. In regard to lexical complexity, CLIL learners significantly outperformed non-CLIL learners in the percentage of nouns produced at T1, T2 and T3 in the interview task and at T1 in the narrative task. The accuracy analyses yielded significant differences in favour of the non-CLIL group in the percentage of correct verb forms at T2 and T3 in the narrative task. Finally, the non-CLIL group significantly outperformed the CLIL group in speech rate in words at T1 and T2 in the interview task and at T2 in the narrative task. No differences between the groups were detected at T3. In brief, practically no significant differences in favour of CLIL learners were found throughout the study. In addition, most of the differences in favour of non-CLIL learners disappeared by the time the study finished (T3).

RQ 2 Do CLIL and non-CLIL young learners' achievement scores in complexity, accuracy and fluency develop significantly from T0 to T1, from T1 to T2, from T2 to T3 and from T0 to T3 keeping amount of English input (hours of instruction) constant between the two groups in the interview and narrative tasks?

In the interview task, both CLIL and non-CLIL learners developed significantly in propositional complexity from T0 to T1 and from T0 to T3 and from T0 to T3 in the

narrative task. Young learners in CLIL and non-CLIL settings also showed significant development in the percentage of coordinate units from T0 to T3 in the interview and narrative tasks. In addition, the CLIL group displayed a significant decrease in the percentage of nouns produced from T0 to T1 and from T0 to T3 in the narrative task. This is the only significant development found in lexical complexity. In regard to accuracy no periods of significant development were observed. Finally, the CLIL group displayed significant development from T0 to T1, from T2 to T3 and from T0 to T3 in speech rate in words in the interview task. The non-CLIL group also developed significantly from T0 to T1, from T1 to T2 and from T0 to T3 in the interview task. In the narrative task, both groups showed significant improvement from T0 to T3, however, the CLIL group also did so from T2 to T3 while the non-CLIL group from T0 to T1. In regard to the percentage of L1 words, the CLIL group displayed a significant decrease for all the time periods in the interview task, whereas the non-CLIL group decreased significantly in all the time periods except from T2 to T3. In the narrative task, both groups displayed significant decreases from T0 to T3. Additionally, the CLIL group also decreased significantly from T2 to T3. To sum up briefly, the two groups displayed significant improvement in practically the same measures and time periods. However, differences were observed in the development of fluency, mainly. From T2 to T3, the CLIL group always showed significant improvement from T2 to T3 in the two fluency measures in the two tasks, whereas the non-CLIL group did not.

RQ 3 Are there statistically significant differences in complexity, accuracy and fluency achievement scores between young learners in the CLIL and non-CLIL groups according to the learners' initial proficiency level keeping amount of English input

(hours of instruction) constant between the two groups at T1, T2 and T3 in the interview and narrative tasks?

RQ 3.1 Are there statistically significant differences in propositional, syntactic and lexical complexity, accuracy and fluency achievement scores between high achievers in the CLIL and non-CLIL groups at T1, T2 and T3 in the interview and narrative tasks?

This study reports statistically significant differences in favour of non-CLIL high achievers in propositional complexity at T1 and T2 in the interview task and at T1 in the narrative task. Additionally, non-CLIL high achievers also performed significantly better in the percentage of coordinate units at T3 in the narrative task. CLIL high achievers, on the other hand, produced a significantly higher number of nouns at T1, T2 and T3 in the interview task. In terms of accuracy, the analyses showed significant differences in favour of non-CLIL high achievers in the percentage of correct error-free units at T3 and in the percentage of correct verb forms at T2 and T3 in the narrative task. Finally, non-CLIL high achievers obtained significantly higher scores in speech rate in words at T1 and T2 in the interview task and at T1, T2 and T3 in the narrative task. Regarding the percentage of L1 words, CLIL high achievers produced significantly higher percentages of L1 words at T1 and T2 in the interview task. Altogether, the comparisons between high achievers in CLIL and non-CLIL groups showed that non-CLIL high achievers obtained a higher number of significant differences in CAF.

RQ 3.2 Are there statistically significant differences in propositional, syntactic and lexical complexity, accuracy and fluency achievement scores between low achievers in the CLIL and non-CLIL groups at T1, T2 and T3 in the interview and narrative tasks?

According to the results, non-CLIL low achievers produced a significantly higher percentage of coordinate units than CLIL low achievers at T2 in the interview task, while CLIL low achievers produced significantly higher percentages of nouns at T3 in the interview task. As for the results in accuracy, significant differences were found in the percentage of error-free units in favour of CLIL low achievers at T3 in the narrative task. Finally, no differences were found between CLIL and non-CLIL low achievers in fluency. Overall, the intergroup comparisons between CLIL and non-CLIL low achievers showed that non-CLIL low achievers obtained statistically significant differences in coordination, whereas CLIL low achievers did so in the total number of nouns and in the percentage of error-free units.

RQ 3.3 Do CLIL low achievers obtain more statistically significant differences in propositional, syntactic and lexical complexity, accuracy and fluency achievement scores than CLIL high achievers when compared to their respective peers in the non-CLIL group at T1, T2 and T3 in the interview and narrative tasks?

As seen from the answers to questions 3.1 and 3.2, very few significant differences were found in favour of CLIL learners according to their proficiency level when compared to their non-CLIL counterparts. On the one hand, non-CLIL high

achievers attained more significant differences than non-CLIL low achievers. As for CLIL learners, no significant differences were found in favour of CLIL high achievers (except in the percentage of nouns and L1 words produced), while CLIL low achievers performed better than non-CLIL low achievers in accuracy and the total number of nouns at T3. According to these results, little evidence was gathered to claim that low achievers benefit more from CLIL instruction than high achievers.

RQ 4 Do CLIL and non-CLIL young learners' achievement scores in complexity, accuracy and fluency develop significantly from T0 to T1, from T1 to T2, from T2 to T3 and from T0 to T3 according to the learners' initial proficiency level keeping amount of English input (hours of instruction) constant between the two groups in the interview and narrative tasks?

> **RQ 4.1** Do CLIL and non-CLIL high achievers' scores in propositional, syntactic and lexical complexity, accuracy and fluency develops significantly from T0 to T1, from T1 to T2, from T2 to T3 and from T0 to T3 in the interview and narrative tasks?

According to the results, CLIL and non-CLIL high achievers developed significantly in propositional complexity from T0 to T3 in the interview and narrative tasks. In addition, CLIL high achievers showed significant development from T0 to T1 in the interview task. Concerning syntactic complexity, CLIL high achievers developed significantly from T0 to T3. In regard to lexical complexity, CLIL high achievers significantly decreased the number of nouns produced from T0 to T3 in the interview and narrative tasks. In terms of accuracy, CLIL high achievers displayed significant

decrease in the percentage of error-free units from T0 to T3 in the interview task. Finally, the fluency analyses report that CLIL high achievers improved significantly in speech rate in words in the interview task from T2 to T3 and from T0 to T3, while non-CLIL high achievers showed significant development from T0 to T1, from T1 to T2 and from T0 to T3. In the narrative task, CLIL high achievers displayed significant development from T0 to T3, whereas non-CLIL high achievers developed significantly from T0 to T1, from T2 to T3 and from T0 to T3. As for the percentages of L1 words, both CLIL and non-CLIL high achievers decreased significantly from T0 to T3 in the interview task and narrative tasks. Additionally, CLIL high achievers also decreased significantly from T2 to T3 in the narrative task. To summarise, CLIL high achievers displayed more periods of significant development than non-CLIL high achievers in complexity and accuracy mainly.

RQ 4.2 Do CLIL and non-CLIL low achievers' scores in propositional, syntactic and lexical complexity, accuracy and fluency develop significantly from T0 to T1, from T1 to T2, from T2 to T3 and from T0 to T3 in the interview and narrative tasks?

The intragroup analyses show that CLIL and non-CLIL low achievers developed significantly from T0 to T3 in propositional complexity in the interview task. In addition, non-CLIL low achievers also showed significant improvement from T0 to T1. In the narrative task, only CLIL low achievers improved significantly from T0 to T3 in propositional complexity. In terms of syntactic complexity, CLIL low achievers also displayed significant development from T0 to T3 in the interview task. Concerning lexical complexity, CLIL low achievers showed a significant decrease of nouns from T0

to T3 in the narrative task. In regard to accuracy, no significant development was found for either CLIL or non-CLIL low achievers. Finally, CLIL and non-CLIL low achievers developed significantly in speech rate in words from T0 to T3 in the interview task and narrative tasks. In addition, CLIL low achievers also showed significant improvement from T2 to T3 in the two tasks. Finally, CLIL and non-CLIL low achievers displayed significant decrease in the percentage of L1 words from T0 to T3 in the two tasks. To sum up, CLIL low achievers obtained more periods of significant development than non-CLIL low achievers in complexity and fluency.

RQ 4.3 Do CLIL low achievers' scores in propositional, syntactic and lexical complexity, accuracy and fluency show greater development from T0 to T3 than those obtained by CLIL high achievers in the interview and narrative tasks?

According to the results, CLIL low achievers displayed greater improvement than CLIL high achievers in coordination, the percentage of adjectives and the percentage of correct verb forms in the interview task. In addition, CLIL low achievers also showed the lowest percentages of decrease in error-free units. In the narrative task, CLIL low achievers attained higher percentages of improvement in propositional complexity, coordination, the percentage of verbs, error-free units and the percentage of correct verb forms. Additionally, CLIL low achievers displayed a higher percentage of decrease in the number of nouns than CLIL high achievers. As for fluency, CLIL low achievers displayed the highest percentage of improvement from T2 to T3 only. In brief, CLIL low achievers seem to show greater improvement than CLIL high achievers in syntactic complexity, lexical complexity and accuracy.

RQ 5 What is the relationship between syntactic complexity, accuracy and fluency achievement scores obtained by young learners in the CLIL and non-CLIL groups from T0 to T3 keeping amount of English input (hours of instruction) constant between the two groups in the interview and narrative tasks?

RQ 5.1 Is there a correlation between syntactic complexity and accuracy achievement scores in CLIL+EFL and EFL settings from T0 to T3 in the interview and narrative tasks?

According to the correlation tests, a strong negative correlation was found between the percentages of coordinate units and the percentages of error-free units in EFL settings from T0 to T3 in the interview task. In CLIL+EFL settings, the results reported a weak negative correlation. The correlation tests could not be conducted for the narrative task due to the low scores in accuracy.

RQ 5.2 Is there a correlation between syntactic complexity and fluency achievement scores in CLIL+EFL and EFL settings from T0 to T3 in the interview and narrative tasks?

The analyses report a strong positive correlation between the percentage of coordinate units and speech rate in words in both CLIL+EFL settings and EFL settings in the interview task. The correlation tests could not be conducted for the narrative task due to the low scores in complexity

RQ 5.3 Is there a correlation between accuracy and fluency achievement scores in CLIL+EFL and EFL settings from T0 to T3 in the interview and narrative tasks?

The results indicate a medium negative correlation between the percentage of error-free units and speech rate in words from T0 to T3 in the interview task in CLIL+EFL and EFL settings. The correlation tests could not be conducted for the narrative task due to the low scores in accuracy.

RQ 6 What are the characteristics, in terms of complexity, accuracy, fluency and oral narrative competence, of the output produced by a selection of young learners with different fluency levels at T0 and T3 in CLIL+EFL and EFL settings in the interview and narrative tasks?

RQ 6.1 What are the characteristics, in terms of syntactic and lexical complexity, accuracy and fluency, of a highly fluent, a fluent and a dysfluent young learner's oral output at T0 and T3 in CLIL+EFL and EFL settings in the interview and narrative tasks?

According to the individual descriptions, some use of coordination was observed in the oral output of highly fluent learners at T0. At T3, in contrast, highly fluent learners extensively use coordination and subordination (i.e. to-infinitive clauses and that-clauses). Fluent learners, on the other hand, achieved good command of coordination from T0 to T3. However, instances of subordination were scarce at T3. As for dysfluent learners, the greatest change observed relates to the use of SVO structures.

At T0, dysfluent learners' output did not contain predicates on many occasions due to simplification strategies, while at T3 learners showed mastery of SVO structures as well as attempts to produce coordination. In addition, none of the learners showed full development in the production of questions or negation. The main difference between CLIL and non-CLIL learners is that CLIL learners started to develop their syntactic system later than non-CLIL learners, as a result, non-CLIL learners showed higher degrees of syntactic complexity in some cases.

With regard to lexical complexity, one of the characteristics of the learners' output at T0 was the limited use of verbs. Highly fluent and fluent learners increased the number of verbs from T1 onwards, while dysfluent learners did so from T2 to T3. As for the use of adjectives, the three groups of learners only used adjectives that were inserted in formulaic sequences. The most relevant difference observed between CLIL and non-CLIL learners was found in the group of highly fluent learners, in which the CLIL learner used prepositional verbs.

Concerning accuracy, the main finding relates to the use of formulaic language and its effects. At T0, learners (regardless of their fluency level) made extensive use of formulaic language, which accounted for the high scores in accuracy learners obtained, especially in the interview task. At T3, however, learners started to analyse and restructure formulaic sequences in order to achieve a specific communicative purpose. In so doing, accuracy levels decreased. This was observed for all the learners, independently of the type of instruction received. Additionally, errors related to the development of syntactic structures were also observed. The most common errors found at T3 were related to the omission, overuse and substitution of certain linguistic features.

Finally, the most noticeable changes in fluency are associated with the amount of language produced. At T0, learners produced fragmented language (mainly dysfluent learners) and very short utterances. At T3, all the learners considerably increased the amount of language produced, thus the sentences became longer. Despite that, comprehension problems as well as L1 use was still present at T2 and T3. As previously mentioned, the main difference between CLIL and non-CLIL learners was that CLIL learners showed the greatest improvement from T2 to T3, while non-CLIL learners already displayed important improvement at T1.

RQ 6.2 What stages of oral narrative development are identified in the oral output of a highly fluent, a fluent and a dysfluent young learner at T0 and T3 in CLIL+EFL and EFL settings in the narrative task?

Both highly fluent CLIL and non-CLIL learners attained stage 7 at T3, featured by the emergence of subordination and narrative discourse features. The CLIL learner moved from stage 5, characterised by the use of complete SVA structures, the present progressive and incipient discourse organisation, to stage 8 with the use of infinitive constructions. The highly fluent control learner, on the other hand, progressed from stage 4, featured by the use of SVO and SVA structures and the conjunction *and*, to stage 7.

The fluent CLIL learner started at stage 2 at T1 with the first prepositions and prepositional phrases and attained stage 5 at T3 with complete SVA structures. At T2, however, the fluent CLIL learner produced an attempt to use the present perfect, which is one of the characteristics of stage 8. Despite this, no use of subordination (stage 7)

was observed. The fluent non-CLIL learner progressed from stage 4 to stage 6, which is characterised by the use of the indefinite article.

Finally, the CLIL dysfluent learner developed from stage 2 with the use of nominal content words to stage 5, while the non-CLIL dysfluent learner progressed from stage 3, featured by the use of the L2 definite article, to stage 7.

Altogether, the results indicate that learners follow the developmental patterns presented in Álvarez (2006): syntactic development at the phrasal level, further syntactic development at the clausal level and morphological development. In addition, the data analysed in the qualitative analysis indicates that, in general terms, those learners with high fluency levels are able to produce subordination at the end of primary school to express complex syntactic relationships between the elements in a story. As for the comparison of the learners according to the type of instruction received, inconclusive results were gathered so as to claim that CLIL+EFL contexts are more beneficial for the learners' narrative development than EFL contexts.

Several conclusions can be drawn from the results obtained in this study. First of all, evidence has been provided to claim that language-related gains in young learners' oral production skills in minimal CLIL exposure programmes are not immediate. This claim is based on the fact that the amount of CLIL instruction during the first year of the study was not enough to enhance young learners' oral production skills when compared to their non-CLIL counterparts. In the second year, in contrast, the results reveal that the amount of CLIL instruction accumulated by the learners started to have a positive effect on CLIL learners' oral output, particularly on fluency, up to the point that significant differences in favour of the non-CLIL group disappeared at T3. Taking this into consideration, it is suggested that the implementation of more intense CLIL

programmes, which provide learners with greater amounts of more regular CLIL input, need to be encouraged in order for CLIL to have a significant impact on young learners' oral production skills. Up to the present, CLIL research has highlighted the superiority of CLIL programmes over more traditional forms of foreign language instruction on the basis of research findings obtained in studies which did not control for amount of English instruction between CLIL and non-CLIL or the comparability of the two groups. This study has acknowledged the potential and benefits of CLIL on several occasions. However, it has also illustrated that when amount of English instruction is kept constant between the two groups, the benefits of minimal CLIL exposure programmes do not reach significance.

As previously mentioned, minimal CLIL exposure programmes are widely implemented in Catalonia. The reasons behind the implementation of this type of CLIL programme relate to the lack of appropriate CLIL training and the lack of institutional support. Even though the Catalan educational authorities encourage the implementation of CLIL, little is being done to provide schools and teachers with the necessary material (i.e. teaching materials) and non-material resources (time for material design or meetings between content and foreign language teachers) to implement good quality CLIL programmes. That is why most Catalan schools decide to implement this type of CLIL programmes before launching stronger forms of CLIL.

Secondly, this study has revealed that in order for CLIL to be effective, a systematic development of learners' oral productive skills needs to be guaranteed. As in other forms of content-based instruction (i.e. immersion programmes and content-based instruction), there seems to be a stronger emphasis on receptive skills over productive skills. As a result, teaching practices for the learning and development of oral (and written) production skills tend to be overlooked. The use of focus on form tasks within

CLIL instruction seems to be a feasible solution for the integration of language objectives and the development of language skills. In so doing, the expectations of CLIL in regard to language outcomes would be fulfilled.

Another conclusion reached from the findings is that CLIL instruction does not necessarily have a negative impact on weak foreign language learners. According to the results obtained, no differences were observed in favour of non-CLIL low achievers at T3 in any of the measures used, which suggests that CLIL and non-CLIL low achievers perform similarly in CAF measures. This is quite a relevant finding within CLIL research and practice, as concerns and reluctant attitudes towards the implementation of CLIL are usually associated with the extra difficulties or problems weak learners will have to cope with in CLIL settings.

In addition, it can also be concluded that CAF measures in CLIL+EFL and EFL settings do not develop simultaneously. As already presented, tradeoff effects were observed in both groups between syntactic complexity and accuracy from T0 to T3, which implies that learners' accuracy levels will decrease as their syntactic complexity increases. Syntactic complexity and fluency, in contrast, do develop simultaneously. The importance of these findings within instructed SLA lies in the attitude towards errors in primary school. As pointed out by many researchers, errors are necessary for the development of the learners' interlanguage. Furthermore, the emergence of errors is very likely to indicate progress among young language learners. Thus, on the basis of these findings it is advisable to encourage risk-taking attitudes to produce syntactically complex structures rather than prime high levels of accuracy.

The next conclusion relates to the importance of fostering the teaching of formulaic language in primary school. According to the qualitative descriptions, formulaic sequences play a crucial role in the development of the learners'

interlanguage, particularly among young language learners. Taking this into account, the teaching of formulaic language needs to be provided, especially in CLIL settings in which formulaic sequences may help learners verbalise content and language objectives. As seen from the interpretations of the qualitative analyses, young language learners start to modify and restructure already learnt chunks of language to fit their communicative purposes during their last year in primary school mainly. In so doing, they succeed in engaging not only in effective communication but also in important acquisition processes such as restructuring of L2 knowledge.

Finally, several generalisations regarding the linguistic features that are to be expected in the L2 English oral output of learners at the end of primary school can also be highlighted. To start with, young learners are predicted to produce complete SVO structures and coordination. In addition, subordination can also be expected among those learners with high fluency levels. Syntactically complex structures like negation and questions are not predicted to be fully developed at the end of primary school. In regard to accuracy, overuse, substitution and omission errors as well as errors related to the restructuring of formulaic language are also likely to occur. As for lexical complexity, the most relevant change during the last year of primary education is associated with the increase in the use of verbs. Eventually, fluency is expected to improve greatly, although some L1 use and comprehension problems may still persist.

8.1 Limitations of the study and further research

In this final part of the dissertation some of the limitations in the present study will be acknowledged. They are mainly related to the generalisation of the findings and the treatment of the oral data in the analyses. Additionally, suggestions for further research will also be provided.

To begin with, the most relevant limitation concerns the generalisation of the findings. Even though the findings of this study are based on the analysis of 416 transcripts (two tasks and four data collection times), the total number of participants in this study was 52. Likewise, the qualitative descriptions analysed the oral output of 6 learners only. Thus, the elaboration of larger research projects (with similar methodological designs) is strongly encouraged in order to corroborate the findings in this study.

In addition to that, several limitations were observed in the treatment of the oral output. Firstly, the use of a completely new *unit* for the analysis of the oral data implied the impossibility to compare findings across research. Secondly, the study of young learners' lexical complexity by looking into lexical diversity seemed to be insufficient. This is why the examination of other dimensions within lexical complexity might be more adequate. For instance, the study of the emergence of functional categories by means of analysing young learners' lexical density might provide more relevant findings in this area. In addition, it is important to note that data collection sessions were carried by a group of researchers and that each of them adopted different approaches and attitudes during the data collection sessions. That might account for the differences in total task times found across participants, groups and data collection times. Finally, another important limitation detected refers to the division of low and high achievers. As described in the Method chapter, the classification of learners according to their proficiency level was based on a written test which assessed receptive skills and grammar knowledge. The use of fluency measures like speech rate in words would have been more reliable indicators of oral proficiency level.

Limitations aside, this study has succeeded in overcoming important methodological shortcomings found in previous CLIL research. In addition, relevant

findings regarding the effectiveness of CLIL instruction and its effects on the development of young learners' oral production skills have been discussed. In the first place, this study has provided enough evidence to claim that minimal CLIL exposure does not render the expected outcomes in regard to speaking skills. Secondly, a number of pedagogical implications on the importance of implementing teaching practices for the development of learners' oral production skills and establishing clear language objectives in CLIL settings have also been outlined. Finally, data tracing the development of L2 English young learners' oral output has been provided in an attempt to characterise English L2 speech among young language learners and inform foreign language pedagogy.

Altogether, this dissertation has contributed with new and valuable data to further explore to what extent minimal CLIL educational practices in addition to EFL instruction are more beneficial than regular EFL instruction on its own. CLIL has the potential to offer the optimal conditions for language learning to take place in instructional contexts. However, clear language objectives as well as greater amounts of CLIL input are needed in order to foster the development of all language skills, especially speaking skills. In addition, the results obtained in this study indicate that the integration of a more systematic use of focus on form tasks into CLIL teaching practices may also be necessary in order to achieve the presumed benefits of CLIL.

REFERENCES

Aguilar, M a nd M uñoz, C . (2014). T he e ffects of pr oficiency on C LIL b enefits i n Engineering s tudents i n S pain. *International Journal of Applied Linguistics,* 24(1): 1-18.

Álvarez, E . (2006). R ate a nd r oute of a cquisition i n E FL na rrative de velopment a t different as ges. In M uñoz, C . (Ed.) *Age and the Rate of Foreign Language Learning.* (pp. 127-155). Clevedon: Multilingual Matters.

Admiraal, W., Westhoff, G., and de Bot, K. (2006). Evaluation of bilingual secondary education i n t he N etherlands: S tudents' l anguage pr oficiency in E nglish. *Educational Research and Evaluation,* 12(1): 75–93.

Basterrechea, M ., García M ayo, M .P. a nd Leeser, M .J. (2014). Pushed out put a nd noticing in a d ictogloss: T ask imp lementation i n th e C LIL classroom. *Porta linguarum*, 22, 7-22.

Bret, A. (2011). Implementing CLIL in a Primary School in Spain: The effects of CLIL on L2 E nglish Learners' O ral P roduction S kills. M aster's F inal R esearch Project, UAB.

Brewster, J ., E llis, G . a nd G irard, D . (2004). *The Primary English Teacher's Guide.* Harlow: Pearson Education Limited.

Broeder, P., Extra, G., and Van Hout, R. (1993). Richness and Variety in the developing lexicon. I n P erdue, C . (Ed.). *Adult language acquisition: cross-linguistic perspectives.* Volume 1: Field Methods. Cambridge: CUP.

Bruton, A. (2011). Is CLIL so beneficial, or just selective? Re-evaluating some of the research. *System*, 39 (4): 523-532.

Bruton, A. (2013). CLIL: Some of the reasons why… and why not. *System*, 41 (3): 587-597.

Bulté, B. and H ousen, A. (2012). D efining a nd ope rationalising L2 c omplexity. In Housen, A, Kuiken, F and Vedder, I. (Eds.). *Dimensions of L2 Performance and Proficiency: Complexity, Accuracy and Fluency.* (pp. 21-46). Amsterdam: John Benjamins.

Cameron, L. (2001). *Teaching Language to Young Learners*. Cambridge: CUP.

Cenoz, J . (2003). The Influence of A ge on t he A cquisition of E nglish: G eneral Proficiency, A ttitudes a nd C ode M ixing. In G arcía M ayo, M .P. a nd García Lecumberri, M . L (Eds.), *Age and the Acquisition of English as a Foreign Language*. (pp. 77- 93). Clevedon*:* Multilingual Matters.

Cenoz, J., Genesee, F. and Gorter, D. (2014). Critical Analysis of CLIL: Taking stock and looking forward. *Applied Linguistics,* 35(5): 243-262.

Cohen, J.W. (1988). *Statistical power analysis for the behavioural sciences.* Hillsdale, NJ: Lawrence Erlbaum Associates.

Coyle, D. (2006). Developing CLIL: Towards a Theory of Practice. In *CLIL in Catalonia, from theory to practice.* APAC Monographs, 6, 5-25.

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 Ruiz de Zarobe, Y. and Lasagabaster, D. (Eds.), *CLIL in Spain. Implementation, Results and Teacher Training.* (pp. vii-viii). Newcastle upon Tyne: Cambridge Scholars.

Cummins, J. (1979). Cognitive/academic language proficiency, linguistic interdependence, the optimum age question and some other matters. *Working Papers on Bilingualism*, 19, 121-129.

Cummins, J. (1981). The role of primary language development in promoting educational success for language minority students. In California State Department of Education (Ed.), *Schooling and Language Minority Students: A Theoretical Framework.* (pp. 3-49). Los Angeles: Evaluation, Dissemination and Assessment Center California State University.

Cummins, J. (1999). BICS and CALP. Clarifying the distinction. Retrieved from http://eric.ed.gov/?id=ED438551.

Cummins, J. (1999). BICS and CALP: Clarifying the distinction, (Report No. ED438551). Washington, DC: ERIC Clearinghouse on Languages and Linguistics. Retrieved from: http://eric.ed.gov/?id=ED438551.

Cummins, J. (2000). Immersion Education for the Millennium: What We Have Learned from 30 Years of Research on Second Language Immersion. Retrieved from: http://www.iteachilearn.com/cummins/immersion2000.html.

Cummins, J. (2008). BICS and CALP: Empirical and Theoretical Status of the Distinction. In Street, B. and Hornberger, N. H. (Eds.). *Encyclopedia of Language and Education,* 2nd Edition, Volume 2: Literacy. (pp. 71-83). New York: Springer Science + Business Media LLC.

Dalton-Puffer, C. (2007). Outcomes and processes in Content and Language Integrated Learning (CLIL): current research from Europe. In Delanoy, W. & Wolkmann, L. (Eds.), *Future Perspectives for ELT.* Heidelberg: Carl Winter.

Dalton-Puffer, C. (2011). Content–and-Language Integrated Learning: From Practice to Principles? *Annual Review of Applied Linguistic,* 31, 182-204.

Dalton-Puffer, C. and Nikula, T. (2014). Content and Language Integrated Learning. *Language Learning Journal.* 42(2): 117-122.

De Graaf, R., Koopman, G., Anikina, Y. and Westhoff , G. (2007). An Observation tool for E ffective P edagogy in C ontent a nd L anguage Integrated Learning (CLIL), *International Journal of Bilingual Education and Bilingualism,* 10(5): 603-624.

De G raaff, R . a nd H ousen, A . (2009). Investigating t he e ffects a nd e ffectiveness of instruction. In Long, M . a nd D oughthy, C . (Eds.), *The Handbook of Language Teaching.* (pp. 726-755). Oxford: Blackwell.

DeKeyser, R . (1998). B eyond f ocus on f orm: C ognitive pe rspectives on l earning a nd practising s econd l anguage grammar. In D oughty, C . and W illiams, J . (Eds.), *Focus on Form in Classroom Second Language Acquisition.* (pp. 4 2-63). Cambridge: CUP.

DeKeyser, R . (2000). T he r obustness of c ritical pe riod e ffects i n s econd l anguage acquisition. *Studies in Second Language Acquisition, 22*, 493-533.

DeKeyser, R. and Larson-Hall, J. (2005). What does the critical period really mean? In J.F.Kroll a nd A .M.B. de G root (Eds.), *Handbook of Bilingualism: Psycholinguistic approaches.* (pp. 88-108). Oxford: OUP.

Doughty, C . a nd W illiams, J . (1998). Issues a nd t erminology. In Doughty, C . a nd Williams, J. (Eds.) *Focus on Form in Classroom Second Language Acquisition.* (pp. 1-11). New York: CUP.

Drew, I. a nd H asselgreen, A . (2008). Y oung l anguage l earner (YLL) research: A n overview of some international and national approaches. *Acta Didactica Norge,* 2(1): Art 5, 1-18.

Edelenbos, P . a nd K ubanek, A . (2009). E arly foreign l anguage l earning: P ublished research good pr actice and m ain pr inciples. In Nikolov, M . (Ed.), *The Age Factor and Early Language Learning.* (pp. 39-58). Berlin: Mouton de Gruyter.

Ellis, R. (1984). *Classroom Second Language Development*. Oxford: Pergamon.

Ellis, R. (1990). *Instructed Second Language Acquisition.* Oxford: Basil Blackwell.

Ellis, R. (1994) *The study of second language acquisition.* Oxford: OUP.

Ellis, R. (1997). *SLA Research and language teaching*. Oxford: OUP.

Ellis, R. (2005). Principles of Instructed Language Learning. *System,* 33, 209-224.

Ellis, R . (2006). C urrent Issues i n t he T eaching of G rammar: A n S LA P erspective. *TESOL Quarterly*, 40(1): 83-107.

Ellis, R. (2008). *The Study of Second Language Acquisition*. Oxford: OUP.

Ellis, R . (2010). S econd l anguage acquisition, t eacher education and l anguage pedagogy. *Language Teaching,* 43(2),182-201.

Ellis, R and Yuan, F. (2005).The effects of careful within-task planning on oral and written task performance. In Ellis, R (Ed.), *Planning and Task Performance in a Second Language*. Amsterdam: John Benjamins.

EU Commission. (1995). *Teaching and learning- Towards the learning society. White Paper on Education and Training*: Brussels. Retrieved from: http://europa.eu/documents/comm/white_papers/pdf/com95_590_en.pdf

EU 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 of 12.08.1995. Retrieved from: http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=CELEX:31995Y0812%2801%29:EN:HTML

Eurydice European Unit (2006). Content and Language Integrated Learning (CLIL) at school in Europe. European Commission. Eurydice, Belgium.

Ferrari, S. (2012). A longitudinal study of complexity, accuracy and fluency variation in second language development. In Housen, A., Kuiken, F. and Vedder, I. (Eds.), *Dimensions of L2 Performance and Proficiency: Complexity, Accuracy and Fluency.* (pp. 277-298). Amsterdam: John Benjamins.

Foster, P. and Skehan, P. (1996). The influence of planning and task type on second language performance. *Studies in Second Language Acquisition,* 18, 299-323.

Foster, P., Tonkyn, A. and Wigglesworth, G. (2000). Measuring Spoken Language: A Unit for All Reasons. *Applied Linguistics*, 21 (3): 354-375.

García Mayo, M.P. (2011). The relevance of attention to form in communicative classroom contexts. *Estudios de lingüística aplicada*, 11, 11-45.

Gass, S. (1989). Language universals and second language acquisition. *Language Learning,* 39, 497-534.

Genesee, F. (1984). Historical and Theoretical Foundations of Immersion. In *Studies on Immersion Education*. (pp. 32-57). Sacramento: California State Department of Education.

Genesee, F. (1994). *Integrating Language and Content: Lessons from Immersion*. National Centre for Research on Cultural Diversity and Second Language Learning.

Georgiu, S. I. (2012). Reviewing the CLIL puzzle. *ELT Journal,* 66(4): 495-504.

Halbach, A. (2009) The primary school teacher and the challenges of bilingual education. In E. Dafouz (ed.), *CLIL across Educational Levels: Experiences from Primary, Secondary and Tertiary Contexts*. Madrid: Santillana. 19-26.

Harley, B. (1998). The role of focus-on form tasks in child L2 acquisition. In Doughty, C. and Williams, J. (Eds.), *Focus on Form in Classroom Second Language Acquisition.* (pp.156-176) Cambridge: CUP.

Housen, A. and Pierrard, M. (2005). Investigating Instructed Second Language Acquisition. In Housen, A. and Pierrard, M. (Eds.), *Investigations in Instructed Second Language Acquisition.* (pp. 1-30). Berlin: Mouton de Gruyter.

Housen, A., and Kuiken, F. (2009). Complexity, accuracy and fluency in Second Language Acquisition. *Applied Linguistics,* 30(4): 461-473.

Housen, A., Schoonjans, E., Janssens, S., Welcomme, A., Schoonheere, E., and Pierrard, M. (2011). Conceptualizing and measuring the impact of contextual factors in instructed SLA – the role of language prominence. *IRAL*, 49, 83-112.

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.

Hunt, K. (1965). *Grammatical structures written at three grade levels*. Champaign, Ill: National Council of teachers of English.

Hunt, K. (1966). Recent measures in syntactic development. *Elementary English,* 43, 732-739.

Hunt, K. (1970). *Syntactic maturity in school-children and adults.* Monograph of the society for research into child development.

Hüttner, J. and Rieder-Bünemann, A. (2007). The effect of CLIL instruction on children's narrative competence. *Vienna English Working Papers,* 16(3): 20-27.

Hüttner, J. and Rieder-Bünemann, A. (2010). A Cross-sectional Analysis of Oral Narratives by Children with CLIL and Non-CLIL Instruction. In Dalton-Puffer, C., Nikula, T. and Ute, S. (Eds.), *Language Use and Language Learning in CLIL Classrooms.* (pp. 61-80). Amsterdam: John Benjamins.

Iwashita, N., Brown, A., McNamara, T. and O'Hagan, S. (2008). Assessed Levels of Second Language Speaking Proficiency: How Distinct?, *Applied Linguistics,* 29(1): 24-49. doi:10.1093/applin/amm017.

Jäppinen, A. K. (2005). Thinking and content learning of mathematics and science as cognitional development in content and language integrated learning (CLIL): Teaching Through a Foreign Language in Finland. *Language and Education*, 19 (2): 147-168.

Järvinen, H. M. (2005). Language learning in Content-based instruction. In A. Housen and M.Pierrard (Eds.), *Investigations in instructed second language acquisition*. (pp. 433-456). Berlin: Mouton de Gruyter.

Jiménez Catalán, R. M., Ruiz de Zarobe, Y. & Cenoz, J., (2006). "Vocabulary profiles of English Foreign Language Learners in English as a Subject and as a Vehicular Language". In Dalton-Puffer, C. and T. Nikula (Eds.) *Vienna English Working Papers (Special Issue: Current Research on CLIL 1),* 15(3): 23-27.

Johnstone, R. (2002). *Addressing the age factor: some implications for language policy.* Retrieved from https://www.coe.int/t/dg4/linguistic/Source/JohnstoneEN.pdf

Juan-Garau, M. (2010). Oral fluency development in secondary education CLIL learners. In Dalton-Puffer, C., Smit, U. & Schiftner, B. (Eds.) *Vienna English Working Papers (Special Issue: Current Research on CLIL),* 19(3), 42-48.

Kormos, J. and Dénes, M. (2004). Exploring measures and perceptions of fluency in the speech of second language learners, *System,* 32(2): 145-164.

Krashen, S. (1981). *Second Language Acquisition and Second Language Learning.* Oxford: Pergamon.

Krashen, S. (1985). *The Input Hypothesis: Issues and Implications*. London: Longman.

Kuiken, F. and Vedder, I. (2012). Syntactic complexity, lexical variation and accuracy as a function of task complexity and proficiency level in L2 writing and speaking. In Housen, A., Kuiken, F. and Vedder, I. (Eds.), *Dimensions of L2 Performance and Proficiency: Complexity, Accuracy and Fluency.* (pp.143-170). Amsterdam: John Benjamins.

Langman, J. (2003). The effects of ESL-trained content-area teachers: Reducing middle-school students to incidental language learners. *Prospect*, 18(1): 14-26.

Larsen-Freeman, D. (2006). The emergence of complexity, fluency, and accuracy in the oral and written production of five Chinese learners of English. *Applied Linguistics,* 27(4): 590-619.

Larson-Hall, J. (2008). Weighing the benefits of studying a foreign language at a younger starting age in a minimal input situation. *Second Language Research,* 24(1): 35-63.

Lasagabaster, D. (2008). Foreign Language Competence in Content and Language Integrated Courses. *The Open Applied Linguistics Journal*, 1, 30-41. doi: 1874-9135.

Lasagabaster, D. and Sierra, J.M. (2010). Immersion and CLIL in English: more differences than similarities. *ELT Journal*, 64(4): 367-375.

Lázaro, A. and García Mayo, M.P. (2012). L1 use and morphosyntactic development in the oral production of EGL learners in a CLIL context. *International Review of Applied Linguistics,* 50, 135-160.

Lenneberg, E. (1967). *Biological Foundations of Language*. New York: Wiley.

Levkina, M. and Gilabert, R. (2012). The effects of task complexity on L2 oral production. In Housen, A., Kuiken, F. and Vedder, I. (Eds.), *Dimensions of L2 Performance and Proficiency: Complexity, Accuracy and Fluency.* (pp. 171-198). Amsterdam: John Benjamins.

Lightbown, P. (2000). Anniversary article. Classroom SLA research and Second Language Teaching. *Applied Linguistics,* 21(4): 431-462.

Lightbown, P. and Spada, N. (2000). Focus on form and corrective feedback in communicative language teaching: Effects on second language learning. *Studies in Second Language Acquisition,* 12(4): 429-448.

Lightbown, P. and Spada, N. (2006). *How languages are learned.* Oxford: OUP.

Lim Falk, M. (2008). Swedish in an English-speaking school context: Subject-related language use in two upper secondary classes. PhD diss., Stockholm University.

Llinares, A and Whittaker, R. (2007). Talking and writing in a foreign language in CLIL contexts: A linguistic analysis of secondary school learners of geography and history. *Revista española de lingüística aplicada,* Vol. extra 1, 83-91.

Llinares, A and Dafouz, E. (2010). Content and Language Integrated Programmes in the Madrid Region: Overview and Research Findings. In Lasagabaster, D. and Ruiz de Zarobe, Y. (Eds.), *CLIL in Spain. Implementation, Results and Teacher Training.* (pp.95-114). Newcastle upon Tyne: Cambridge Scholars Publishing.

Long, M. (1990). Maturational constraints on language development. *Studies in Second Language Acquisition,* 12, 251-285.

Long, M. (1991). Focus on Form: a design feature in language teaching methodology. In de Bot, Ginsberg, and Kramsch (Eds.), *Foreign language research in cross-cultural perspective.* (pp. 39-52). Amsterdam: Benjamins

Long, M. (2009). Methodological principles for language teaching. In Long, M. and Doughthy, C. (Eds.), *The Handbook of Language Teaching.* (pp. 373-394). Oxford: Blackwell.

Long, M. and Robinson, P. (1998). Focus on form: Theory, research and practice. In Doughty, C. and Williams, J. (Eds.), *Focus on Form in Classroom Second Language Acquisition.* (pp. 15-41). Cambridge: CUP.

Lorenzo, F. (2007). An analytical framework of language integration in L2 content-based instruction courses: The European dimension. *Language and Education,* 21(6): 502-214.

Lorenzo, F. (2010). CLIL in Andalusia. In Lasagabaster, D. and Ruiz de Zarobe, Y. (Eds.), *CLIL in Spain. Implementation, Results and Teacher Training.* (pp.2-11). Newcastle upon Tyne: Cambridge Scholars Publishing.

Lorenzo, F., Casal, S. and Moore, P. (2010). The effects of content and language integrated learning in European education: key findings from the Andalusian bilingual sections evaluation project. *Applied Linguistics,* 31(3): 418–42.

Luoma, S. (2004). *Assessing Speaking.* Cambridge: CUP.

Lyster, R. (1998). Immersion Pedagogy and Implications for Language Teaching. In Cenoz, J. and Genesse, F. (Eds.), *Beyond Bilingualism.* (pp. 64-95). Clevelond: Multilingual Matters.

Lyster, R. (2007). *Learning and Teaching Languages through content: A counterbalanced approach*. Amsterdam: John Benjamins.

Lyster, R. and Sato, M. (2013). Skill acquisition theory and the role of practice in L2 development. In Mayo, M., Gutierrez-Mangado J. and Adrian, M. (Eds.), *Contemporary approaches to second language acquisition.* (pp. 71-92). Amsterdam: John Benjamins.

Marsh, D. (1994). *Bilingual Education and Content and Language Integrated Learning. International Association for Cross-cultural Communication, Language Teaching in the Member States of the European Union (Lingua).* University of Sorbonne. Paris.

Marsh, D. (2002). *CLIL/EMILE – The European Dimension: Action, Trends and Foresight Potential.* Strasbourg: European Commission.

McKay, P. (2006). *Assessing Young Language Learners.* Cambridge: CUP.

Merisuo-Storm,T. (2007). Pupils' attitudes towards foreign-language learning and the development of literacy skills in bilingual education. *Teaching and Teacher Education,*23, (2): 226-235.

Met, M. (1998). Curriculum Decision-making in Content-based Language Teaching. In Cenoz, J. and Genesse, F (Eds.), *Beyond Bilingualism. Multilingualism and Multilingual Education.* (pp. 35-63). Clevelond: Multilingual Matters.

Mewald, C. (2007). A Comparison of Oral Language Performance of Learners in CLIL and in Mainstream Classes at Lower Secondary Level in Austria. In Dalton-Puffer et al. (Eds.), *Syntax, Style and Grammatical Norms.* (pp. 139-173). Frankfurt: Peter Lang.

Miret, M. (2009). CLIL effects on writing performance and overall proficiency. MA thesis. Unpublished manuscript. Barcelona. University of Barcelona.

Mohan, B. (1986). *Language and Content*. Reading, Massachusetts: Addison-Wesley Publishing Company.

Mora, J., (2006). Age Effects on Oral Fluency Development. In Muñoz, C. (Ed.) *Age and the Rate of Foreign Language Learning.* (pp. 65 - 88). Clevedon: Multilingual Matters.

Muñoz, C. (2003). Variation in Oral Skills Development and Age of Onset. In García Mayo, M.P. and García Lecumberri, M. L (Eds.), *Age and the Acquisition of English as a Foreign Language*, (pp. 161- 181). Clevedon: Multilingual Matters.

Muñoz, C. (Ed.). (2006). *Age and the Rate of Foreign Language Learning.* Clevedon: Multilingual Matters.

Muñoz, C. (2007). CLIL: Some thoughts on its psycholinguistics principles, *Revista española de lingüística aplicada,* Volumen monográfico 1, 17-26.

Muñoz, C., Tragant, E. and Torras, M.R. (2010). Los inicios de la producción oral en lengua extranjera de alumnos de primaria en Europa, *ELIA,* 10, 11-39.

Myles, F. (2012). Complexity, accuracy and fluency: The role played by formulaic sequences in early interlanguage development. In Housen, A., Kuiken, F. and Vedder, I. (Eds.), *Dimensions of L2 Performance and Proficiency: Complexity, Accuracy and Fluency.* (pp. 171-198). Amsterdam: John Benjamins.

Navés, T. (2009). Effective Content and Language Integrated Programmes. In Ruiz de Zarobe, Y. and Jimenez Catalán, R. (Eds.), *Content and Language Integrated Learning Evidence from Research in Europe*. (pp. 22 -40). Clevedon: Multilingual Matters.

Navés, T. and Victori, M. (2010). CLIL in Catalonia: An overview of research studies. In Lasagabaster, D. and Ruiz de Zarobe, Y. (Eds.), *CLIL in Spain. Implementation, Results and Teacher Training.* (pp. 30 -54). Newcastle upon Tyne: Cambridge Scholars Publishing.

Nikolov, M. (2009). The age factor in context. In Nikolov, M. (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. (2005). English as an object and tool of study: interactional effects and pragmatic implications. *Linguistics and Education*, 16 (1): 27-58.

Nikula, T. and Marsh, D. (1998). Terminological considerations regarding Content and Language Integrated Learning. *VALS-ALSA,* 67, 13-18.

Norris, J. and Ortega, L. (2000). Effectiveness of L2 instruction: A research synthesis and quantitative meta-analysis. *Language Learning,* 50 (3): 417-528.

Norris, J. and Ortega, L. (2009). Measurement for understanding: Towards an Organic Approach to Investigating CAF in instructed SLA: The Case of Complexity. *Applied Linguistics,* 30(4): 555-578. doi:10.1093/applin/amp044.

Oh, S. (2006). Investigating the Relationship between Fluency Measures and Second Language Writing Placement Test Decisions. Unpublished Master's Scholarly Paper. University of Hawai'i.

Ortega, L. (2009). Sequences and processes in language learning. In Long, M. and Doughthy, C. (Eds.), *The Handbook of Language Teaching.* (pp. 91-105). Oxford: Blackwell.

Palloti, G. (2009). CAF: Defining, Refining and Differentiating Constructs. *Applied Linguistics.* 30(4): 590-601.

Paradis, M. (2004). *A Neurolinguistic Theory of Bilingualism.* Amsterdam: John Benjamins.

Pérez-Cañado, M.L. (2012). CLIL research in Europe: Past, present and future. *International Journal of Bilingual Education and Bilingualism*, 15 (3): 315-341.

Pérez-Vidal, C. (2007). The Need for Focus on Form (FoF) in Content and Language Integrated Approaches: an Exploratory Study. *Revista española de lingüística aplicada,* Volumen monográfico 1, 39-45.

Piaget, J. (1926). *The language and thought of a child*. London: Routledge.

Pica, T. (1983). Adult acquisition of English as a second language under different conditions of exposure. *Language Learning,* 33, 465-497.

Pica, T., Halliday, L., Lewis, N. and Morgenthaler, L. (1989). Comprehensible outputs as an outcome of linguistic demands on the learner. *SSLA,* 11(1): 63-90.

Pienemann, M. (1998). *Language Processing and second language development: Processability theory*. Amsterdam: John Benjamins.

Pienemann, M., Johnston, M., Brindley, G. (1988). Constructing an acquisition-based procedure for second language assessment. *Studies in Second Language Acquisition,* 10, 217-43.

Pinter, A. (2006). *Teaching Young Language Learners.* Oxford: OUP.

Piske, T. (2010). Positive and Negative Effects of Exposure to L2 Orthographic Input in the Early Phases of Foreign Language Learning: A Review. In B. Diehr and J. Rymarczyk (Eds.), *Researching Literacy in a Foreign Language among Primary School Learners*. (pp. 37-50). Frankfurt/Main: Peter Lang.

Pladevall Ballester, E. (in press) CLIL subject selection and young learners' listening and reading comprehension skills. *International Journal of Applied Linguistics.*

Pladevall-Ballester, E. (2014) Exploring primary school CLIL perceptions in Catalonia: students', teachers' and parents' opinions and expectations, *International Journal of Bilingual Education and Bilingualism*, 2014: 1-15. doi: 10.1080/13670050.2013.874972.

Robinson, P. (1995). Task complexity and second language narrative discourse. *Language Learning,* 45(1): 99-140.

Robinson, P. (2001). Task complexity, cognitive resources, and syllabus design: a triadic framework for examining task influences on SLA. In Robinson, P. (Ed.), *Cognition and Second Language Instruction.* (pp. 287-318). Cambridge: CUP.

Robinson, P. (2007). Task complexity, theory of mind, and intentional reasoning: Effects on L2 speech production, interaction, uptake and perceptions of task difficulty. *IRAL, 45,* 193-213.

Robinson, P., Cadierno, T. and Shirai, Y. (2009). Time and Motion: measuring the effects of the conceptual demands on tasks on second language speech production. *Applied Linguistics,* 30(4): 533-554.

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 competences benefit from CLIL? An insigight into applied linguistics research. In Ruiz de Zarobe, Y., Sierra, J.M., Gallardo de l Puerto, F. (Eds.), *Content and Foreign Language Integrated Learning: Contributions to Multilingualism in European Contexts.* (pp. 129-153). Bern: Peter Lang.

Ruiz de Zarobe, Y. and Lasagabaster, D. (2010). CLIL in a Bilingual Community: The Basque Autonomous Country. In Lasagabaster, D. and Ruiz de Zarobe, Y. (Eds.), *CLIL in Spain. Implementation, Results and Teacher Training.* (pp.12-29). Newcastle upon Tyne: Cambridge Scholars Publishing.

Schmidt, R. (1990). The role of consciousness in second language learning. *Applied Linguistics,* 11, 129-158.

Schmidt, R. (2001). Attention. In Robinson, p. (ed.), *Cognition and second language instruction*. (pp. 3-33). Cambridge: CUP.

Scovel, T. (1988). *A time to speak: A psycholinguistic inquiry into the critical period for human speech.* Rowley, MA: Newbury House.

Seikkula-Leino, J. (2007). CLIL learning: achievement levels and affective factors. *Language and Education*, 21(4): 328-341.

Serra, C. (2007). Assessing CLIL at primary school. A longitudinal study. *The International Journal of Bilingual Education.*, 10(5):582-602.

Short, D.J. (1993). Assessing Integrated Language and Content Instruction. *TESOL Quarterly,* 27(4): 627-656.

Singleton, D. (2003). Critical Period or General Age Factor(s). In García Mayo, M.P. and García Lecumberri, M.L (Eds.), *Age and the Acquisition of English as a Foreign Language*. (pp. 3-22). Clevedon*:* Multilingual Matters.

Singleton, D. (2005). The critical period hypothesis: A coat of many colours. *International Review of Applied Linguistics*, *43*, 269-285.

Skehan, P. (1996). A framework for the implementation of task-based instruction. *Applied Linguistics,* 17(1): 38-62.

Skehan, P. (1998). *A cognitive approach to language learning.* Oxford: OUP.

Skehan, P. (2003). Task-based instruction. *Language Teaching.* 36(1): 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. (2012). Complexity, accuracy, fluency and lexis in task-based performance: a synthesis of the Ealing research. . In Housen, A., Kuiken, F. and Vedder, I. (Eds.), *Dimensions of L2 Performance and Proficiency: Complexity, Accuracy and Fluency.* (pp. 199-220). Amsterdam: John Benjamins, 2012

Snow, M., Met, M., and Genesee, F. (1989). A Conceptual Framework for the Integration of Language and Content in Second/Foreign Language Instruction. *TESOL Quarterly*, 23 (2): 201-217.

Somers, T. and Surmont, J. (2012). CLIL and Immersion: how clear-cut are they? *ELT Journal,* 66(1): 113-116.

Spada, N. and Tomita, Y. (2010). Interactions between type of instruction and type of language feature: A meta-analysis, *Language Learning,* 60(2): 263-308.

Stoller, F.L. (2004). Content-Based Instruction: Perspectives on Curricular Planning. *Annual Review of Applied Linguistics,* 24, 261-283.

Stotz, D. and Meuter, T (2003). Embedded English: Integrating content and language in a Swiss primary school project. In *Le Bulletin Swisse de Linguistique Apliquée* 77, 83-101.

Swain, M. (1985). Communicative competence: some roles of comprehensible input and comprehensible output in its development. In Gass, S., Madden, C. (Eds.), *Input and Second Language Acquisition.* Newbury House, Rowley, MA, pp. 235–252.

Swain, M. (1988). Manipulating and complementing content teaching to maximize language learning. *TESL Canada Journal,* 6(1): 68-83.

Swain, M. (1993). The Output Hypothesis. Just Speaking and Writing aren't Enough, *Canadian Modern Language Review,* 50(1): 158-164.

Swain, M. (1998). Focus on Form through Conscious Reflection. In Doughty and Williams (Eds.), *Focus on Form in Classroom Second Language Acquisition.* (pp. 64-82). New York: Cambridge.

Sylvén, L.K. (2004). Teaching in English or English Teaching? On the Effects of Content and Language Integrated Learning on Swedish Learners' Incidental Vocabulary Acquisition. PhD diss., University of Gothenburg.

Sylvén, L.K. (2013). CLIL in Sweden – why does it not work? A metaperspective on CLIL across contexts in Europe. *International Journal of Bilingual Education and Bilingualism*, 16:3, 301-320.

Vallbona, A. (2014). L2 competence of young language learners in science and arts CLIL and in EFL instruction contexts. A longitudinal study. PhD diss.. UAB.

Vallbona, A. (2011). Implementing CLIL in the Primary Classroom. Research and Future Challenges .In Urmeneta, C. E., Evnitskaya, N., Moore, E., and Patiño, A. Implementing CLIL in the primary classroom: Results and future challenges. *AICLE–CLIL–EMILE: Educació plurilingüe: Experiencies, research and polítiques* 2. (151-157).

Vallbona, A. (2009). Implementing CLIL Methodology in Primary Education. A Case Study. MA thesis, Bellaterra: UAB.

Van de Craen, P., Mondt, K., Allain, L. and Gao, Y. (2007). Why and how CLIL Works? An outline for a CLIL theory. *Vienna English Working Papers.* 16(3): 70-78.

Várkuti, A. (2010). Linguistic Benefits of the CLIL Approach: Measuring Linguistic Competences, *International CLIL Research Journal*, 1(3): 201-221.

Vermeer, A. (2001). Breadth and depth of vocabulary in relation to L1/L2 acquisition and frequency of input. *Applied Psychilinguistics,* 22, 217-234.

Wannagat, U. (2007). Learning through L2-Content and Language Integrated Learning (CLIL) and English as a Medium of Instruction (EMI), *The International Journal of Bilingual Education and Bilingualism,* 10(5): 663-682.

Wolfe-Quintero, K., S. Inagaki, and H.-Y. Kim. (1998). *Second Language Development in Writing: Measures of Fluency, Accuracy, and Complexity*. University of Hawai'i, Second Language Teaching and Curriculum Center.

Wolff, D. (2007). CLIL: Bridging the Gap between School and Working Life. In D. Marsh and D. Wolff (Eds.), *Diverse Contexts- Converging Goals: CLIL in Europe*. (pp. 15-26). Frankfurt: Peter Lang.

Wood, D. (2002). Formulaic language in acquisition and production: Implications for teaching, *TESL Canada Journal*, 20(1): 1-15.

Zydatiß, W. (2007). *Deustch.Englische Zúge in Berlin: Eine evaluation des bilingualen sachfachunterrichts an gymnasien. Kontext, kompetenzen, konsequezen.* Frankfurt- am-Main: Peter-lang.

APPENDICES

Appendix A: Interview task

1. Hi. How are you?
2. My name's… What's your name and surname?
3. How old are you?
4. Tell me about your family…
   4.1 Have you got any brothers or sisters?
   4.2 What is your sister's /brother's name?
   4.3 How old is your brother / sister?
   4.2 What is your parents' job? For example, I am a teacher, what's your mother's/ father's job?
5. What do you usually do at the weekends (on Saturday and Sunday)?
   5.1 Do you visit your grandparents at the weekends?
   5.2 Do you practise any sports? What sport do you play?
   5.3 Do you go out with your parents? What do you do ?
6. What did you do last weekend?
7. And what are you going to do next Sunday?
   7.1 What are your plans for next weekend?
8. Now your turn: Ask me a few questions.

Appendix B: Picture-elicited narrative task

This is a story of a boy, a girl, a mother and their dog. Have a look at the pictures and tell me the story.

Appendix C: Complete transcripts (highly fluent CLIL and control learners)

Highly fluent CLIL learner: T0

Interview

1 @Begin

2 @Languages: en, ca

3 @ Participants: SUB_Subject: H ighly f luent C LIL le arner Subject, INV _Investigator

4 @Coder: ……

5

6 *INV: ok hi how are you ?

7 *SUB: I am good .

8 *INV: my name is …..and what is your name and surname ?

9 *SUB: ……...

10 *INV: ok how old are you ?

11 *SUB: ten years old .

12 *INV: tell me about your family .

13 *INV: for example have you got any brothers or sisters ?

14 *SUB: one sister and one father and one mother .

15 *INV: what is your sister's name ?

16 *SUB: ……...

17 *INV: how old is she ?

18 *SUB: eight years .

19 *INV: what are your parents' job ?

20 *SUB: 0 [= says nothing]

21 * INV: I am for example I am a teacher I teach English .

22 *INV: what is your mother's job ?

23 *SUB: my f ather i s a t eacher m y f ather n o@s:c s é@s:c co m@s:c es @s:c diu@s:c.

24 *INV: now what do you do at the weekends ?

25 *INV: on Sturday and sunday what do you do ?

26 *SUB: no .

27 *INV: you do nothing .

28 *INV: you stay at home ?

29 *SUB: yes .

30 *INV: do you practise any sports ?

31 *SUB: yes .

32 *INV: what do you do ?

33 *SUB: taewkondo football and tennis .

34 *INV: do you go out with your parents ?

35 *SUB: yes go to bicycle .

36 *INV: ok last weekend what did you do ?

37 *INV: las saturday and sunday what do you do ?

38 *SUB: in home .

39 *INV: ok and what are you going to do next next weekend ?

40 *INV: what are your plans for next weekend nest saturday and sunday ?

41 *INV: what are your plans ?

42 *INV: have you got a lot of plans ?

43 *SUB: no .

44 *INV: no .

45 *INV: ok now it is your turn ask me questions now .

46 *SUB: in the weeks of the ara@s:c no@s:c em@s:c surt@s:c .

47 *SUB: how old are you ?

48 *INV: I am thirty years old .

49 *SUB: and your family ?

50 *INV: I have got two brothers and my parents are retired .

51 *SUB: who is the job of your fathers ?

52 *INV: my father used to be a mechanic and my mother used to be a cheff .

53 *INV: ok .

54 @End

Narrative T0

1 @Begin

2 @Languages: en, ca

3 @ Participants: SUB_Subject: H ighly f luent C LIL le arner Subject, INV _Investigator

4 @Coder: ……

5 *INV: now this is the story of a boy a girl a mother and their dog

6 *INV: have a look at the story and tell me the story .

7 *SUB: this boys are cooking hmm@p a cake .

8 *SUB: in the picture two the dog is looking in the basket and the mum is [//] and the mum and the boys are looking one map .

9 *SUB: the picture three the boys go out the home and his mothers hmm@p says bye .

10 *SUB: in the picture four the boys are in the mountains in the camp@s:c two cows and xxx .

11 *SUB: and the boy and the girl open the basket and the dog is go out .

12 *SUB: the girl are surprise .

13 *SUB: the boys are looking in the basket and in the basket no have [/] no have the cake .

14 *SUB: and the dog is go out .

@End

Highly fluent CLIL learner: T1

Interview

1 @Begin

2 @Languages: en, ca

3 @ Participants: SUB_Subject: H ighly f luent C LIL le arner Subject, INV _Investigator

4 @Coder: ……..

5 *INV: ok hi how are you ?

6 *SUB: good .

7 *INV: ok my name is …….what's your name ?

8 *SUB: …….

9 *INV: and your surname ?

10 *SUB: ……..

11 *INV: ho [//] how old are you ?

12 *SUB: eleven .

13 *INV: good now Arnau tell me about your family

14 *SUB: oof

15 *INV: have you got any brothers or sisters ?

16 *SUB: one sister and one father and one mother .

17 *INV: ok and what's your sister's name ?

18 *SUB: …….

19 *INV: and how old is she ?

20 *SUB: eight years .

21 *INV: eight # s he's e ight ok # and w hat a bout your m other ? w hat a bout your mother's job ?

22 *SUB: teacher .

23 *INV: she's a teacher # an English teacher ?

24 *SUB: no .

25 *INV: for primary or secondary school ?

26 *SUB: hmm@p primary .

27 *INV: and your father ? what's your father's job ?

28 *SUB: it i s a cap @s:c d e@s:c em pressa@s:c n o@s:c s e@s:c co m@s:c es@s:c diu@s:c .

29 *INV: ok he's a manager .

30 *INV: ok A rnau w hat do you usually do a t t he w eekends ? on Saturdays a nd Sunday what do you usually do ?

31 *SUB: in s aturday I go # t o play foot+ball at e leven o'clock and the other time no I +...

32 *INV: you play playstation ?

33 *SUB: no .

34 *INV: no ? you don't like it ? ok

35 *SUB: 0 [= says nothing]

36 *INV: ok a nd hm m@p w hat did y ou do l ast w eekend ? l ast S aturday a nd Sunday what did you do ?

37 *SUB: I go in [/] # in Barcelona and in [/] Lloret de Mar # in the beach .

38 *INV: and what about next weekend ? next next ?

39 *SUB: I don't know .

40 *INV: you don't yet ? don't you have plans ?

41 *SUB: no .

42 *INV: ok and hmm@p do you visit your grandparents ? are you going to visit your grandparents ?

43 *SUB: no # I have one .

44 *INV: you've got one grandfather or grandmother?

45 *SUB: grandmother .

46 *INV: ok A rnau # a sk m e qu estions # now you a re the one w ho w ill a sk questions

47 *SUB: hmm@p how old are you ?

48 *INV: I'm twenty+eight .

49 *SUB: what is your job ?

50 *INV: I'm a school teacher I'm an English teacher .

51 *SUB: and your family ?

52 *INV: my family ? what do you want to know about my family ?

53 *SUB: sisters ? brothers ?

54 *INV: yes I have one brother # his name is …….

55 *SUB: es@s:c el@s:c que@s:c va@s:c a@s:c venir@s:c ?

56 *INV: no hahahaha he's not my brother .

57 *SUB: how years # have # he ?

58 *INV: my brother my brother is twenty+five ok Arnau that was good

59 @End

Narrative T1

1 @Begin

2 @Languages: en, ca

3 @ Participants: SUB_Subject: H ighly f luent C LIL le arner Subject, IN V _Investigator

4 @Coder:…….

5 *INV: now this is the story of a boy a girl a mother and their dog # have a look at the pictures and tell me the story

6 *SUB: the boy and the girl preparate a picnic

7 *INV: uhhuh@i

8 *SUB: and the dog go [//] goes in [//] goes on the +…

9 *INV: [“] basket

10 *SUB: ++ [“] basket

11 *SUB: In picture three the boy is going camp@s:c and his mother says bye

11 *INV: uhhuh@i

12 *SUB: in the camp the boys open the basket and the dog goes out

13 *INV: uhhuh@i

14 *SUB: and the # food # que@s:c estava@s:c in the basket the dogs the dog eats

15 *INV: ok well done ……very good

@End

Highly fluent CLIL learner: T2

Interview

1 @Begin

2 @Languages: en, ca

3 @Participants: SUB_Subject: H ighly f luent CLIL l earner S ubject, INV _Investigator

4 @Coder: …….

5 *INV: how are you ?

6 *SUB: I'm good .

7 *INV: my name is ……what is your name and surname ?

8 *SUB: ……..

9 *INV: how old are you ……. ?

10 *SUB: I'm eleven years old .

11 *INV: tell me about your family .

12 *SUB: I have two brothers [//] bueno@s:c one sister and one brother one father one mother and six more parents@s:c but I don't know the name of tiet@s:c .

13 *INV: ok and what's your sister [/] sister's name ?

14 *SUB: …….

15 *INV: and how old is she ?

16 *SUB: I think so its eight .

17 *INV: and your brother ? what's his name ?

18 *SUB: fort .

19 *INV: sorry .

20 *SUB: …….. he has forty days .

21 *INV: forty days ?

22 *SUB: yes .

23 *INV: and

24 *SUB: my father is a boss of empresa@s:c and my mother is a teacher of maths .

25 *INV: what do you usually do at the weekends ?

26 *SUB: I don't know .

27 *INV: what do you do generally .

28 *SUB: it's different every day is different .

29 *INV: yes for example ?

30 *SUB: sometimes I go to the mountain to run bicycle .

31 *INV: ok and what did you do last weekend ?

32 *SUB: I go to the mountain .

33 *INV: and what are you going to do next weekend ?

34 *SUB: I don't know .

35 *INV: ok now it's your turn you ask me questions .

36 *SUB: what about your parents ?

37 *INV: I have parents my father's called …….. and my mother is called ……

38 *SUB: and your job ?

39 *INV: I'm [//] I work at a university I teach the history and sociology of sport and english .

40 *SUB: I don't have more .

41 *INV: just one .

42 *SUB: what's the job of your parents ?

43 *INV: my father was a teacher he taught english and my mother was a nurse .

44 *SUB: 0.

Narrative T2

1 @Begin

2 @Languages: en, ca

3 @Participants: S UB_Subject: H ighly f luent CLIL l earner S ubject, INV _Investigator

4 @Coder: …….

5 *INV: ok very good …….now this is the story of a boy a girl a mother and their dog have a look at the pictures and tell me the story ok .

6 *SUB: picture number the boy has to prepare a picnic .

7 *INV: uhhuh@i .

8 *SUB: and the dog look the bueno@s:c the boys .

9 *INV: uhhuh@i .

10 *SUB: in pi cture num ber t wo w hen m um no [//] when t he bo ys l ook i nto t he map the dog goes into the # [/] into the cest@s:c .

11 *INV: ok .

12 *SUB: the picture number three the boys go out of the home .

13 *INV: uhhuh@i .

14 *SUB: in four he are in the picnic

15 *SUB: in five he open the # bueno@s:c [//] in five the dogs go out of the picniC cest@s:c and in six no are +…

16 *INV: ok

17 *SUB: ++ no are more food in the cest@s:c .

18 *INV: why is the basket empty ?

19 *SUB: ah ["] basket # can you repeat ?

20 *INV: why is the basket empty ?

21 *SUB: the dog eat the food .

22 *INV: ok well done .

@End

Highly fluent CLIL learner: T3

Interview

1 @Begin

2 @Languages: en, ca

3 @Participants: SUB_Subject: Highly fluent CLIL learner Subject, INV _Investigator

4 @Coder: …….

5

6 *INV: hi how are you ?

7 *SUB: good I think so I'm good .

8 *INV: uhhuh@i my name is ….. what's your name ?

9 *SUB: Arnau .

10 *INV: and your surname ?

11 *SUB: Montero .

12 *INV: uhhuh@i how old are you ?

13 *SUB: I have twelve years old and I have the birthday last # month .

14 *INV: oh really happy birthday then .

15 *SUB: thanks .

16 *INV: tell me about your family….

17 *SUB: I have my father works in a factory and he had I think so he had fifty forty+nine years .

18 *INV: and your mother ?

19 *SUB: and m y m other w orks in a s chool i s a m aths i s a pr of@s:c [//] s orry teacher of maths .

20 *INV: uhhuh@i .

21 *INV: do you have any brothers or sisters ?

22 *SUB: the sister had nine years old and she [/] she works too in [/] in this school.

23 *INV: ok hmm@p …….what do you usually do at the weekends ?

24 *SUB: I play foot+ball with my team or I do homework # nothing more .

25 *INV: ok what did you do last weekend ?

26 *SUB: I also do homework and go to the swimming pool with my family .

27 *INV: uhhuh@i good and what about next weekend ? what are your plans for next weekend ?

28 *SUB: I don't know .

29 *INV: are you going to play foot+ball ?

30 *SUB: no .

31 *INV: ok …….. now it's your turn you ask me questions .

32 *SUB: which is your work ?

33 *INV: I am a teacher I teach English .

34 *SUB: in +... ?

35 *INV: In ……. in a school called …..

36 *SUB: and about your family do you have some other teachers in your family ?

37 *INV: No I'm the only teacher in my family .

38 *SUB: Is it because you like teachers work or because your fathers tell you do teachers work ?

39 *INV: No because I like teachers [//] teaching I like teaching my parents work in a hospital and they wanted me to become a doctor and I said I no no no

40 I don't want to work in a hospital I don't want to be a doctor I want to teach and I want to teach English .

41 *SUB: I have very many medics@s:c in my family five or six because I had a lot of parents .

42 *INV: really ? ok very good ……very good.

Narrative T3

1 @Begin

2 @Languages: en, ca

3 @Participants: SUB_Subject: Highly fluent CLIL learner Subject, INV _Investigator

4 @Coder: …….

5 *INV: this is the story of a boy a girl a mother and their dog have a look at the pictures and tell me the story ok .

6 *SUB: first picture the boys are preparing the breakfast and the dog looks about it.

7 *SUB: and in picture number two mum tells him to go to the park to eat the breakfast and the dog is jumping in the entering the cest@s:c

8 *SUB: and in picture three the boys are saying bye to mum and they are going to the park.

9 *SUB: in picture number four the boys are looking for [/] for seat.

10 *SUB: in number five the boys are looking into [/] into the cest@s:c and the dog goes out

11 *SUB: and in number six in the cest@s:c aren't any [/] any food and the dog runs out .

12 *INV: excellent .

@End

Highly fluent control learner: T0

Interview

1 @Begin

2 @Languages: en, ca

3 @Participants: S UB_Subject: H ighly f luent c ontrol learner S ubject, INV _Investigator

4 @Coder: ……

5

6 *INV: how are you ……..[/] how are you ?

7 *SUB: hmm@p I am fine .

8 *INV: ok my name is …….what is your name and surname ?

9 *SUB: yes I am…….

10 *INV: ok and how old are you…….?

11 *SUB: I ten years old .

12 *INV: ok tell me about your family .

13 *SUB: hmm@p my mum his name is ……..and my father ……

14 *INV: any brothers or sisters ?

15 *SUB: hmm@p yes it one sister and his name is ….. .

16 *INV: ok good and what are your parents' jobs ?

17 *SUB: hmm@p .

18 *INV: ok now what do you usually do at the weekends ["] ?

19 *SUB: weekend ["] .

20 *INV: weekend Saturday Sunday what do you usually do ?

21 *SUB: I do [//] I ride the horse and do to the horse and ## nothing .

22 *INV: ok what did you do last weekend ?

23 *SUB: hmm@p similar of all .

24 *INV: and what are you going to do next weekend ?

25 *SUB: similar .

26 *INV: now your turn you ask me questions .

27 *SUB: hmm@p how old are you ?

28 *INV: I am forty+five .

29 *SUB: hmm@p xxx@s:c what is your name ?

30 *INV: my name is……..

31 *SUB: hmm@p ## what is your favourite subject ?

32 *INV: my favourite subject is English .

33 *SUB: what is your mother name's ?

34 *INV: my mother's name is ……..

35 *SUB: and how old are she ?

36 *INV: she is seventy .

37 *SUB: hmm@p what is your favourite sport ?

38 *INV: my favoutite sport is basketball .

39 *INV: ok that is fine .

40 @End

Narrative T0

1 @Begin

2 @Languages: en, ca

3 @ Participants: SUB_Subject: H ighly f luent c ontrol learner S ubject, INV _Investigator

4 @Coder: ……

5 *INV: ok now this is the story……...

6 *INV: this is the story and it is the story of a boy a girl a mother and a dog ok .

7 *INV: can you tell me the story a little bit ? .

8 *SUB: hmm@p it is one boy and one girl and a dog .

9 *SUB: his mum hmm@p ## hmm@p explain the route of this map and hmm@p the dog is in the castel .

10 *INV: uhhuh@i .

11 *SUB: the two boys go to the [/] to the forest

12 *INV: uhhuh@i

13 *SUB: and his dog eat the food and run in the forest .

@End

Highly fluent control learner: T1

Interview

1 @Begin

2 @Languages: en, ca

3 @ Participants: SUB_Subject: H ighly f luent c ontrol learner S ubject, INV _Investigator

4 @Coder: ……

5 *INV: hi how are you ?

6 *SUB: I am fine thank you .

7 *INV: my name is ……and what is your name ?

8 *SUB: I am ……...

9 *INV: and your surname ?

10 *INV: …… .

11 *INV: how old are you Arnau@s:c?

12 *SUB: I am hmm@p ten years old .

13 *INV: ok now tell me about your family .

14 *SUB: my family I say +...

15 *INV: have you got any brothers or sisters ?

16 *SUB: yes I have a siter his name is ……. and he [//] she have hmm@p seven years and my mother is ……and she have hmm@p forty years old .

17 *SUB: my father is …….. and he have hmm@p forty+five years old and I have my grandfathers and grandmothers aunts and uncles .

18 *INV: ok and what are your parents' jobs ?

19 *SUB: hmm@p m y m other's work i n a [/] i n a dentist an d m y f ather's i s a doctor.

20 *INV: now …….. what do you usually do at the weekends ?

21 *INV: on Saturdays and Sundays what do you usually do ?

22 *SUB: hmm@p I go to ride a horse and we went to [/] to ……. with [/] with a cousins we [/] we stay in his house .

23 *INV: ok and what did you do last weekend ?

24 *INV: last saturday and sunday what did you do ?

25 *SUB: hmm@p I ride a horse all saturdays I go to this and I do not remember but ## I think it is I go to no@s:c I do not remember .

26 *INV: you do not remember what you did yesterday .

27 *INV: what did you do yesterday ?

28 *SUB: this ?

29 *INV: yesterday .

30 *INV: today is monday yesterday was sunday what did you do yesterday ?

31 *SUB: but I explain later .

32 *SUB: in this sunday I go to……...

33 *INV: ok and next weekend what are you going to do next weekend?

34 *SUB: next weekend ride a horse of course and in Sunday I think that we go to … .

35 *INV: ok ……..it is your turn ask me questions .

36 *SUB: hmm@p how are you ?

37 *INV: I am fine thank you .

28 *SUB: you are welcome .

39 *SUB: how old are you ?

40 *INV: I am twenty+seven .

41 *SUB: and what is your favourite sport ?

42 *INV: my favourite sport is yoga .

43 *SUB: when you born ?

44 *INV: I was born on 17 March in 1984 .

45 *SUB: and hmm@p I do not know .

46 *SUB: hmm@p today you go to more schools or ?

47 *INV: No when I finish here I will go back to…….

48 *INV: one more question .

49 *SUB: hmm@p what is your favourite animal ?

50 *INV: my favourite animal is the dog .

51 *INV: thank you very good ……..

@End

Narrative T1

1 @Begin

2 @Languages: en, ca

3 @ Participants: SUB_Subject: Highly f luent c ontrol learner S ubject, INV _Investigator

4 @Coder: ……

5 *INV: this is the story of a boy a girl a mother and their dog .

6 *INV: there are six pictures in this story have a look at them and tell me what the story is about .

7 *SUB: hmm@p in th e pi cture num ber one t he bo y a nd the g irl is d oing th e breakfast I think .

8 *SUB: and in picture number two his mum is [//] bueno@s:c he have a map and he point hmm@p the [/] the bueno@s:c the mountain the forest .

9 *SUB: in the number three hmm@p the two boys are in the street and he says bye+bye for him mum .

10 *SUB: in the number four they are arrive in the forest and this .

11 *SUB: in number five hmm@p they look that in [/] in her nest there are a dog and +...

12 *INV: and ?

13 *SUB: ++ in the end in number six he see [//] they see that in the nest do not have the breakast the dog eats .

14 *INV: ok very good .

@End

Highly fluent control learner: T2

Interview

1 @Begin

2 @Languages: en, ca

3 @ Participants: SUB_Subject: H ighly f luent c ontrol learner S ubject, INV _Investigator

4 @Coder: ……

5 *INV: hi how are you ?

6 *SUB: hello I'm fine .

7 *INV: my name is …. what is your name and surname ?

8 *SUB: my name is ……..and my surname are [/] are ……. and … .

9 *INV: how old are you ?

10 *SUB: I'm eleven years old

11 *INV: ok can you tell me about your family ?

12 *SUB: my father are …. and he works in a [/] in a [/] in a hospital .

13 *INV: uhuhh@i

14 *SUB: and my mother is a dentist and his name is ……...

15 *INV: uhhuh@i

16 *SUB: and I have one sister that his name is [//] she's name is …..

17 *INV: uhhuh@i how old is she ?

18 *SUB: he ai [//] she is eight years old .

19 *INV: ok can you tell me what do you usually do at the weekends ?

20 *SUB: in this weekend we [/] we went to [/] to ……..

21 *INV: uhhuh@i

22 *SUB: we have a house in # well +...

23 *INV: ++ in the city of <Girona> ?

24 *SUB: <in the city yes> and we stay there and bueno@s:c we stay there and no@s:c mas@s:c

25 *INV: this is last weekend yes ? normally normally usually what do you do ?

26 *SUB: normally we went to the house of [/] of [/] of ……..this is a farm there are horses and we +...

27 *INV: ok and do you visit your grandparents ?

28 *SUB: what ?

29 *INV: do you visit your grandparents normally ?

30 *SUB: yes well a lways b ueno@s:c # yes always th e [//] th ey [/] th ey h ave a house hmm@p in [/] in front of my house .

31 *INV: do you practise any sports ?

32 *SUB: yes I ride a horse and I do karate .

33 *INV: ok c an you t ell m e i f you ha ve a ny pl ans f or t he l ast [//] f or t he ne xt weekend ?

34 *SUB: no hmm@p I don't have any plans .

35 *INV: ok now its your turn you can ask me a few questions three questions .

36 *SUB: good # how old are you ?

37 *INV: I'm thirty years old .

38 *SUB: what is your favourite colour ?

39 *INV: my favourite colour is black .

40 *SUB: and what is your favourite sport ?

41 *INV: my favourite s port I t hink i s f oot+ball a nd t his weekend w as i mportant no?

42 *SUB: uhhuh@i

@End

Narrative T2

1 @Begin

2 @Languages: en, ca

3 @ Participants: SUB_Subject: H ighly f luent control learner S ubject, INV _Investigator

4 @Coder: ……

5 *INV: now we're going to move onto the second part I'm gonna show you some pictures and then I'm gonna give you a minute it's a story about a boy a girl their mum and their dog so have a look at it tell me when you're ready ok ?

6 *SUB: hmm@p in the firs pic [/] first pic [/] first picture there are a boy and a girl and his mum and they prepare sandwich for a picnic

7 *SUB: in th e s econd p icture h mm@p h er [/] h er [//] h is mu m they [/] they hmm@p they give him a map but the dog he puts in the cist [/] cist [//] castle bueno@s:c cistell@s:c.

8 *SUB: in the other picture hmm@p the boy and the girl they go to the forest and in the forest they [/] they [/] they [/] they they app [//] app [//] they appear the look the dog but it [//] they are in the +...

9 *INV: ++ in the basket .

10 *SUB: ["] in the basket

11 *SUB: the dog go to run

12 *SUB: the boys hmm@p look in the cistell@s:c and they are [//] they aren't the sandwich.

13 *SUB: the dog eat the sandwich .

14 *INV: ok thank you …...

@End

Highly fluent control learner: T3

Interview

1 @Begin

2 @Languages: en, ca

3 @ Participants: SUB_Subject: H ighly f luent c ontrol learner S ubject, INV _Investigator

4 @Coder: ……

5 *INV: hi how are you ?

6 *SUB: I'm fine thank you and you ?

7 *INV: I'm fine thanks my name is …. what is your name ?

8 *SUB: I'm ….

9 *INV: and your surname ?

10 *SUB: ……..

11 *INV: how old are you …….?

12 *SUB: I am twelve years old .

13 *INV: tell me about your family …...

14 *SUB: my f ather i s a [//] he w orks i n t he hos pital a nd hi s na me i s …., …. in English .

15 *INV: yes ……. is ……..

16 *SUB: and my mother is a doctor all my family doctors and his name is …..

17 *INV: ok right …..

18 *SUB: and I have a small sister younger than me .

19 *INV: <yes>

20 *SUB: <and> his name is ……and he do third course .

21 *INV: ok very good and …….. what do you usually do at the weekends ? on saturdays and sundays what do you do ?

22 *SUB: in saturday I go to ride a horse and sundays I do what I want .

23 *INV: and what did you do last weekend ? last saturday and sunday ?

24 *SUB: nothing ride a horse and the normal .

25 *INV: and what are your plans ? what are you going to do for next weekend ?

26 *SUB: hmm@p we go to …...

27 *INV: oh ?

28 *SUB: yes the the the horse running a the competition .

29 *INV: ok well done now ... can you ask me a few questions it's your turn now .

30 *SUB: ok hmm@p how old are you ?

31 *INV: I'm twenty+eight .

32 *SUB: what is your favourite sport ?

33 *INV: my favourite sport well I like yoga very much .

34 *SUB: yoga ? uhhuh@i interesting you relax ?

35 *INV: yes I relax and I do exercise as well .

36 *SUB: and you [/] you go to different schools to do this ?

37 *INV: yes I I'm working here and in ……

38 *SUB: uhhuh@i

39 *INV: there's a school in ……..aswell

40 *SUB: where you live ?

41 *INV: I'm from ……live in ……. .

42 *SUB: xxx

@End

Narrative T3

1 @Begin

2 @Languages: en, ca

3 @ Participants: SUB_Subject: H ighly f luent c ontrol learner S ubject, INV _Investigator

4 @Coder: ……

5 *INV: this is the story of a boy a girl a mother and their dog have a look at the pictures and tell me the story .

6 *SUB: ok i n t he pi cture num ber one w e c an l ook t he bo ys pr eparing t he breakfast and his dog looking.

7 *SUB: in the number two his da [//] her mum hmm@p are showing the map of [/] of the [/] of the [/] yes of the world and his dog is looking to the to the basket.

8 *INV: uhhuh@i

9 *SUB: in the number three the boys walking in the street and say good+bye to his mum .

10 *SUB: in number four the boys are climbing a [//] one small cliff .

11 *INV: uhhuh@i

12 *SUB: in number five they open the basket and they see her [//] his dog .

13 *SUB: in [/] in this and in number six they look another time the basket and they see that the breakfast isn't [/] isn't there .

14 *INV: yeah so ?

15 *SUB: the dog eat [/] the dog eat the breakfast .

16 *INV: ok well done .

@End

Appendix D: Coding symbols

PAUSES

0.5 seconds
greater tan 0.5 seconds
0[=says nothing] participant not answering a question

REPETITIONS

[/] Repetition of words
[//] Retracing with correction

EXPLANATIONS AND ALTERNATIVES

[“] quotation mark

UTTERANCE TERMINATORS

+… uncompleted utterance
++ continuation after interruption / self-completion
+/? Interruption of a question
<….> overlap

UNINTELLIGIBLE MATERIAL

xxx

L1 USE

@s:c

EXCLAMATION AND INTERACTIONAL MARKERS

hmm@p Thinking, waiting
uhhuh@i Yes