A STUDY OF THE ACQUISITION OF ENGLISH AS A FOREIGN LANGUAGE: INTEGRATING CONTENT AND LANGUAGE IN MAINSTREAM EDUCATION IN BARCELONA.

Helena Roquet i Pugès

TESI DOCTORAL UPF / ANY 2011

DIRECTORA DE LA TESI

Dra. Carmen Pérez Vidal (Departament de Traducció i Ciències del Llenguatge)

UNIVERSITAT POMPEU FABRA

Als meus estimadíssims pares

ACKNOWLEDGEMENTS

It is only too difficult to thank all those who have helped me through the long process of elaboration of this doctoral thesis. This research would not have been possible without the cooperation and support of a number of people to whom I would like to express my gratitude.

First of all, I am specially obliged to all those who took part in the tests, as well as the staff at the school involved who made the tests possible and administered them. I am particularly indebted to Ms Maite Díaz, for her coordination of the whole CLIL programme in the school and for her contribution to this scientific research.

My special thanks go to my tutor, Dr. Carmen Pérez-Vidal. She has always been an exceptional teacher and her capacity to work, and to transmit enthusiasm for learning never stop to surprise me. She not only has guided my research with an invaluable knowledge and critical comments even across distances during my stay in Vienna, but has been warm and encouraging when my spirits were low, understanding with my limitations, and demanding towards my research project. I will never forget her wise advice and her personal involvement. It has been a privilege to be her student for such a long time and I very much hope that I can follow her example in my academic life as a teacher for many years to come.

I also owe my deepest gratitude to Prof. Christiane Dalton-Puffer. My stay and research in Vienna would not have been possible without her kind welcome from the very first moment and I am particularly grateful for her clever ideas as well as her warm assistance and support. It has been an honor for me to be a visiting member of the CLIL research group at her department, and to be able to attend the CLIL research seminar there. She has also made possible my consultations with in-house and visiting experts on CLIL and language testing in Vienna.

Together with her, I would like to show my gratitude to Ms Veronika Schindelegger, not only for taking part in this research project and helping me evaluating the tests to find a common criteria to set up a level for the qualitative assessment, but also for her kindness and for becoming a friend during my stay in Austria.

This stay in Vienna was possible thanks to a scholarship from the Departament de Traducció i Ciències del Llenguatge at Universitat Pompeu Fabra (Barcelona) to whom I would like to express my gratitude.

ABSTRACT

The present study seeks to contribute new evidence on the effects of the CLIL approach (Content and Language Integrated Learning) on young EFL learners' productive and receptive skills in a school set in Barcelona (Catalonia). Catalonia is a bilingual community in which both Catalan and Spanish are official languages and English represents the first foreign language included in the curriculum. For that purpose, two groups (n= 50 each) involving 100 bilingual Catalan / Spanish students aged 12 to 15 were analysed longitudinally over two academic years in two different types of exposure contexts: FI (Formal Instruction of English as a foreign language school subject, control group) and CLIL (English as medium of instruction when learning Science, experimental group). Data were elicited both for productive and comprehension skills and were statistically analysed quantitatively and also qualitatively using a posttest design at the end of each academic year. Results obtained confirm the effectiveness of the CLIL programme, however not in all domains and to the same degree as significant benefits did not accrue in all skills and measurements. Concerning receptive skills, when contrasting the differential effects of the two programmes on the participants' linguistic progress, the group in the FI+CLIL improved their reading competence significantly more than the other, as was expected, but not their listening competence. As for productive skills, our findings show a significant improvement in the case of the FI+CLIL group, something which we had not hypothesised, as the subjects' writing and particularly so accuracy, significatively progressed and so did lexico-grammatical abilities. This is in contrast with findings published in previous studies. Results also tend to confirm that age had an impact and thus the older, the better as far as progress made by our subjects. Finally, our results show that the CLIL approach did not seem to erase the differences observed in traditional foreign language teaching contexts when gender is considered: contrary to expectations, female participants still outperformed their male counterparts not only in a FI context but also in a CLIL context. In conclusion, it can be stated that the effectiveness of a CLIL context of learning in this dissertation is confirmed but that it does not suffice to improve the participants' overall linguistic competence as, whereas some levels of language competence made substantial progress, some other levels did not seem to follow the same path.

RESUM

L'objectiu d'aquest estudi és aportar noves dades sobre els efectes de l'enfocament EICLE (Enfocament Integrat de Continguts i Llengües Estrangeres) en les habilitats productives i receptives de subjectes aprenents d'anglès com a llengua estrangera, d'una escola concertada de la ciutat de Barcelona (Catalunya). Catalunya és una comunitat bilingüe on el català i l'espanyol són llengües oficials i l'anglès representa la primera llengua estrangera del currículum educatiu. Amb aquest objectiu s'ha dut a terme un estudi comparatiu de l'adquisició de l'anglès com a llengua estrangera en dos contextos d'aprenentatge. L'estudi adopta un disseny longitudinal (al llarg de dos anys) i compara dos grups (n= 50 cada grup) d'aprenents bilingües català espanyol d'edats compreses entre els 12 i els 15 anys. El grup de control aprèn l'anglès amb l'enfocament convencional en aules d'instrucció formal (IF) i el grup experimental rep l'enfocament EICLE a l'assignatura de naturals en anglès (*Science*) a més de seguir les classes convencionals en context IF. S'han recollit dades relatives a les habilitats de producció i de comprensió i s'han analitzat quantitativament i qualitativament mitjançant un disseny post-test al final de cada any acadèmic tot aplicant anàlisi estadística. Els resultats obtinguts confirmen que el programa EICLE és efectiu tot i que no en tots els àmbits ni amb la mateixa intensitat, ja que no apareixen millores significatives en totes les habilitats i mesures adoptades. Pel que fa a les habilitats receptives, quan es contrasten els efectes diferencials dels dos programes en el progrés lingüístic dels subjectes, el grup EICLE millora significativament més en la prova de comprensió escrita, com estava previst, però no en la de comprensió oral. Pel que fa a les habilitats productives, els nostres resultats demostren una millora significativa en el grup que segueix l'enfocament EICLE. A diferència de la recerca publicada anteriorment, els nostres participants milloren de manera significativa en la producció escrita, especialment en la correcció, i en les habilitats lèxico-gramaticals. Els resultats també tendeixen a confirmar que el factor edat és rellevant i que com més grans millor. Finalment, els resultats demostren que l'enfocament EICLE no ajuda a disminuir les diferències observades en contextos tradicionals d'ensenyament de llengües estrangeres pel que fa al gènere dels subjectes: a diferència del que s'havia previst, els subjectes femenins obtenen millors resultats que els subjectes masculins no només en el context d'IF sinó també en el context EICLE. En conclusió, en aquesta tesi doctoral es pot confirmar l'efectivitat d'un context d'enfocament EICLE. Ara bé, aquest enfocament no és una garantia suficient per a la millora de la competència lingüística general dels aprenents.

TABLE OF CONTENTS

List of Tables

List of Figures

1. INTRODUCTION

Globalisation and internationalisation are making increasing demands on the foreign language skills of European citizens. As a consequence, one of the key features in the European strategy towards multilingualism is an interdisciplinary approach to education. This is a policy in which lesser-known languages, either majority or minority languages in the community are established within a school programme as the medium of instruction for content subjects. This approach has recently been known as Content and Language Integrated Learning (CLIL). In these so-called CLIL classes a language other than the L1 of the students is used in teaching a non-language subject matter, the aim being to increase the students' exposure to the language and to create a motivating low-anxiety environment in which attention is paid to the message conveyed rather than to form. In this way the students' language competence is to be enhanced and they are to be better prepared for life and work in a globalised society and economy, where English in particular dominates as the Lingua Franca of today's world.

The interest of investigating the effects of CLIL contexts of acquisition on linguistic outcomes and processes seems undeniable and undiminishing. One of the main questions in relation to CLIL as a new educational approach arises when analysing to what extent the increased exposure to a target language brought about when adopting CLIL to teach one or more subjects in a school translates into tangible improvement in the quality of language output. In addition, another relevant issue is what aspects of language proficiency are most likely to be affected. This is more particularly interesting when CLIL is analyzed and contrasted together with other contexts of acquisition, with different input conditions, such as formal instruction (FI) in the foreign language classroom.

This research study[1] presents CLIL as a central feature in the European strategy towards multilingualism and its impact on second language acquisition by secondary English as a foreign language (EFL) learners and compares it to FI. Indeed, CLIL and FI allow a close examination of the effects of different contexts of learning on foreign language

---

[1] This study has been developed within the framework of the ALLENCAM research group: Language Acquisition in Multilingual Catalonia, Universitat Pompeu Fabra (www.upf.edu/dcl/recerca)

leamers' linguistic outcomes and attitudes, a question currently under scrutiny in second language acquisition (SLA) research. In this study the terms foreign language acquisition (FLA) and SLA are used alternatively. And only when relevant will the central difference between a foreign and a second language be made.

It is hypothesised that because each of the two contexts has differential patterns of input exposure and offers different opportunities for interaction, both quantitatively and qualitatively, their effect on the participants' communicative and motivational development will also be different.

More specifically, this research study focuses on whether or not the acquisition of a language which is only heard and practised in the language classroom as the object of instruction, i.e. FI, presents significant differences with respect to the acquisition of a language which is only heard and practised in the language classroom as the vehicle of instruction, i.e. CLIL. And, likewise, it focuses on whether or not the degree of infl1uence of individual factors such as age, or gender infl1uence the level of competence achieved in each different context. Since research studies which have covered age and gender issues in the last decades throw contradictory results, we believe that analysing the effects these two individual variables have may contribute with useful data to the FLA research field.

The reason for focusing on the classroom, however, is not merely to shed light on how FLA takes place. Being myself a teacher, this research is also motivated by a desire to discover what classroom conditions or contexts are most likely to facilitate acquisition, whether CLIL or FI, and what exactly does CLIL contribute to education in general and language acquisition in particular. It is important to remember that CLIL initiatives are becoming very popular all over the world. Where it was once believed that the formula for successful foreign language acquisition was "the earlier the better", on the face of the fact that no generalized improvement has come about in EFL in our contexts, nowadays CLIL initiatives may well be seen as a second alternative to the success in FLA (Pérez-Vidal, forthcoming). In other words, there is an inherent indirect goal in the study which falls within the domain of teaching and learning EFL. However, such a pedagogic purpose is not biased towards didactics as in methodological handbooks for teachers, where the aim is to suggest specific techniques or activities that teachers can use. This research considers didactics not in

terms of "techniques" or "activities", but in terms of what kinds of classroom behaviours teachers need to engage in to promote learning in CLIL contexts as opposed to FI. Indeed, we believe that the ultimate goal of research in language acquisition studies is that findings be taken into consideration by educational policies and the educational community in general (both top down and bottom up forces).

Hence, in order to accomplish its main objective, the present dissertation has been structured around seven chapters. Chapter 1 corresponds to the present chapter, the introduction, and explains the main objective of this research, justifies its relevance and describes how it is organised.

In chapter 2 the theoretical background to the study of LA in Formal Instruction Contexts (from now on FIC) is presented from an applied linguistics perspective using a highly canonical structured. The chapter has three main parts. The first one is an overview of SLA/FLA research from the 50's until today. The second one analyses LA and Multilingualism in detail, focusing on Bilingual and Third LA, the issue on focus in our study, while dealing with Catalan/Spanish learners of English as a third language. The third part presents the specific context of acquisition scrutinised: CLIL.

Chapter 3 deals with the research questions and the method used to carry out the present research study. The organisation of the chapter is as follows: Section 3.1 presents the objectives of the investigation. In section 3.2 the research question and the hypotheses used to address the issue analysed are explained. After this, a section with the method is offered. In it the context and participants of the study, the design, treatment, and instruments used, and the data collection procedure are explained in detail. The last part within the method section (3.3) is a large description of the quantitative and qualitative measures used to analyse the data, and the statistics adopted.

Chapter 4 presents the results of the analyses performed in order to answer the research question formulated in Chapter 3. This research question enquired how context of learning affects the oral and written development of young bilingual secondary education EFL learners when contrasting a group experiencing FI only and a group experiencing FI in combination with CLIL. Three specific issues are of interest in this field of research when contrasting the two contexts. The first one, related to general language development and presented in

section 4.1, is whether all linguistic abilities develop accordingly or differently. The subjects examined are measured as far as their writing abilities, their reading abilities, their listening comprehension abilities, and their lexico-grammatical abilities. The second issue, related to individual differences and presented in section 4.2, is whether changes occur irrespective of age differences. The third and last issue corresponds to section 4.3 and relates to another individual variable, gender, which has caught the interest of research in recent years.

Chapter 5 contains the discussion of results. This chapter is organized into different sections each one dealing with one of the three hypotheses established in relation to the main question. Firstly, section 5.1 tackles the issue of language progress and skill development. It discusses the results obtained in order to address the first hypothesis. Secondly, section 5.2 deals with the the issue of language progress and age: whether changes occur irrespective of age differences. In the third place, section 5.3 discusses the results concerning the issue of the impact of gender differences in a FI and a FI + CLIL context in relation to each of the different skills measured.

Finally, based on a summary of the main findings and the discussion, Chapter 6 offers the conclusions reached after carrying out the investigation. It also identifies the limitations of the study presented and suggests several issues for further research.

To conclude, chapter 7 is a list of the bibliographical references included throughout the dissertation. After this, 3 appendices are added in the end. The first one presents the tests administered to the participants. The second one shows a table with the school's CLIL programme development. Finally, in the third and last one, the rating scale used for assessment of the writing task from Friedl/Auer (2007) is shown.

It is very much hoped that the investigation presented here will be able to discretely contribute to the field of SLA research and at the same time be of interest to all those involved in education. This is the right place to play tribute to all the schools which are offering their CLIL programmes as sources of empirical evidence such as the school in which data were collected for the present dissertation. Their contribution to scientific research is tantamount to their capacity for innovation and service to the community. To finish on a personal note, given the effectiveness of the CLIL programme analysed in this

research as far as specially writing skills and accuracy, I have begun to pay attention to content topics in as much as to form in many of my EFL FI classes. As a consequence, I have started to use more authentic texts and topics and in addition chosen by students and thus overcoming the artificiality of FI contexts in EFL. More often than before my students ask questions just for pleasure of finding out more about the topic, and it happens that when students like what they are doing, classes become in turn a much more stimulating experience for the teacher. Because as we all know, there is nothing that equals the gratification of working with motivated students.

I. BACKGROUND TO THE STUDY

2. LANGUAGE ACQUISITION LITERATURE: AN OVERVIEW

In this chapter the theoretical background to the study of LA in Formal Instruction Contexts (from now on FIC) is presented from an applied linguistics perspective. The chapter has three main parts. The first one is an overview of SLA/FLA research from the 50's until today. The second one analyses LA and Multilingualism in detail focusing on Bilingual and Third LA, the issue on focus in our study, dealing with Catalan/Spanish learners of English as a third language. The third part presents the specific context of acquisition scrutinized: CLIL.

The notion of language acquisition (LA from now on) can be approached from a variety of perspectives: sociolinguistic, educational, neurolinguistic, psycholinguistic or linguistic. Each of these approaches involves different assumptions, methods and goals, albeit they all aim at contributing to a better understanding of the processes underlying LA. The present study focuses on the linguistic account of language acquisition, more specifically Second Language Acquisition and Foreign Language Acquisition (from now on SLA and FLA) in relation to a specific European context of learning: Content and Language Integrated Learning (CLIL from now on).

The task awaiting the study of SLA and FLA is vast. It must account for failure as well as success and must cover both naturalistic and formal learning, at all ages, in second and foreign language environments, with monolingual, bilingual or multilingual speakers and with all the variables that these macro considerations imply. For example, consider a group of monolingual students in a state secondary school in Spain, who may be learning a FL, that is a language not spoken in the environment in a conventional Formal Instruction classroom, or following a bilingual immersion programme, compared with a group of multilingual speakers learning a language in the target language community during a Study Abroad (SA) Period. If what is being proposed is a FLA theory, it must be applicable in any of these different contexts of learning. In addition, it must cover all aspects of language: form, meaning, sound, use.

2.1 Formal Instruction Contexts

In this part an overview of S/FLA research is presented. First of all, following a chronological order, four different periods are identified

and described: i. the structuralist-behaviorist period, ii. the Chomskyan period, iii. the social-interactionist period, and iv. the cognitive period. After this, four key hypotheses that provide a rough sketch of the features which come to play a role in the process of acquiring languages are introduced: i. the Input Hypothesis, ii. the Interaction Hypothesis, iii. the Output Hypothesis, iv. and the Noticing Hypothesis. Next, a section dealing with focus on form and acquisition and the current state of research in this domain is presented. Finally, the chapter ends with a section on individual differences in SLA.

2.1.1 Four Different Periods

In this section, developments in the field of Linguistics applied to the acquisition of languages are succinctly presented and organised in four chronological periods, roughly corresponding to four differentiated models. As Pérez-Vidal (2001a) summarises, three distinct periods can be identified since work in the field of Linguistics, understood as the study of language as a system, began around the 50's: a structuralist-behaviorist period, a period named after its main figure, Chomsky, and the social-interactionist period. The main focus in this author's account is placed on the role of input in the process of language acquisition and how it is viewed differently in each period. This is a view which specifically relates to contextual differences in language acquisition, the main focus of the study presented here. After these three, a fourth final period follows, a cognitive period which is included in this presentation in order to show the most recent developments in the field of Linguistics applied to the acquisition of languages, together with a summary of Skill Acquisition Theories.

2.1.1.1 The Structuralist-Behaviorist Period

In the first period of research on LA, conventionally called the structuralist-behaviorist period, the structuralists linguists were close to the behaviorist school of psychology (Skinner, 1957). They believed that learning takes place as a habit formation process, a stimulus-response reaction followed by a positive or a negative reinforcement. It was posited that children learn languages by being encouraged or reinforced when they speak properly, and by being discouraged when they do not speak properly. This implies a view in which adults or carers speak to children and model specific linguistic forms and patterns so that children imitate them and internalise those patterns. If they imitate well, they will be reinforced with a Very good! Right!, if

they don't, they will be corrected. In the case of an L2, native speakers (from now on NSs) will speak to non-native speakers (from now on NNSs); and, in classrooms, teachers will speak to students to help them learn in a similar way to how adults do with children.

What this means, in terms of the role assigned to language learners, is that acquisition is manipulated from outside by choosing what one says (input), and providing appropriate stimuli to speak. It is thus *controlled* from the outside and the learner is *passive* and learns by *analogy*, not by analysis.

2.1.1.2 The Chomskyan Period

In the following period, to put it very succinctly, firstly Chomsky objected to the view that human learning, and specifically language learning, can be explained as the stimuli-response chain (Chomsky, 1959). He rejected that learning is a habit formation process, a process of imitation and analogy, where children acquire a language by imitating more and more complex structures. He objected basically on two accounts. On the one hand, he argued, although children do imitate certain words and structures, they cannot possibly imitate structures that they have never heard before, such as *GO-ED* for past of *go* instead of *WENT*, in the case of English.

Thus, given that imitation cannot explain some of the language produced by children, we can easily say that the structuralist behaviourist paradigm does not work as an explanation of the language acquisition process. It is true that imitation does play a role, and today we know that it is a strategy used by child and adult learners: there is no doubt that some children imitate a great deal, although some imitate much less, and that the same happens with adults in natural and formal acquisition in the classroom. However, the crucial objection put forward in this period is that imitation does not take into consideration a much more essential process which is going on underneath imitation, the complex process of acquiring a language.

Secondly, Chomsky also objected to the idea that children are reinforced by what they say well, and corrected when they do not speak well. Chomsky accumulated evidence that proved that when adults talk to children, what they are mainly concerned about is communication. If something is corrected it is to make sure that the child's contribution to

communication is true, not false, but blatant grammatical errors tend to go uncorrected provided that what the child says is true.

> "He a girl" (a child says of her mother pointing at her)
> The mother accepts in spite of:
>   - wrong gender
>   - and no verb (Pérez-Vidal, 1996: diary record)

As a result of this view, and in contrast with the previous one, the role assigned to the learner is that of being engaged in a mental activity process whereby, in the case of a first language, children use what adults say, the input, to form and test hypothesis about how language is organised, hence trying to infer the rules from the language spoken to them by their carers. Further development of language ability is the process from basic rules to more refined one which will allow the child to incorporate more and more of the language he or she hears. In this way, what is called the interlanguage of the learner develops, that is the type of language produced in the process of learning the second/foreign language. Hence, language acquisition is a creative and rule governed activity. And this, Chomsky claimed, is possible because we are all genetically programmed, we have an inner mechanism which is different from all the others we inherit (so it is language-specific) and unique to humans (species-specific), that allows us to proceed in this way. This inner mechanism includes knowledge of the properties that are common in the basic structure of all languages.

Accordingly, babies learning their first languages and second language learners alike are *active* in the process of learning languages, they proceed by *analysis*, rather than analogy, and they are *creative*, for they are able to use a set number of rules to produce as many sentences as they wish.

Since those days, empirical research on input addressed to children has proved that input is generally correct, although modified, and somehow simplified by adults whose role is to interact with children and help them say what they would not be able to say on their own. Such a kind of input has received several names: 'motherese' baby talk and more recently child directed speech (CDS). This type of assistance has received the name of *scaffolding*. It was proved that such a type of communication contributes to the process of how a child acquires a first language (Gallaway and Fichards, 1994). As it will be further explained in more detail in the following sections, a very powerful model was put

forward by Stephen Krashen (1985) who, on the basis of this evidence, formulated the same principle in relation to second language acquisition. His Monitor Theory, resting on five hypotheses of which the basic one was "The input hypothesis", stated that input which is accessible to the learner, because being simplified to just one level above his or her competence level, might become intake, that is, acquired language. However, this idea that only comprehensible input is necessary for acquisition to take place was subsequently criticised as is explained further below. Other authors became more interested not only in the input addressed to learners, but in the discourse interaction in which learners participate, also drawing from the research in first language acquisition. Hatch (1978) applied the construct of scaffolding to adults learning second languages, and subsequently other authors carried out empirical research in order to prove to what extent the type of input addressed to adult learners was similar to that addressed to children, and whether it also favoured acquisition (Larsen-Freeman and Long, 1991). The terms teacher-talk and foreigner-talk or 'foreignese' came to be used to refer to this specific type of input.

2.1.1.3 The Social- Interactionist Period

The third period in this account of SLA research is the social-interactionist period. The idea that input is necessary in the process of acquiring languages has not changed in this period, however it has been refined in three respects. Firstly, in the sense that input is no longer considered sufficient; it is even seen as insufficient by some authors (Long, 1985). It is understood that in addition to input, specific interactional adjustments are displayed not only by the native speaker in the situation of communication, as was assumed in the preceding period, but also by the learner within the microcosm in which learning takes place. It is proposed that adults learn an L2 by participating in interaction where modifications at the level of language and discourse take place. In this way they benefit from comprehensible input, which allows them to incorporate new syntactic structures in their interlanguage. In addition to that, negotiation of meaning takes place.

Consequently, NSs contribute to the process of language acquisition of adult learners in two ways: by allowing them to negotiate meaning with them, and by adjusting their discourse with the type of modifications needed. Negotiation of meaning is a key phenomenon in the model put forward by social interactionism. It takes place when there are communication problems either due to limitations in the learners'

competence, or the complexity of the situation. In this sense, negotiation is a kind of problem solving strategy applied to language, motivated by a pressure to communicate imposed in a particular situation, something which has recently been stressed by authors such as Gass (1997).

Secondly, if we go back to the mental activity which goes on when trying to establish hypotheses about the rules of the language we hear, in the model put forward by Chomsky it would appear as if all input can be processed with the same degree of efficiency, yet this is not the case. The order of acquisition and the stages of acquisition studies proved that when learning a particular language we first learn some aspects of it and then others (see Ellis, 1994, for a throurough summary or Dulay *et al.*, 1982). And this we all do following the same order. It was also clear that children make the same mistakes on route towards adult competence, and so do adults gaining native-like competence throughout their different stages of interlanguage. Input studies first acknowledged the need for input, and, through this third period, they also investigated the manifold reasons why we learn some bits of language before others. These reasons can be summarised as follows:

1. Learners get positive evidence in the form of input.
2. But this is not sufficient, because they will only notice some phenomena in this evidence, and not others, those utterances they notice will trigger rules which will be language specific (Schmidt, 1990).
3. The learner's output produced will be contrasted with the input, which may contain negative evidence, correction that will help them disconfirm hypothesis they have made.

Thirdly, there is conclusive research showing that learners, both children and adults, who receive either explicit corrections or implicit corrections perform better (Long, 1996). The previous three successive stages, which in real life may not take place in such a neat order, are only possible when learners are engaged in interaction with an interlocutor.

What is the role of the learner in the process of acquiring a language according to this view? It is the *conversational adjustments* and interactional modifications in which they are *actively* engaged when *negotiating meaning* which promote acquisition.

2.1.1.4 The Cognitive Period

Since the late 1980s, there has been a revival of interest in psychological theories of language learning. Cognitive psychologists see no reason to assume that language acquisition requires specific brain structures used uniquely for language acquisition. Rather, they hypothesize that second language acquisition, like other learning, requires the learner's attention and effort –whether or not the learner is fully aware of what is being attended to. Some information processing theories suggest that language, like other skilled activity, is first acquired through intentional learning of what is called 'declarative knowledge' and that, through practice, the declarative knowledge can become 'proceduralized' and, with further practice, it can become 'automatic' (De Keyser, 2003). Other theorists make a similar contrast between 'controlled' and 'automatic' processing (Segalowitz, 2003). The difference is that controlled processing is not necessarily intentional. Controlled processing occurs when a learner is accessing information that is new or rare or complex. Controlled processing requires mental effort and takes attention away from other controlled processes. For example, a language learner who appears relatively proficient in a conversation in a familiar topic may struggle to understand an academic lecture, because the effort and attention involved in interpreting the language itself interferes with the effort and attention needed to interpret the content. Automatic processing, on the other hand, occurs quickly and with little or no attention and effort. Indeed, it is argued that we cannot prevent automatic processing and have little awareness or memory of its occurrence. Thus, once language itself is largely automatic, attention can be focused on the content. The information processing model offers a useful explanation as to why learners in the initial phases of learning seem to put so much effort into understanding and producing language.

Thus, according to the information processing model, learning occurs when, through repeated practice, declarative knowledge becomes automatic. In addition to practice, it is also hypothesized that a process referred to as 'restructuring' may result in learners appearing to have made quite sudden changes in their interlanguage systems rather than gradually increasing the speed with which they use constructions that were already present. Restructuring is a cognitive process in which previously acquired information that has been somehow stored in separate categories is integrated and this integration expands the learner's competence (McLaughlin, 1990; McLaughlin and Heredia,

1996). Sometimes the restructuring can lead learners to make errors that had not previously been present. For example, when a learner comes to understand that English question forms require inversion, there might be a period in which embedded questions (*Do you know what the children are doing?*) would be produced with inversion as well (**Do you know what are the children doing?*).

Some researchers working within information processing models of SLA have argued that nothing is learned without 'noticing'. That is, in order for some feature of language to be acquired, it is not enough for the learner to be exposed to it through comprehensible input. The learner must actually notice what it is in that input that makes the meaning. This idea has raised a considerable amount of interest in the context of instructed second language learning (Schmidt, 1990, 2001). The next section (2.1.2), dealing with the hypotheses in SLA which account for how languages are learnt, further develops it as the Noticing Hypothesis.

The implicit/explicit dichotomy is also one of the central issues in the cognitive view of SLA. The underlying question here is whether adults can learn a language fully through the same implicit learning mechanisms used by the child in learning a first language. According to De Keyser (2003), 'implicit learning' can be defined as learning without awareness of what is being learned as opposed to a more explicit process whereby there is no lack of consciousness of the structure being learned.

It is important, furthermore, to distinguish implicit learning from two concepts it is often confused with in the second language literature: inductive learning and implicit memory. Inductive learning (going from the particular to the general, from examples to rules) and implicit learning (learning without awareness) are two orthogonal concepts (see table 1 below). Via traditional rule teaching, learning is both deductive and explicit. When students are encouraged to find rules for themselves by studying examples in a text, learning is inductive and explicit. When children acquire linguistic competence of their native language without thinking about its structure, their learning is inductive and implicit. The combination of deductive and implicit is less obvious, but the concept of parameter setting in Universal Grammar could be seen as an example; supposedly learners derive a number of characteristics of the language being learned from the setting of the parameter, and this clearly happens without awareness.

Table 1. The inductive/deductive and implicit/explicit dimensions

	Deductive	Inductive
Explicit	Traditional teaching	Rule discovery
Implicit	Using parameters	Learning L1 from input

Source: Implicit and Explicit Learning (De Keyser, 2003: 314)

In the same vein, implicit memory and implicit learning are in principle independent concepts. Even though implicitly aquired knowledge tends to remain implicit, and explicitly acquired knowledge tends to remain explicit, explicitly learned knowledge can become implicit in the sense that learners can lose awareness of its structure over time, and learners can become aware of the structure of implicit knowledge when attempting to access it, for example for applying it to a new context or for conveying it verbally to somebody else. (De Keyser, 2003).

Finally, we can not finish a section devoted to cognitive psychology in language acquisition without briefly summarising Skill Acquisition Theories. This is what the next section is about.

Skill Acquisition Theories

Skill Acquisition Theories of language acquisition draw on the distinction between declarative and procedural knowledge (Anderson, 1983) or between controlled and automatic processes (McLaughlin, just presented, 1987). That is, they are based on the view that language learning is characterized by a progression from an initial declarative knowledge stage involving controlled processing, to a final procedural stage where knowledge is automatic. Skills are learnt as a result of "practice". Practice, however, needs to be skill-related. So the development of skill in listening requires practice in processing input while the development of speaking requires practice in oral production (DeKeyser and Sokalski, 1996). According to this view, procedural knowledge is uni-directional; that is, automatization of one skill, such as listening, does not directly assist automatization of a different skill, such as speaking. However, automatization of one skill may have an indirect effect on a different skill by improving and strengthening declarative knowledge which is bi-directional (i.e. can be utilized in the development of different skills). Van Patten and Cadierno disagreed with this view and purported that input practice alone can lead to

improved output, although not viceversa (Van Patten and Cadierno, 1993).

According to skill-acquisition theorists, L2 learners achieve proceduralization through extensive practice in using the L2. However, "practice" is a relatively crude concept, especially when applied to language learning. What exactly does it entail? The traditional view is that practice involves the process of repeatedly and deliberately attempting to produce some specific target feature. It was this view that led to the use of the mechanical drills found in the audilingual and oral-situational methods of language teaching (Richards and Rodgers, 2001). What was missing from this view, according to DeKeyser (1998), was recognition of the importance of practice directed at "behaviour" rather than at "structures". Ellis (1988) showed that practice is often not effective in enabling learners to use new structures autonomously. This is because practising a structure in a mechanical way reifies the structure by decontextualizing it and thus does not affect long-term memory or lead to any change in behaviour. To change behaviour (i.e. develop automatic processes) it is necessary to provide practice of the actual behaviour itself. In the case of language learning, "behaviour" must entail attempts to communicate. Thus, for practice to work for the development of the speaking skill it must involve learners producing the target structure in the context of communicative activities.

According to this view, then, communicative practice serves as a device for proceduralizing knowledge of linguistic structures that have been first presented declaratively. Instruction that incorporates such practice can be seen as an attempt to intervene directly in the process by which declarative knowledge is proceduralized. DeKeyser (1998) drew on Anderson's skill-learning theory to argue for such an intervention:

> …proceduralization is achieved by engaging in the target behaviour –or procedure- while temporarily leaning on declarative crutches…
> Repeated behaviours of this kind allow the restructuring of declarative knowledge in ways that make it easier to proceduralize and allow the combination of co-occurring elements into larger chunks that reduce the working memory load (DeKeyser, 1998: 49).

Johnson (1988, 1996) also drew on skill-learning theory to justify practice. He emphasized the importance of feedback in the learning process, suggesting that the instructional sequence is best seen as one of "learn → perform → learn" rather than the traditional sequence of "learn → perform". During (or perhaps after) the "perform" stage learners must have the opportunity to receive feedback. This feedback, Johnson suggested, should consist of "mistake correction" (i.e. negative evidence about the misuse of features that the learners already have knowledge of but cannot yet use automatically). Johnson emphasized that for feedback to be effective learners "need to see for themselves what has gone wrong in the operating conditions under which they went wrong" (1988: 93). He suggested that this can probably be best achieved by means of extrinsic feedback (i.e. feedback from an outside source) that shows the learner what is wrong by modelling the correct form while they are attempting to communicate.

Skill acquisition theories of the kind promoted by DeKeyser and Johnson underlie mainstream accounts of how to teach grammar. Ur (1996), for example, proposed a sequence of practice activities designed to lead a learner from "accuracy" (i.e. performance based on declarative knowledge) to "fluency" (i.e. performance based on procedural knowledge). This sequence involves "controlled drills", "meaningful drills", "guided meaningful practice", "structure-based free sentence composition", "structure-based discourse composition" and "free discourse". It should be noted, however, that such a sequence finds a place for mechanical as well as communicative practice, seeing the former as a way of preparing for the latter, and, as such, does not conform with DeKeyser's and Johnson's views about the need to ensure that the practice involves "behaviour" in "real operating conditions".

There can be little doubt that language learning, in part at least, does involve skill-learning in the sense that practice aids the process by which L2 knowledge is automatized. However, skill-acquisition theories are problematic in two related aspects. First, they provide no explanation for the orders and sequences of acquisition. As Mitchell and Myles (1998) commented "the route followed by L2 learners is not convincingly explained by such approaches" (p. 99). Second, it is difficult to accept that the acquisition of all L2 features begins with declarative knowledge. This implies a role for metalinguistic awareness in L2 acquisition that far exceeds that sketched out in this section.

Four different periods regarding developments in the field of Linguistics applied to the acquisition of languages have been presented. One different line of enquiry in those years was the analysis of individual differences and their impact on LA to which a whole section is devoted. Before that, however, what follows now is an overview of four major FLA/ SLA hypotheses which tried to further account for how languages are learnt in the decades subsequent to the periods just described.

2.1.2 Four Key Hypotheses

In this section four key hypotheses related to the process of LA are presented: the Input Hypothesis, the Interaction Hypothesis, the Output Hypothesis, and the Noticing Hypothesis. None of these hypotheses alone purport to account for FLA. Yet taken as pieces of a model in construction they should provide a rough sketch of the features which come into play in the process of acquiring languages. The role of input and context of acquisition, the focus of this dissertation, is also central to all of them.

2.1.2.1 The Input Hypothesis

Advanced by Stephen Krashen since the 1980s (Krashen, 1982, 1985, 1998), the basic claim of the input hypothesis states that acquisition will take place automatically if learners receive 'comprehensible input'. We progress along the natural order of acquisition by understanding input that contains structures at our next 'stage' of language competence –structures that are slightly beyond our current level of competence. (We move from i, our current level, to $i + 1$, the next level along the natural order, by understanding input containing $i + 1$;…). Therefore, input is the essential environmental ingredient for LA to take place.

Krashen's hypothesis, within his monitor theory of FLA, proscribes traditional instruction devices (grammar teaching, linguistic grading, error correction, etc.) due to the so-called "non-interface" concerning any potential relationship between learned and acquired knowledge. This author claims that knowledge resulting of consciously learned language is distinct in memorial representation from unconsciously acquired language, that only the latter type of knowledge can be deployed in spontaneous language use, and, furthermore, that there can be no interaction between these two independent knowledge systems

(i.e., the so-called *learning / acquisition* distinction, where the former implies a conscious effort and the latter a naturalistic, to some degree, unconscious process). The non-interface position states that *learned* knowledge can never become *acquired* knowledge.

Krashen's views have had a notable impact on FLA and also on language pedagogy. Although intuitively attractive, however, Krashen's model was extremely criticised as lacking theoretical or empirical foundations. In fact, it was criticised to such an extent that the dichotomy acquisition/learning is generally dispensed with and linguists tend to use the terms as synonyms, as is the case in the present study. It was also argued that incomprehensible input was also vital to the process as it triggered learners' awareness of gaps in their knowledge (Gass, 1997). Long looked more carefully at the notion of comprehensibility and how it had to be negotiated in NSs / NNSs conversations (Long, 1983) and came to the conclusion that while the input hypothesis might be open to criticism one thing is uncontroversial: that without comprehensible input there is no learning, so from that perspective the hypothesis holds. This author's views are further presented in the following section.

2.1.2.2 The Interaction Hypothesis

In the early 1980s, as was already presented above (see 2.1.1.3), Long first advanced the argument that in order to understand the nature and usefulness of input for FLA more fully, greater attention had to be paid to the interactions in which learners were engaged (Long 1983). Like Krashen, Long viewed comprehensible input as a source of acquisition. However, he differed from Krashen in that he emphasized one particular way of obtaining comprehensible input: negotiation of meaning. Long argued that these interactions did not have to be seen simply as a one-directional source of target language input, feeding into the learner's presumed internal acquisition device. Instead, when learners engaged with their interlocutors in negotiations around meaning, the nature of the input might be qualitatively changed. That is, the more the input was queried, recycled and paraphrased, to increase its comprehensibility, the greater its potential usefulness as input, because it should become increasingly well-targeted to the particular developmental needs of the individual learner. "Learning does not happen outside performance; it occurs *in* performance" (Swain and Lapkin, 1998: 321). This view has become known as the Interaction Hypothesis (Long 1983, 1996). This author proposed that

environmental contributions to acquisition are mediated by selective attention and the learner's developing L2 processing capacity, and that these resources are brought together most usefully, although not exclusively, during negotiation for meaning.

He also stated that *negotiation for meaning*, and especially negotiation work that triggers *interactional* adjustments by the NS or more competent interlocutor, facilitates acquisition because it connects input, internal learner capacities, particularly selective attention, and output in productive ways. (Long, 1996: 451).

In addition, he proposed a role for different kinds of feedback. According to him, negative feedback obtained during negotiation work or elsewhere may be facilitative of L2 development, at least for vocabulary, morphology, and language-specific syntax, and essential for learning certain specifiable L1-L2 contrasts. (Long, 1996: 414). He defined negative feedback as input that provides "direct or indirect evidence of what is ungrammatical" (Long, 1996: 413).

The Interaction Hypothesis thus leans towards a social constructivist model of language lacking in FI contexts. According to this model, there are a variety of possible classroom interaction patterns between the Student(s) – S(s) and the Teacher(s) – T: S(s)↔S(s), S(s)→T, T→S(s) etc which have generated a number of studies (for example Allwright, 1984; Ellis, Tamaka and Yamasaki, 1994; Hall and Verplaetse, 2000). The analysis of interaction between T and S(s) has led to a renewed interest in corrective feedback as a means of negotiating meaning (and learning language) (Lyster, 2002; Lyster and Ranta, 1997; Mackey, 2006; Mohan and Beckett, 2003; Oliver and Mackey, 2003). These studies have mostly been small scale classroom based studies and as a consequence findings have tended to be mixed – different methods, assumptions and questions are almost bound to produce different results (Oliver and Mackey, 2003: 520), although, in Mackey's words, they do tend to be "interesting, complex and positive" (2006: 405). A rough consensus around the linguistic benefits of negative evidence seems to have been reached, as is further developed below (see table 2 about FoF).

Identifying the specific linguistic and / or learning gains of (negotiated) interaction is made difficult by the fact that the interest primarily lies in *incidental* rather than *intentional* learning (Ellis, 1999a: 4). Incidental

leaning covers a wealth of possibilities and can be extremely difficult to pinpoint.

> Incidental learning is unintentional or unplanned learning that results from other activities. It occurs often in the workplace, during the use of computers, and in the process of completing tasks. Incidental learning occurs in many ways, including the following: through observation, repetition, social interaction, and problem solving; from implicit meanings in the classroom or workplace policies or expectations; by watching or talking to colleagues or experts about tasks; and from being forced to accept or adapt to situations (Kerka, 2000:1).

Gass, Mackey and Pica recognise that "the precise role of interaction in actual development and internalisation of L2 knowledge has continued to challenge researchers" (1998: 299) although, as is the case with input and output, it is probably safe to say that there is a general consensus that it is a necessary component of LA and that corrective feedback seems to lead to higher levels of linguistic development.

2.1.2.3 The Output Hypothesis

A second challenge to Krashen was put forward by Merrill Swain. Her work with immersion students experiencing content-based second language French instruction in Canadian schools led her to question the claim that comprehensible second language input is sufficient to ensure all-round interlanguage development. Whereas Krashen had seen no role for speaking in L2 acquisition, Swain considered learner output an important mechanism in the acquisition process. This author advanced a set of claims about the relationship between language use and language learning, the so-called Output Hypothesis (Swain, 1985). The immersion students studied by Swain and her colleagues were exposed to French-medium instruction for extended periods of time, and achieved comprehension abilities in French as a second language that were close to native speaker level. However their productive ability lagged behind, something which Swain attributed to the fact that their classroom use of French mostly involved reading and listening to second language input, without corresponding expectations that they themselves would speak or write in French at a high level. Swain argued that students could often succeed comprehending second language texts, while only partly at processing them, that is, concentrating on semantic processing. In her view, only second

language production (i.e. output) really forces learners to undertake complete grammatical processing, and thus drives forward most effectively the development of second language syntax and morphology.

The Output Hypothesis goes beyond the idea of output as practice and addresses it as a learning tool (Swain, 1993). Swain (1995: 128) proposes three functions for output:

1. the notice / triggering function which, through a process akin to consciousness-raising, allows learners to become aware of gaps in their interlanguage
2. the hypothesis-testing function, which allows learners to try out language and see if it works
3. the metalinguistic function, which allows learners not just to try out their hypotheses but also to discuss them and reflect upon them

The Output Hypothesis has generated significant quantities of research, much of it examining the minutiae of the three points above (eg. De Bot, 1996; Izumi, 2003; Izumi, Bigelow, Fujiwara, and Fearnow, 1999; Shehadeh, 2003; Swain and Lapkin, 1982, 1995, 1998; Swain, 2005). Naturally the three functions above are not always guaranteed; rather they are seen as expressing an optimal scenario. Thus, the idea of output, like input, being necessary to language learning seems to be a working hypothesis, and a fairly sound one.

2.1.2.4 The Noticing Hypothesis

Although these previous hypotheses are presented in linear order, they have not evolved in sequence; to a certain extent (apart, perhaps from the original version of the input hypotheses) their relationship might be seen as symbiotic, feeding into and off each other and often grappling with the same concerns. One of the primary concerns which they all share relates to the role of consciousness within SLA. The idea of *noticing* being essential to language learning was first put forward by Schmidt, who posited three possible models of learning: subliminal, incidental and implicit (1990). He rejected the first, accepted the second and was uneasy about the third; and, in justifying his decision, conceived the Noticing Hypothesis, which, stated in general terms, is as follows:

> SLA is largely driven by what learners pay attention to and notice in target language input and what they understand the significance of noticed input to be (Schmidt, 2001).

In specifying the Noticing Hypothesis beyond its general formulation, Schmidt has claimed that learners *must* pay attention to what he terms "surface elements" in order to acquire them. More specifically, he states that:

> the order of attention and noticing are elements of the surface structure of utterances in the input – instances of language, rather than any abstract rules or principles of which such instances may be exemplars (Schmidt, 1990: 5).

Noticing structural regularities, forming hypotheses, and making comparisons is a level beyond. Precisely what these "surface" elements of language input are is, as yet, little understood. However, Schmidt is clear about how these elements should *not* be construed:

> Noticing is therefore used here in a restricted sense, as a technical term roughly equivalent to "apperception" (Gass *et al.*, 1998), to Tomlin and Villa's (1994) "detection within selective attention"… My intention is to separate "noticing" from metalinguistic awareness as clearly as possible (Schmidt, 1995: 5).

The key point is that metalinguistic awareness and noticing are to be considered different mental processes.

Though the Noticing Hypothesis is made difficult to evaluate due to its conceptual nature, again there is a commonsense consensus that noticing things (on some level of consciousness) must help. This begs the practical question: if learners need to notice, there must be ways in which the noticing can be *pushed*. This has given rise to the idea of enhanced input – a kind of third way between explicit teaching and implicit learning which seeks to provide opportunities for noticing. Input can be enhanced by various means:

1. textual or typographical enhancement refers to visual stimulus in printed materials by underlining, italics, colours, etc (Sharwood-Smith, 1993). Research has found, however, that while such enhancement does promote noticing, it does not necessarily

promote learning (Izumi, 2002; Jourdenais, Ota, Stauffer, Boyson and Doughty, 1995; Leow, 2001).

2. proactive verbal enhancement implies the teacher choosing in advance to model a wider range of types of input, for example making an effort to not always use the same discourse markers or request formulae (Moore, 2007). The supposition is that learners will acquire something from the teacher's models by a process akin to osmosis.

3. reactive verbal enhancement, known as Focus on Form. This is the area of input enhancement which has probably received the most attention. It:

> consists of an occasional shift of attention to linguistic code features –by the teacher and/or one or more students – triggered by perceived problems with comprehension or production (Long and Robinson, 1998: 23).

> There has been a wealth of research into Focus on Form enhancement although, once again, findings are mixed as to the benefits for language development (see section 3.1 and see Ellis, 1999b; Doughty and Williams, 1998; Long and Robinson, 1998) and the debate continues (Davies, 2006; Loewen, 2004; Zyzik and Polio, 2008).

The four key hypotheses outlined above serve more than one purpose. On the one hand, they serve to pinpoint four of the commonly agreed basic requirements of language learning theory: input, output, interaction and noticing (or some form of conscious involvement). On the other hand, even in the rapid and arguably superficial treatment above, it becomes clear that fine-tuning these hypotheses is stymied in the face of learner and situational variables, an issue dealt with further below (see 2.1.4).

The preceding two sections in this chapter have dealt with questions of general import to the study of second language acquisition, in an endeavour to review the current state of thinking around LA. In the following section, however, the focus is on classroom formal instruction contexts of LA as the relevant ones in the study presented here. Furthermore, the distinction between SLA and FLA will be presented.

2.1.3 Instructed Foreign Language Acquisition. Focus on Form

While there may be many similarities in the ways that learners learn a FL inside an outside a classroom, there are also differences with regard to the nature of the input they are exposed to, the kinds of interactions they participate in, and, crucially, the extent to which they attend to form as opposed to meaning. In naturalistic settings, learners will primarily treat the target language (from now on TL) as a tool for communicating. In classrooms settings, it is also possible that the TL will be viewed as a tool for communicating (as is the case in task-based teaching) but it is also likely that learners will approach the TL as an object to be studied and intentionally learnt. These differences have led researchers to specifically identify FI as a context of LA worth of attention.

In fact, the study of how acquisition takes place in a classroom context has entailed revisiting many general issues, some of them already presented in 1.1.1 and 1.1.2, such as the role of interaction in shaping learning, the difference between implicit and explicit knowledge, the role of corrective feedback, the significance of acquisition orders and sequences, and the role of individual learner differences. In the following lines, Focus on form as a feature of the input and language exposure learners receive will be presented. We adopt Ellis' perspective (Ellis, 2008) on this matter. After this, a presentation of the studies that have investigated the effects of Focus on Form Instruction (from now on FFI) on SLA is offered.

Before that, the distinction between FLA and SLA needs to be clarified. In the case of SLA the language plays an institutional and social role in the community (i.e. it functions as a recognized means of communication among members who speak some other language as their mother tongue). For example, English as a second language is learned in the United States, the United Kingdom, and countries in Africa such as Nigeria and Zambia. In contrast, FLA takes place in settings where the language plays no major role in the community and is primarily learnt only in the classroom. Examples of FLA are English learnt in France, Spain or Japan. This distnction is best treated as a sociolinguistic one rather than a psycholinguistic one. Somewhat confusingly, the term SLA is often used as a superordinate term to cover both types of learning. In this study, we take the standpoint of

using both, in an attempt not to forget their differences, albeit the many similarities.

2.1.3.1 Focus on form and acquisition

Ellis (2008) stressed the role of FoF in the process of LA in FI contexts. According to this author, one reason why learners fail to achieve high levels of competence in communicative classrooms may be their failure to attend to form. That is, because the activities they engage in are meaning-focused, they do not notice features such as past tense markings or unusual word order, or have many opportunities for "pushed output". Such an interpretation is compatible with the Noticing and Output Hypotheses summarised above. This has led researchers to investigate "focus-on-form" instruction. Long (1991) provided the following definition:

> …focus-on-form…overtly draws students' attention to linguistic elements as they arise incidentally in lessons whose overriding focus is on meaning or communication (Long, 1991: 45-6).

One macro-distinction that has figured strongly in recent form-focused instruction research is Focus-on forms versus Focus-on-form (Long, 1991; Doughty and Williams, 1998). Focus-on-forms refers to instruction that seeks to isolate linguistic forms in order to teach them one at a time as when language teaching is based on a structural syllabus. Focus-on-form, as defined above, involves "alternating in some principled way between a focus on meaning and a focus-on-form" (Long, 1991) and involves the use of tasks as opposed to exercises.

The definition of focus-on-form and focus-on-forms types of instruction is debated, however. For example, Doughty and Williams (1998) characterised focus-on-form as follows:

> …a focus-on-form entails a focus on the formal elements of language, whereas focus-on-forms is not limited to such a focus …the fundamental assumption of focus-on-form instruction is that meaning and use must be evident to the learner at the time that attention is drawn to linguistic apparatus needed to get the meaning across (Doughty and Williams,1998:4).

Further, Doughty and Williams have argued that both types can include explicit instruction and that the distinction between the two types constitutes a continuum rather than a dichotomy.

The theoretical rationale for focus-on-form is as follows:

1. To acquire the ability to use new linguistic forms communicatively, learners need the opportunity to engage in meaning-focused language use.
2. However, such opportunity will only guarantee full acquisition of the new linguistic forms if learners also have the opportunities to attend to form *while* engaged in meaning-focused language use. Long (1991) argued that only in this way can attention to form be made compatible with the immutable processes that characterize SLA and thereby overcome persistent developmental errors.
3. Given that learners have a limited capacity to process the new language and have difficulty in simultaneously attending to meaning and form they will prioritize meaning over form when performing a communicative activity (VanPatten, 1990).
4. For this reason, it is necessary to find ways of drawing the learners' attention to form *during* a communicative activity. As Doughty (2001) noted "the factor that distinguishes focus-on-form from other pedagogical approaches is the requirement that focus-on-form involves learners briefly and perhaps simultaneously attending to form, meaning and use during one cognitive event" (Doughty, 2001: 211)

Doughty and Williams (1998) offered a taxonomy of focus-on-form tasks and techniques based on whether they were unobtrusive (for example, recasts) or obtrusive (for example, consciousness-raising tasks).

Ellis, Basturkmen, and Loewen (2002) summarized the various options for inducing attention to form in the context of meaning-focused language use (see table 2 below).

Table 2. Options for focus-on-form

Options	Description
A Reactive focus-on-form	The teacher or another student responds to an error that a student makes in the context of a communicative activity.

1 Negotiation a Conversational b Didactic	The response to the error is triggered by a failure to understand what the student meant. It involves "negotiation of meaning". The response occurs even though no breakdown in communication has taken place; it constitutes a "time-out" from communicating. It involves "negotiation of form".
2 Feedback a Implicit feedback b Explicit feedback	The teacher or another student responds to a student's error without directly indicating an error has been made, e.g. by means of a recast. The teacher or another student responds to a student's error by directly indicating that an error has been made, e.g. by formally correcting the error or by using metalanguage to draw attention to it.
B Pre-emptive focus-on- form	The teacher or a student makes a linguistic form the topic of the discourse even though no error has been committed.
1 Student-initiated	A student asks a question about a linguistic form.
2 Teacher-initiated	The teacher gives advice about a linguistic form he/she thinks might be problematic or asks the students a question about the form.

Source: Doing focus on form. System 30 (Ellis, Basturkmen, and Loewen, 2002: 429-30)

2.1.3.2 Research on Form-focused instruction

There are now a large number of studies that have investigated the effects of FFI on SLA but clear conclusions are difficult to arrive at. Norris and Ortega (2000, 2003) noted problems with the research methods employed, relating to both the overall design of the studies (for example, no control group or no pre-test) and to the instruments used to measure learning outcomes (for example, the failure to demonstrate validity and reliability). Thus the following conclusions

regarding what research tells us about FFI must be viewed as programmatic (Ellis, 2008: 900).

a. FFI is effective in helping learners acquire an L2

The case for form-focused instruction is strengthening. FFI results in increased accuracy and accelerates progress through developmental sequences. It is effective in promoting both the learning of grammar and L2 pragmatics. Two important caveats, however, are that FFI may not be necessary for learning many of the features of an L2 (i.e. many of the features can be learnt naturally) and FFI may not ensure that learners achieve full target-language competence. In general, however, learners who receive FFI will learn faster and progress further than those who do not. Also, certain "marked" L2 features may only be acquirable with FFI.

b. The effects of FFI are not always positive

A number of studies have shown that FFI can sometimes have negative effects. FFI directed at features that are formally simple but functionally complex may result in their overuse. FFI directed at a feature similar to one previously taught and learnt may lead to confusion and loss of learning.

c. FFI facilitates natural language acquisition

This is a more contentious claim. As presented above (see 2.1.2.1), Krashen (1981) mantained that FFI assists "learning" but plays no role in "acquisition". However, there is now clear evidence that although FFI may be powerless to alter a developmental sequence it facilitates progress through it. Initially, this finding led to the claim that learners had to be at the stage immediately preceding the stage targeted by the instruction, but this may no longer be the case. Instruction directed to more than one stage ahead can enable learners to progress, although they will still follow the sequence. The available evidence suggests that FFI can work by facilitating the processes involved in natural L2 acquisition.

d. FFI also offers an alternative mode of learning

In addition to facilitating natural language acquisition, FFI can teach learners metalinguistic facts about the L2 and thereby contribute to their explicit knowledge. Explicit knowledge is of value in itself as it is available for formulating and monitoring utterances, especially in planned language use. There is also growing evidence that explicit knowledge developed through instruction can assist learners'

acquisition of implicit knowledge. That is, teaching "rules" (or assisting learners to discover rules for themselves) leads ultimately to improved accuracy in unplanned as well as planned language use.

e.   FFI can result in implicit as well as explicit L2 knowledge

This claim is, according to Ellis (2008), a key one. It addresses the criticism leveled at much FFI research by Doughty (2003) –namely, that the choice of measurement in many studies is biased towards explicit knowledge. It remains uncertain, however, how FFI results in implicit knowledge. One possibility is that it enables learners to convert explicit knowledge into implicit knowledge through practice in accordance with skill-building theory. Another possibility is that it serves to facilitate the processes involved in natural language acquisition in accordance with the claims that have been advanced in favor of focus-on-form instruction or input-processing instruction.

f.   There are constraints on the teachability of specific features

There are constraints on whether FFI works. Factors such as the degree of markedness, form-function transparency, and the nature of the processing operations involved determine how difficult different structures are to teach. Thus, it does not follow that FFI will always be effective. However, little is currently known about whether these constraints apply to explicit knowledge as well as implicit knowledge or whether there are different constraints that apply to the two types of knowledge.

g.   The effects of FFI may or may not be evident immediately and may or may not be durable

A number of studies have found that the effects of the FFI do not appear in the immediate post-test but do emerge some time later in a delayed post-test. Instruction raises learners' consciousness about a feature which is then attended to selectively in subsequent input, resulting in acquisition. Other studies have found that instruction can have an immediate effect but that this may not last (i.e. it disappears in a delayed post-test). An explanation of this phenomenon is that the instruction resulted only in explicit (declarative) knowledge which then atrophies because the learner was not developmentally "ready" to acquire it or because of no subsequent communicative exposure to it and lack of adequate amounts of practice.

h. Both focus-on-forms and focus-on-form instruction are effective

This claim is also contentious. Doughty (2003), for example, argued that "the completely decontextualized nature of explicit focus-on-forms… promotes a mode of learning that is arguably unrelated to L2 acquisition… in that the outcome is merely the accumulation of metalinguistic knowledge about language" (p.271). In other words, Doughty claimed that focus-on-forms only results in explicit knowledge. There is also growing evidence that focus-on-form instruction facilitates acquisition. However, it is not possible to claim that one kind of instruction is superior to the other. Ellis (2008) has argued that given the disagreements in the definitions of these two constructs and the fact that they are composites involving a number of distinct options it may not be possible to conduct a convincing comparison.

i. The type of instruction influences learning outcomes

What does research tells us about the relative effects of a number of different instructional options such as input-based, explicit as opposed to implicit, inductive versus deductive, error-inducing production practice, or corrective feedback? According to Ellis (2008), it is not easy to reach clear conclusions due to the fact that most of the studies did not investigate discrete options but rather combination of options, making it difficult to determine what aspect of the instruction was effective. There is clear evidence that input-based instruction can assist acquisition although it may be premature to claim that this is more effective than production-based instruction. There is clearer evidence (Norris and Ortega, 2000) that explicit instruction (especially when combined with practice activities) is more effective than implicit instruction (i.e. practice activities alone). Both inductive and deductive explicit instruction appears to work with no clear evidence in favour of either. Inducing errors in order to correct also appears effective. Arguably, some of the best research has examined corrective feedback, providing an accumulation of evidence to suggest that explicit types of feedback (for example, metalinguistic explanation) are more effective than implicit types (for example, recasts) and that output-prompting types (for example, elicitation) are more effective than input-provided (for example, recasts) at least for features that learners have already partially acquired (Sanz, 2003; Lyster, 2004; Sheen, 2006). Ultimately, however, trying to establish which type of instruction is most effective may be a mistaken enterprise as it may depend on contextual and individual learner factors.

j.   Individual difference factors mediate the effects of FFI

Researchers have been interested in examining whether, to what extent, and in what ways individual factors such as learning style, language aptitude, memory, anxiety, age, motivation, and learner attitudes interact with different types of FFI. There is clear evidence that they do, although, again, it may be premature to offer any conclusions. Learners' language analytical abilities influence their capacity to process instruction, especially when this is of the more formal, explicit kind. Working memory and affective factors such as motivation and anxiety have been shown to have an impact when the instruction is of the more implicit kind.

The research we have considered in the previous section summarised the effects of FFI on learners in general. The underlying assumption is that it is possible to identify whether instruction or what types of instruction work best for all learners. However, it would seem likely that learners differ in the kind of instruction they are best equipped to benefit from. In other words, individual factors may mediate the effects of instruction (as explained in 2.1.3.2. "j"). Therefore, in the next section these individual differences are examined in detail.

2.1.4 Individual Differences

No account of L2 acquisition is complete without due consideration of individual differences (ID) in learners. There is a real plethora of individual learner variables which researchers have identified as influencing learners outcomes and these are presented in this section. Dörnyei (2005) defines them broadly as "enduring personal characteristics that are assumed to apply to everybody and on which people differ by degree" (p.4).

Table 3 lists the main variables mentioned in three surveys: Skehan (1989), Robinson (2002), and Dörnyei (2005). The three authors include language aptitude, motivation, and anxiety among others. These, then, can be considered "core variables" as appear on the table in bold.

Table 3. Factors listed as influencing individual learner differences in language learning in three surveys

Skehan (1989)	Robinson (2002)	Dörnyei (2005)

1. Language aptitude
2. Motivation
3. Language learning strategies
4. Cognitive and affective factors
  - extroversion / introversion
  - risk-taking
  - intelligence
  - field independence
  - **anxiety**
	1. Intelligence
	2. **Motivation**
	3. Anxiety
	4. Language Aptitude
	5. Working Memory
	6. Age
		1. Personality
		2. Language Aptitude
		3. Motivation
		4. Learning and cognitive styles
		5. Language learning strategies
		6. Other learner characteristics
		  - **anxiety**
		  - creativity
		  - willingness to communicate
		  - self-esteem
		  - learner beliefs

Source: The Study of Second Language Acquisition (Ellis, 2008: 644)

According to Ellis (2008), what has been lacking in this area of SLA, however, is a framework for examining these factors. This is because the factors overlap in vague and indeterminate ways. Ellis' attempt to impose some order on this field of enquiry (see Ellis, 2004) is shown in table 4 below. This author distinguished factors according to whether they constitute (I) "abilities" (i.e. cognitive capabilities for language learning that are relatively immutable); (2) "propensities" (i.e. cognitive and affective qualities involving preparedness or orientation to language learning that can change as a result of experience); (3) learner cognitions about L2 learning" (i.e. conceptions and beliefs about L2 learning), and (4) "learner actions" (i.e. learning strategies).

Table 4. Factors responsible for individual differences in L2 learning

Category	Factors
A Abilities	1 Intelligence 2 Working memory 3 Language aptitude
B Propensities	1 Learning style 2 Motivation 3 Anxiety 4 Personality 5 Willingness to communicate
C Learner cognitions about L2 learning	Learner beliefs
D Learner actions	Learning strategies

Source: Individual differences in language learning (Ellis, 2004: 530)

Dörnyei's 2005 classification is presented and discussed in detail in order to identify ID's relevant to our study of CLIL acquisition. This

author's view has been chosen for two different reasons. It follows the structure of the seminal book by Peter Skehan (1989) *Individual Differences in Second Language Learning* and it provides a consistent and comprehensive review of the most up-to-date studies in this field. In section 2.2.1.4 (Individual variables associated with bilingualism) reference to some other IDs will be made as they are relevant to the present study of CLIL acquisition. This is, for example, the case of some variables associated with bilingualism and trilingualism, such as language use, and balance.

According to Dörnyei (2005), the core variables in ID research comprise personality, language aptitude, motivation, learning/cognitive styles, and learning strategies.

   a.   Personality
It can be defined as those characteristics of a person that "account for consistent patterns of feeling, thinking and behaving" (Dörnyei, 2005). The relationship between personality variables and L2 learning is not yet clear. There is some evidence to show that extroverted learners are advantaged in the development of the kind of language associated with basic interpersonal communication skills. Extroverted learners may also be more likely to participate actively in oral communication. This broad sub-domain of personality is not further explored here, partly because it is so extensive, and also because progress in this area has been slow, in terms of both methodology and systematic patterns of results. A review of the area can be found in Dewaele and Furnham (2000).

   b.   Aptitude
People differ in the extent to which they possess a natural ability for learning an L2. This ability, known as *language aptitude*, is believed to be in part related to general intelligence but also to be in part distinct. Language aptitude involves both an underlying language learning capacity and a capacity to handle decontextualized language. Both quantitative and qualitative differences in language aptitude have been found. These relate to the development of both linguistic and communicative L2 abilities.

Early work by John Carroll (1991) led to the identification of a number of components of language aptitude. These are:

- Phonemic coding ability – capacity to code an unfamiliar sound so that it can be retained over more than a few seconds and subsequently retrieved or recognized.
- Grammatical sensitivity – capacity to identify the grammatical functions that words fulfil in sentences.
- Inductive language learning ability – capacity to extract syntactic and morphological patterns from a given corpus of language material and to extrapolate from such patterns to create new sentences.
- Associative memory – capacity to form associative bonds in memory between L1 and L2 vocabulary items.

Skehan (1998) proposed that different components of aptitude could be related to stages of information processing. Phonemic coding ability can be related to input processing; language analytic ability (grammatical sensitivity, inductive language learning) can be related to central processing; and memory-as-retrieval can be related to output and fluency. Such a set of linkages shows how aptitude, at a fairly general level, is consistent with a cognitive view of SLA.

For many years, aptitude has been isolated from the wider area of foreign language learning and acquisition. It has been perceived as effective as a predictor, but undemocratic with respect to learners, out of date conceptually, and of little explanatory value. Research over the last years has indicated that this judgment is unwarranted. Aptitude may well be a central construct when there is a focus in form in SLA, precisely the condition many SLA researchers now call for. If we accept that there is a critical period for second language learning and that totally meaning-based acquisition is a hazardous undertaking, then aptitude may well represent a constellation of individual differences which bear upon the effectiveness with which learners are able to focus on form when the conditions for doing so are operative (Skehan, 1991). According to Dörnyei (2005), age is precisely one ongoing issue in language aptitude research and therefore a special mention is devoted to this factor. In addition, this factor, together with gender, are very relevant for the present dissertation because their possible impact on differential gains in favour of a FI or a CLIL learning context is going to be measured.

b.1 Age
Children generally enjoy an advantage over adults in L2 learning because of their age, particularly in pronunciation. However, this will only become evident after substantial exposure to the L2. In the short

term, adults may learn faster. The evidence relating to the existence of a critical period for L2 acquisition, after which full competence is not possible, is mixed, with no definite conclusion possible. Children and adults manifest similar processes of learning.

If we consider research carried out in the particular setting where our CLIL study is conducted, the Barcelona Age Factor Project (Muñoz, 2006) investigated the effects of age of onset in an instructional setting. This project examined the acquisition of English by classroom learners in Catalonia, comparing those who began their study at the age of 8, 11, and 14 and controlling for exposure to English outside the classroom. Data from a battery of tests providing measures of both implicit and explicit types of knowledge were collected on three occasions –after 200 hours of instruction, 416 hours, and 726 hours. The main finding was that the older learners progressed faster than the younger learners. The younger learners did not catch up over time. These results, which contrast with those of previous studies, must be interpreted in the light of the distinctions established above between SLA and FLA, and the input and output hypotheses: the number of hours available for learning in this FI foreign language context was insufficient to enable them to do so. However, there was evidence that age had a differential effect on the acquisition of different aspects of the L2. Thus, the advantage for the older learners was strong and durable on measures of grammar and least evident in the case of measures of speech perception, listening comprehension, and oral fluency. In the latter measures, no statistically significant differences between the young and older starters were evident on the final measurement. Overall, the research supports the conclusion that learners who start learning an L2 in adolescence or as adults learn more rapidly than those who start in childhood.

However, to fully understand the results of this research it is useful to distinguish the effects of age on the rate of acquisition in terms of the distinction between implicit and explicit learning (DeKeyser, 2000). The greater cognitive development of older learners is advantageous where explicit learning is concerned. In contrast, they do not necessarily outperform early-starters in the long-term where implicit learning is involved.

These results are preceded by those of a wide range of studies investigating the effects of age on L2 acquisition in a SLA context summarised by Singleton (1989, 2005):

> Concerning the hypothesis that those who begin learning a second language in childhood in the long run generally achieve higher levels of proficiency than those who begin in later life, one can say that there is some good supportive evidence and that there is no actual counter evidence (Singleton, 1989: 137).

As for the Critical Period Hypothesis (CPH), another Spanish based project, the Basque Age Factor Project (García-Mayo and García-Lecumberri, 2003) claims that there is a fixed span of years during which language learning can take place naturally and effortlessly, and after which it is not possible to be completely successful. There is, however, no clear consensus on when the "window of opportunity" for language learning ends. Singleton (2005), in a survey of the literature that has addressed this issue, reports claims ranging from near birth to late adolescence. Also, it has become clear that, if there is a critical period, this varies depending on the level of language under examination with the end point coming earlier for pronunciation than for grammar.

c. Motivation

This factor involves the attitudes and affective states that influence the degree of effort that learners make to learn an L2. Various kinds of motivation have been identified: instrumental, integrative, resultative, and intrinsic. (Clément and Gardner, 2001; Gardner, 2007; Dörnyei, 2005).

Instrumental motivation

Learners may make efforts to learn an L2 for some functional reason – to pass an examination, to find a better job, or to obtain a place at university.

Integrative motivation

Some learners may choose to learn a particular L2 because they are interested in the people and culture represented by the target-language group (the motivation that many English speaking Canadians have for learning French, for example).

Resultative motivation

It is also possible that motivation is the result of learning. That is, learners who experience success in learning may become more motivated to learn. In a context like Canada, success in learning French may intensify English-speaking learners' liking for French culture.

Intrinsic motivation

It is possible that many learners do not hold distinct attitudes, positive or negative, towards the target-language group. Such is probably the case with many foreign language learners. It does not follow, however, that such learners are unmotivated. They may find the kinds of learning tasks they are asked to do intrinsically motivating. According to this view, motivation involves the arousal and maintenance of curiosity and can ebb and flow as a result of such factors as learners' particular interests and the extent to which they feel personally involved in learning activities.

These four types of motivation should be seen as complementary rather than as distinct and oppositional.

Strength of motivation serves as a powerful predictor of L2 achievement, but may itself be the result of previous learning experiences. Learners with either integrative or instrumental motivation, or a mixture of both, will manifest greater effort and perseverance in learning. Other internal sources of motivation, such as self-confidence, may be more important than either type of motivation in some contexts. Motivation can also take the form of intrinsic interest in specific learning activities and, as such, may be more easily influenced by teachers than goal-directed motivation.

Finally, Dörnyei (2005) proposed a process-oriented conceptualization of motivation. As so it is defined as the dynamically changing cumulative arousal in a person that initiates, directs, coordinates, amplifies, terminates, and evaluates the cognitive and motor processes whereby initial wishes and desires are selected, prioritized, operationalized, and (successfully or unsuccessfully) acted out.

d. Cognitive and learning style

According to Dönyei, confusing labels, ineffective measurement instruments, lack of distinction between style and other constructs (e.g., personality) make the research on learning style and cognitive style a "quagmire" (Dönyei, 2005: 120). The autor tries to at least clarify the difference between learning style and cognitive style using colors as an analogy:

> cognitive styles can be seen as equivalents of the colors proper, whereas learning styles are the manifesations of the colors in the real world" (Dönyei, 2005:160).

Cognitive style can be defined as a predisposition to process information in a characteristic manner. In contrast, learning style can be defined as a typical preference for approaching learning in general. The former, in other words, is more restricted to information-processing preferences, while the latter embraces all aspects of learning.

The major interpretation of cognitive style has been through studies of the constructs of field independence and field dependence (FI/FD from now on). FI are seen as more likely to analyze information into its component parts, and to distinguish the essential from the inessential. FD, in contrast, are more likely to deal with information structures as wholes, or "gestalts". At a personal level, FI are portrayed as aloof, preferring to find solutions to problems for themselves. FD, in contrast, are sociable and work well in groups. Each of these putative preferences could have advantages in language learning: the former should link with a capacity to analyze linguistic material, and perhaps learn systematically; the latter to engage in communicative language use, and to "talk to learn".

e. Learning strategies

This concept reflects the learner's active contribution to enhancing the effectiveness of his or her own learning. In other words, the students' own active and creative participation in the learning process through the application of individualized learning techniques. According to Skehan (1991), learning strategies can be classified in the following way: cognitive strategies (repetition, summarizing, using images…); metacognitive strategies (analyzing, monitoring, evaluating, planning, organizing one's own learning process…); social strategies (initiating interaction with native speakers, cooperating with peers…); and affective strategies (taking control of the emotional conditions and experiences that shape one's subjective involvement in learning), what Dörnyei refers to as the 'process oriented' conceptualisation of motivation.

There have been various attempts to discover which strategies are important for L2 acquisition. One way is to investigate how "good language learners" try to learn. One of the best-known and frequently cited study is Naiman *et al.* (1978/1996). A main finding of such studies is that successful language learners pay attention to both form and meaning. Good language learners are also very active (i.e. they use strategies for taking charge of their own learning), show awareness of

the learning process and their own personal learning styles and, above all, are flexible and appropriate in their use of learning strategies. They seem to be especially adept at using metacognitive strategies.

Other studies have sought to relate learners' reported use of different strategies to their L2 proficiency to try to find out which strategies are important for language development. Such studies have shown, not surprisingly, that successful learners use more strategies than unsuccessful learners. They have also shown that different strategies are related to different aspects of L2 learning. Thus, strategies that involve formal practice (for example, rehearsing a new word) contribute to the development of linguistic competence whereas strategies involving functional practice (for example, seeking out native speakers to talk to) aid the development of communicative skills. Successful learners may also call on different strategies at different stages of their development. However, there is the problem with how to interpret this research. Does strategy use result in learning or does learning increase learners' ability to employ more strategies? For the moment being, there are more questions than answers in the literature.

Finally, apart from the core variables which Dörnyei highlights, this author also includes a discussion of less-researched learner characteristics also regarded as important in the ID research picture. Among them, I will make reference to 'affective state' since it includes 'anxiety', a factor highlighted not only by Dörnyei (2005) but also by Skehan (1989) and Robinson (2002). Since the effect of gender is also measured in the present dissertation, in the following lines mention to this individual variable is also made. These suposed less prominent variables open a large window onto future research into this field that may also significantly impact second language acquisition studies.

    f.   Affective state

Affective factors relate to the learner's emotional state and attitude towards the target language. Learners' affective states vary dynamically and have a significant impact on their ability to learn. The affective aspect that has received the most attention in SLA is anxiety. Anxiety arising out of poor performance, communication apprehension, tests, and fear of negative evaluation is likely to have a debilitating effect on L2 learning, but it can also have a facilitative effect. How anxiety affects learning will depend on its strength and the situational context. There has been an attempt to examine experimentally how language anxiety affects language processing (MacIntyre and Gardner, 1994).

Ethnographic studies based on rich descriptions of learners' reactions to their learning situations have also begun to appear (Spielmann and Radnofsky, 2001). The research has addressed three key issues: the sources of language anxiety, the nature of the relationship between language anxiety and language learning, and how anxiety affects learning.

It can be concluded that anxiety (its presence or absence) is best seen not as a necessary condition of successful L2 learning, but rather as a factor that contributes in differing degrees in different learners, depending in part on other individual difference factors such as their motivational orientation and personality. Research has attempted to relate language anxiety to the developmental aspects of language learning and to a model of language processing. These are both positive aspects of anxiety research.

One study we would like to mention here is related to the effects of language anxiety on learners of Spanish as a FL, Juan-Garau and Marcos-Llinàs study (submitted). The study was conducted with 134 American college learners of Spanish. Anxiety was measured through the Foreign Language Classroom Anxiety Scale (FLCAS) consisting of 33 statements that assess communication apprehension, test anxiety, and fear of negative evaluation in the FL classroom. Results show an interrelation between language anxiety and academic achievement. Thus, students with high levels of anxiety exhibit lower academic achievement in comparison to students with low levels of language anxiety. This finding seems to confirm previous studies on anxiety and language learning. Nevertheless, results also showed a medium level of language anxiety among most participants with no negative effect on academic achievement.

    g. Gender

As mentioned before, since the effect of gender on language competence will also be analyzed in this dissertation, attention will be paid to the gender variable in the following lines. Research studies, which have widely covered gender issues during the last three decades, suggest that gender plays a significant role in foreign language performance, as there seems to be a female oriented culture that spreads the idea that learning foreign languages is a feminine terrain (Kobayashi, 2002) and, subsequently, male students feel less confident and obtain worse scores (Onwuegbuzie, Bailey and Daley, 2001; Oxford, 1993). Research studies undertaken in different contexts show

that women are more inclined to study foreign languages and second languages and they usually outperform their male counterparts (Pavlenko and Piller, 2008; Sunderland, 2000). However, CLIL programmes seem to help blur these gender-based differences. A study of foreign language learners by Schmidt, Boraie and Kassagby (1996) may help to explain this. They concluded that females were better foreign language learners because they were more intrinsically motivated, whereas males expressed more ex- trinsically motivated reasons. This is why CLIL programmes may help balance out gender differences, as male students might feel more motivated to learn both the language and the subject matter, enabling them to obtain higher scores in the subject concerned (Lasagabaster, 2008).

To conclude, it is important to note that two views regarding the nature of the relationship between IDs and L2 learning are possible. According to the aggregate view, the one represented by the Naiman *et al*. study (1978/1996), success is the result of the accumulative effect of facilitative IDs. For example, a learner with low anxiety, high overall language aptitude, an inclination to be analytic, a strong integrative motivation, and an outgoing personality could be considered likely to succeed. In contrast, according to the alternative view, there are many ways to achieve success and it is not possible to draw up a single profile of the successful language learner. This view of IDs has important implications for language instruction because it acknowledges the need to take account of learner-instruction interactions (i.e. to recognize that different learners can achieve the same level of success if the instruction matches their own preferred approach to learning).

By way of a final cautionary word, research into individual differences would do well to heed Larsen-Freeman's (1997) warning:

> Progress in understanding L2 acquisition will not be made simply by identifying more and more variables that are thought to influence language learners (Larsen-Freeman, 1997: 156).

As has been noted at the beginning of this section, individual factors may mediate the effects of instruction. An increasing number of studies have examined the inter-relationships between ID variables, formal instruction, and learning outcomes. There is now empirical evidence that cognitive and affective IDs do mediate the effects of FFI. For example, Bialystok (1985) noted commonsensically that there needs to

be a "minimal congruity" between the learner's preferred way of learning and the type of instruction for L2 acquisition to proceed efficiently.

Regarding the studies investigating the effect of affective factors in LA, it is tempting to hypothesize that these may be more important when attention to form occurs in instruction that is primarily meaning-oriented. Two studies support this hypothesis. Takahashi (2005) found that intrinsic motivation led to higher achievement in input-based instruction, while Sheen (2006) reported that learners with low language classroom anxiety were able to benefit from corrective feedback in the form of recasts to a greater extent than those with high anxiety.

Other studies investigated the effects of cognitive variables such as language aptitude (especially language analytical ability). DeKeyser (1993) found no interaction between language analytical ability and instruction (possibly because this study did not investigate specific grammatical targets) but three other studies (which targeted specific structures) did find a relationship. Robinson (1997) reported that learners with high analytical ability benefited in three of the conditions he investigated –implicit, rule-search, and instructed. Erlam (2005) found that language analytic ability was a factor when learners were taught by means of an inductive or structured input instruction. Sheen (2006) found that learners with high language analytical ability achieved higher scores on tests of English articles as a result of a treatment that included metalinguistic explanations but not as a result of the same treatment with recasts. Thus, in the case of Robinson and Sheen it would appear that language analytical ability is of benefit when the instruction encourages direct attention to form but not when learner's attention is primarily focused on meaning.

For the purpose of our study, research related to age (García-Mayo and García-Lecumberri, 2003, and Muñoz, 2006) has already been presented in this section. Furthermore, in 2.2.1.5 and 2.2.2.1 empirical evidence concerning individual variables related to bilingualism and trilingualism (language use, biliteracy and balance) is also summarised since the dissertation presented here deals with Catalan/Spanish learners of English as a third language in multilingual Catalonia.

In sum, in this section the core variables in ID research mainly following Dörnyei's (2005) view have been presented. These comprise

personality, language aptitude, motivation, learning/cognitive styles, and learning strategies. Following the thread of the core variables, I have elaborated on each of them individually. It is important to note that age has also been considered and has been included as a factor within language aptitude research. The same occurs with affective state, which, although suposed a less prominent variable, it is one of the three common factors highlighted both by Skehan (1989), Robinson (2002), and Dörnyei (2005) together with motivation and language aptitude. After the presentation of the core variables, the current state of research in this domain has been briefly presented.

To conclude, an assessment of ID research seems to offer a mixed picture. One can even delineate interesting connections and interactions among different ID variables which can give a better understanding of the complicated phenomenon of ID variables in L2 acquisition. For example, it is clear that foreign language aptitude and cognitive style have some degree of relationship. This connection is accounted for by the way each draws upon the common underlying factor of intelligence. In slight contrast, Skehan (1998) argues that the connection arises because within aptitude one can propose an analytic learner type and a memory-oriented learner type. This is related to, but not identical with the analytic-holistic contrast in the style literature. Skehan argues that if cognitive style is interpreted as not one continuum but two, this, combined with a style vs. predisposition interpretation, can acomodate, separately, both aptitudinal and style concepts. We can see another connection between learning style and learning strategies, in that style relates to consistency of strategy use accross contexts. More interesting, perhaps, is the potential connection between motivation and learning strategies: the effective use of learning strategies may be precisely the sort of behaviour that causes motivational levels to be sustained within the learning situation (Dörnyei, 2001). Their use may give encouragement to the learner, provide benchmarks for evaluation and progress, and enable motivational goal setting to be accomplished. If strategies are viewed in this way, they may re-emerge within a more elaborated theoretical framework.

2.1.5 Summary

So far, in this first part of chapter 2 an overview of research on FI has been presented. Four different chronological periods concerning four differentiated models of language acquisition have been explained: the

structuralist-behaviorist period, the Chomskyan period, the social-interactionist period, and the cognitive period with a section devoted to skill acquisition theories. After them, four key hypotheses related in our view to language acquisition have been introduced: the Input Hypothesis, the Interaction Hypothesis, the Output Hypothesis, and the Noticing Hypothesis. A section on focus on form and acquisition and the current state of research in this domain has also been presented. The section ends with a review on individual differences in SLA.

To conclude, we have seen that taken together, the fields of skill acquisition theory and second language acquisition theory have established a series of conditions for the acquisition of knowledge and development of skill. A relevant question for the study presented here is whether a CLIL context creates different learning conditions to those of FI. Purportedly, in contrast to FI, which limits exposure to the target language to what takes place in the classroom, CLIL offers more hours of both input and output production. In addition, communication is a central feature, hence interaction is fostered (Pérez-Vidal, 2007). Empirical evidence is beginning to show that this seems to be the case (see 2.3 below). As far as the cognitive side of learning processes, research is scarce.

What follows now, included in the second part of this chapter, is a review of current thinking about bilingual and trilingual contexts since the study presented here deals with Catalan/Spanish learners of English as a third language in multilingual Catalonia.

2.2 Bilingual and Trilingual Contexts.

This second part of chapter two provides a focus on key issues at the intersection of multilingualism and second language acquisition. An overview of research on Bilingual Language Acquisition and Third Language Acquisition (BLA and TLA henceforth) will be presented after an introduction of what bilingualism implies.

2.2.1 Bilingual Language Acquisition

In this section we introduce definitions and major typologies for classifying bilingualism. We then discuss theories on the cognitive effects of bilingualism. After that, the major individual variables associated with bilingualism are explained and, finally, an overview of research on BLA is presented.

2.2.1.1 Definitions of bilingualism

According to Butler and Hakuta (2004, p.114), bilinguals are often defined as individuals or groups of people who obtain the knowledge and use of more than one language. However, bilingualism is a complex psychological and socio-cultural linguistic behaviour and has multi-dimensional aspects. There is no agreed-upon definition of bilingualism among researchers. What do we mean by "knowing" two languages? As is often believed, bilinguals could be defined as individuals who have "native-like control of two languages" (Bloomfield, 1933, p.56). However, this strict view of bilingualism limits the number of individuals and groups that could be classified as bilingual, not to mention the fact that such a definition makes it difficult to operationalize "native-like fluencies".

On the other hand, many researchers employ a broader view of bilingualism and include in their defintion of bilinguals those individuals who have various degrees of proficiency in both languages (Macnamara, 1967; Hakuta, 1986; Valdés and Figueroa, 1994; Mohantly and Perregaux, 1997). Broader definitions of bilingualism have an advantage in that they incorporate the developmental processes of SLA into the scope of studies on bilingualism (Hakuta, 1986). Grosjean (1999) and Grosjean (2008), for instance, focuses on the daily use of two languages, and distinguishes bilinguals who use more than two languages in their daily life from "dormant bilinguals" who retain knowledge of different languages but no longer use them in daily life.

In this investigation, I adopt a broader notion of bilinguals which corresponds to the recent shift of focus among bilingual researchers away from the acquisition of formal rules of language and onto communicative skills (Mohantly and Perregaux, 1997). The present authors define bilinguals as individuals or groups of people who obtain communicative skills, with various degrees of proficiency, in oral and/or written forms, in order to interact with speakers of one or more languages in a given society. Accordingly bilingualism can be defined as psychological and social states of individuals or groups of people that results from interactions via language in which two or more linguistic codes (including dialects) are used for communication. Hammers and Blanc (2000) called individual bilingualism "bilinguality" and distinguish it from societal bilingualism. This is the first classification that is explained in the next part.

2.2.1.2 Typologies

Rather than attempting to provide a definition of bilingualism, most specialists prefer to work within the framework of a typology of bilingualism which allows for a clear delimitation of the particular area of investigation within a larger field. We are going to highlight the following typologies mainly following Baetens Beardsmore (1986) point of view: societal and individual bilingualism, natural and secondary bilingualism, balanced and dominant bilingualism, and additive and subtractive bilingualism.

Societal and individual bilingualism

According to Baetens Beadsmore (1986), in societal bilingualism the investigator is placing the accent primarily on understanding what linguistic forces are present in a community, their inter-relationships, the degree of connection between political, economic, social, educative and cultural forces and language. On the other hand, individual bilingualism focuses on the individual and this broad field attempts to classify bilinguals into different categories depending on linguistic, cognitive, developmental, and social dimensions.

For the student of bilingualism the societal aspects often form the background canvas which determines the relevance of his enquiry by clarifying the historical and social processes which lead to the existence of bilingual individuals. Even the microlinguistic case study of one bilingual speaker must normally be prefaced by a contextualization of the elements which brought about the presence of two or more languages in that one speaker, often in the form of a simple case history, but sometimes leading to a complex description of both the background and the ways the two languages form part of the person's everyday life (as we have done in the present study section 2.3.2.4).

Natural and secondary bilingualism

By natural bilingual, also known as primary bilingual, we understand someone who has picked up two languages by force of circumstances, either in the home as a child or by moving to a community where the speaker is obliged to work with more than one language, but where no systematic instruction in two languages has been provided.

On the other hand, secondary bilingualism describes the situation when a second language has been added to a first language via instruction. (Baetens Beardsmore, 1986: 8)

Balanced and dominant bilingualism

This distinction is based on the relationship between the proficiencies of the respective languages that bilinguals master (Peal and Lambert, 1962). Balanced bilinguals are those who acquire similar degrees of proficiency in both languages. On the other hand, dominant (or unbalanced) bilinguals are individuals whose proficiency in one language is higher than that in the other language

Additive and subtractive bilingualism

Lambert (1974) focused on how one's L2 affected the retention of one's L1. Bilinguals who can enhance their L2 without losing L1 proficiency have been referred to as additive bilinguals, whereas those whose L2 was acquired or learned at the expense of losing their L1 have been referred to in the literature as subtractive bilinguals. To be additive bilinguals, both of the languages learned by bilingual individuals must be valued in the society in which they reside.

In sum, bilingual individuals can be classified on the basis of different dimensions both at the individual and social levels, and thus can be classified into different types of bilinguals depending on which dimensions of their bilingual characteristics are the focus of attention.

2.2.1.3 Theories dealing with the cognitives effects of bilingualism

As will be explained more in detail in section 2.2.1.5, since the 1960s, research on the impact of bilingualism on cognition has associated bilingualism as the source of the cognitive differences observed when comparing bilingual and monolingual subjects. There are a series of authors that maintain that bilingualism is an important reason for these differences. One of them is Cummins with the Threshold Hypothesis (1976, 1979) and Interdependence Hypothesis (1981).

The Threshold Hypothesis (Cummins)

Cummins elaborated the Threshold Hypothesis (1976) in order to explain contradictory results found in studies on cognitive advantages (and disadvantages) of bilingualism. This hypothesis establishes a relationship between cognition and language proficiency. It attempts to describe the underlying mechanisms leading to individual differences in terms of positive and negative cognitive consequences among bilinguals. His original proposal stated that bilingual individuals can enjoy cognitive advantages if they attain "native-speaker competence" in both languages, therefore a balance is necessary. In other words, Cummins does not claim that only balanced bilinguals benefit from the bilingual experience, but that an upper threshold must be reached to observe cognitive benefits.

However, if they have not attained such competence in either of the languages, they may fall into a state of "semilingualism" and may not be able to avoid negative consequences in their cognitive and academic development. As Cummins admitted:

> The term semilingualism has no explanatory or predictive value but is rather a restatement of the equally ill-defined notion of 'limited proficiency in two languages' (Cummins, 2000: 104)

Whatever term is used and whatever "semilingualism" refers to, there are indeed individual differences in academic performance among L2 learners as well as monolingual students (see section 2.1.4). The key questions seem to be (1) to what extent such individual differences in academic performace among L2 learners can be attributed to their "language proficiency" as opposed to their ability to master academic content knowledge and skills ("achievement"); and/or (2) whether or not individual variations in academic performance are better explained by qualitatively different constructs (i.e. separately for monolingual students and L2 learners). These questions leads us to the following section (2.2.1.4) about individual variables associated with bilingualism (in section 2.1.4 an overview of the individual differences in SLA in general was presented).

The Interdependence Hypothesis (Cummins)

According to Cummins' Interdependence Hypothesis (1981), children learn to use language as a symbolic system in the process of acquiring

literacy skills in their first language, and that results in ability to classify, abstract, and generalise linguistic information in a way that can be transferred to subsequent language learning contexts.

This hypothesis, stating that academic proficiency in L1 and L2 are interdependent, provided a theoretical framework for understanding the mechanism of bilingual proficiency in academic contexts. This "common underlying proficiency", composed of both conceptual and procedural knowledge and skills, enables bilinguals to transfer academic skills from one language to another.

2.2.1.4 Individual variables associated with bilingualism

Individual bilinguals can be classified differently according to different dimensions, such as the relative relationships between L1 proficiency and L2 proficiency, the age of exposure to a given language, and the status of a particular language in a given society. However, such typologies can capture only a small subset of the many aspects of bilingualism. Bilingualism is indeed very dynamic and entails multi-dimensional, continuous variables. It is therefore very important to identify the factors that contribute to individual variation not only in language learning in general as we have seen in section 2.1.4 but also within bilinguals. The present section focuses on the two major variables associated with bilingualism (Baker, 1993): *language use* and *balance* (*age* is also another important factor within this field but we have already explained it in 2.1.4)

Language Use (frequency and context)

Baker (1993) points out the importance of incorporating language use and not just language knowledge into any study dealing with bilingualism. The difference between knowledge of the two languages and actual use of the two languages can be especially acute in Barcelona, the context of the present study, where the unequal status of Catalan vs. Spanish is more greatly felt. It is commonplace in Catalan sociolinguistics to refer to the difference between the language of the classroom (Catalan) and the language of recess (Spanish).

Biliteracy and bilingualism (balance)

Some scholars attribute the cognitive benefits of bilingualism to literacy in two languages (Bialystok, 2004; Cook, 1997; Cummins,

1981; Swain *et al.,* 1990). Swain and colleagues investigated the effect of L1 (a number of Romance and non-Romance languages) literacy on L3 (French) learning among 319 eight-graders for whom English was the L2 in Toronto. Results of this study suggest that the crucial factor in successful L3 acquisition is development of heritage language literacy skills, rather than exclusively oral skills. As has already been detailed in section 2.2.1.3, their conclusion supports Cummins' (1981) linguistic interdependence hypothesis, according to which children learn to use language as a symbolic system in the process of acquiring literacy skills in their first language, and that results in ability to classify, abstract, and generalise linguistic information in a way that can be transferred to subsequent language learning contexts.

2.2.1.5 Research on BLA: Background on the evidence regarding effects of bilingualism

The claim that bilingual children of many language backgrounds show academic or intellectual deficiencies was widespread through most of the 20th century (Saer, 1923; Jones and Stewart, 1951; Macnamara, 1967). The studies purporting to demonstrate such deficiencies typically showed a correlation between bilingual status and low scores on academic or intelligence tests. The causes of these apparent deficiencies demand evaluation. Might the bilingual child be hampered by the extra cognitive / linguistic burden imposed by multiple language learning? In the USA, the key facts that have been invoked to support this view have long been based upon the well-documented tendency oh Hispanic-American children to perform poorly on various tests of achievement when compared with monolingual children. While it is true that the average Hispanic child scores below the mean for the nation on academic tests, it is also true that the average Hispanic child in the United States is of lower socio-economic status than the average child as measured across the entire population. In studies comparing academic performance of Hispanic children of low socio-economic status with non-Hispanic children of similar socio-economic status, Hispanic children do not trail academically (Lambert, 1981; Peal & Lambert, 1962). These results suggest that poor academic performance could be the result of factors other than bilingualism: poverty is associated with low educational levels in parents, poor nutrition, domestic violence, a sense of diminished status and self-worth, and lower levels of linguistic stimulation than are available to children of higher socioeconomic status (August & Hakuta, 1997).

In fact, many bilingual children do well in school. A thorough analysis of the evidence on educational and linguistic outcomes for children educated in two languages yields a complex picture that suggests bilingual education is sometimes advantageous. Based in part on the results of the Canadian studies of bilingualism through French immersion in elementary school for children from English-speaking homes (Lambert and Tucker, 1972; Lapkin *et al.,* 1980; Swain and Lapkin, 1991), it appears that, for some children, successful adapatation to the needs of bilingualism may produce academic and social advantages in comparison with monolingual peers.

It is important to note that for the past 40 years there have been a spate of investigations suggesting that bilingual children and adults actually possess significant and consistent advantages over monolinguals on a variety of metalinguistic and / or cognitive tasks. More specifically, since the 1960s, research on the impact of bilingualism on cognition has associated bilingualism with positive effects on a number of internal variables, including intelligence (Peal & Lambert, 1962), metalinguistic awareness (Ben-Zeev, 1977; Bialystok, 1991), cognitive flexibility and processing mechanisms (McLaughlin & Nayak, 1989; Nation & McLaughlin, 1986; Nayak, Hansen, Krueger, & McLaughlin, 1990).

Concerning the most recent studies related to cognitive consequences of bilingualism, one of the most exciting developments comes from studies that show that a life as a bilingual confers a set of benefits to cognition within the realm of executive function. A now compelling body of literature shows that there are benefits of bilingualism on attentional control that extend from young bilingual children to young adult bilinguals and to elderly bilinguals (for example, Bialystok, 2005; Costa, Hernandez and Sebastián-Gallés, 2008). Most notably, these benefits are observed in simple cognitive tasks that do not explicitly involve language. The data on older bilinguals are particularly striking because bilingualism appears to provide a measure of protection against the normal effects of cognitive aging (for example, Bialystok, Craik, Klein and Viswanathan, 2004). Elderly bilinguals outperform their monolingual counterparts on tasks that require them to ignore irrelevant information or to resolve conflict in the face of stimulus-response incompatibility. The hypothesis is that a life spent negotiating cross-language competition fine tunes a set of cognitive skills that benefit the ability to select targeted information, regardless of whether the context is linguistic or not. Thus far, the available data are correlational. It will

remain to be seen in the next period how studies of language processing in bilinguals might be related to the observed cognitive consequences to provide a causal account of the way in which the resolution of cross-language competition might create these changes in cognitive performance.

All these findings encourage further evaluation of the relatively good performance of bilingual learners in academic domains. Such is the case of the subjects in the study presented here. Their performance analysed in the following sections can be categorised as the development of bilingual learners in a third target language, a topic developed in the next section.

2.2.2 Third Language Acquisition

In this section a special focus on research in third language acquisition (TLA henceforth) will be presented specially highlighting the effects of bilingualism on third language acquistion. After this, the main models in TLA research will be characterised.

The field of research on TLA represents a rather young discipline within linguistics which has, however, been gaining more and more interest over the last years (Clyne, 1997; Cenoz and Genesee, 1998; Hufeisen and Lindemann, 1998; Cenoz and Jessner, 2000; Dentler, Hufeisen and Lindemann, 2000). Although the number of studies on the acquisition of a third or a fourth (or more) languages is still very limited, this research area has already established itself as a field of its own by emphasising the differences between TLA, BLA and SLA as well as pointing out that other aspects of learning a third language have to be seen as similar to SLA.

2.2.2.1 Effects of bilingualism on third language acquisition

It seems to be widely known that under certain circumstances life with two or more languages can lead to advantages, not only with regard to language knowledge but also in terms of cognitive and sociopragmatic development. Cummins' Threshold Hypothesis (Cummins, 1976) presented above (2.2.1.3) states that a certain level of proficiency in both languages has to be attained in order to profit from the cognitive advantages which are related to a heightened level of metalinguistic awareness, creative or divergent thinking, communicative sensitivity and further language learning.

In some studies bilingualism has already been proved to be beneficial in the learning process of the third language. This belief corroborates informal observations (Larsen-Freeman, 1983 in Zobl, 1992) and different studies that compare bilinguals' and monolinguals' acquisition of a foreign language (Cenoz & Valencia, 1994, with data from the Basque Country; Safont-Jordà, 2005, as well as Sanz, 2000, 2008, with Catalan data, and Swain *et al*. 1982, 1990, with Canadian data).

Similarly, Ringbom (1987), one of the first scholars to show interest in TLA, reports of the advantages of Swedish speaking Finns over monolingual Finnish students when acquiring English in Finland. Ringbom concludes that, among other factors, like language typology or linguistic experience resulting from bilingualism, which play an important role in this particular language learning context, the high degree of automaticity as that found in experts learners is influential.

These studies show that in a sociolinguistic situation that promotes additive bilingualism, like the one for students in immersion programmes in Canada, the Basque Country and Catalonia, bilingualism appears to exert a positive effect on third language learning.

In Sanz's (2000) study, bilingualism (biliteracy, to be precise) also results in more efficient language learning. Her comparison of achievement in English by students in Spain following instruction through Catalan and Spanish in Catalonia yielded evidence in favour of bilingualism and bilingual education as positive contributors to foreign language learning. Sanz (2005) carried out a subsequent study prompted by a need to explore the role of individual variables in L3 learning and specially the role of biliteracy in enhancing L3 learning. The author focused on a group of high-school junior bilinguals in Catalan and Spanish learning English as a foreign language to identify and explain those factors associated with bilingualism as well as general factors identified in the SLA literature that may predict successful acquisition of a third language (L3). The results obtained lead Sanz to conclude that motivation and exposure are the most important variables for success in L3 acquisition by bilinguals. Once these variables are controlled, however, a higher level of biliteracy –the ability to read and write in Catalan and Spanish- is associated with a higher level of English proficiency. The author interprets these results as a confirmation of the existence of cognitive benefits of bilingualism for cognition as it pertains to subsequent language learning, and

specifically of the existence of a Threshold Level (Cummins, 1976) associated with equal ability to either read or write in both Catalan and Spanish. These results also agree with those in Muñoz (2000), Lasagabaster (2000), Sagasta (2003), and Roquet (2005) and take them further by showing that it is not overall L1 and L2 proficiency but biliteracy that contributes to cognitive benefits resulting in enhanced ability to learn languages.

In Lasagabaster's study (2000), Spanish Basque bilingual students also outperformed Spanish monolinguals in the acquisition of English as an L3. Furthermore, Lasagabaster applied Cummins's Threshold Model to these trilingual children in the Basque Country and found support for the relationship between the varying levels of proficiency in the three languages and the stages in cognitive development.

Cenoz (2003) found a tendency towards mixed results in studies on the effects of bilingualism on further language learning which she related to the diversity of the studies concerning the specific aspects of proficiency, methodology used and the testing context. Summarizing, she pointed out that the majority of studies on general proficiency indicated a positive effect of bilingualism on TLA and that this effect was linked to metalinguistic awareness, language learning strategies and communicative ability, in particular in the case of typologically close languages

Roquet (2005) analysed the effects of the "regular reading language" variable (Catalan, Spanish, or both) on the learners' written ability in a third language, namely English. At the same time, the variable upper threshold of bilingual proficiency in Catalan / Spanish was analysed to check its correlation with their written ability in English. The study was carried out while also looking at the impact of some of the individual variables that have proven to be essential in previous research (age, motivation, L1, or L3 degree of exposure). For that purpose, 58 bilingual Catalan / Spanish students aged 13 to 17 were analysed with a questionnaire gauging personal linguistic profile, reading and writing regular languages, attitudes, beliefs, and motivation. They were also administered three written tests in Catalan, Spanish, and English by means of a composition. They were analysed following an adapted version of the Celaya, Pérez-Vidal, & Torras (2001) matrix, which profiles lexical and syntactic complexity, fluency and accuracy features.

The results in Roquet demonstrated again the importance of reading regularly in both the L1 and L2 in order to attain balanced bilingualism with an upper threshold of bilingual proficiency (regardless of the rest of the individual variables). At the same time, results showed statistically significant superiority in complexity, fluency and accuracy features of written English tests of learners who read regularly in both Catalan and Spanish.

Safont Jordà (2005) also compared Valencian / Spanish bilinguals and monolinguals learning an L3 and this study showed that the main effect for experience was prevalent, constant, and uncomplicated by possible interactions. Instruction did not level the field: bilinguals retained their advantage even after instruction focused on specific pragmatic functions.

Sanz (2005) concluded, like Thomas' classic study (1988), that the key variable for the successful acquisition of a third language in bilinguals was the ability to read and write in two languages. Thomas showed that English-Spanish bilingual students in USA performed significantly better than their monolingual peers when learning French in the classroom. Ongoing research by Ellen Bialystok and her team at York University also shows a strong relationship between biliteracy and bilingualism.

2.2.2.2 Main models in TLA research

Most of the models used in research on multilingualism have been developed from those presented for BLA research and also SLA. Research on TLA is supposed to bridge the gaps between the areas of study of SLA and bilingualism. Models from both fields have been taken into consideration in the study of TLA: Bilingual and multilingual production models (De Bot, 1992, 2004; Clyne, 2003), The activation/inhibition model (Green, 1998), The language mode hypothesis (Grosjean, 1998, 2001), The multilingual processing model (Meissner, 2004). Furthermore, a new dynamic systems theory approach has been adopted in the study of language acquisition in three, four or more languages. The dynamics of the processes involved in individual progression and regression and the complex interdependences between the factors involved in the language acquisition process are focused on in a dynamic view of language acquisitions and multilingualism, the dynamic model of multilingualism (DMM henceforth), a model created by Herdina and

Jessner (2002) as a specific realisation of such a new approach to the study of multilingualism.

According to these authors, language acquisition can be seen as inherently dynamic, something that exhibits change and flux, and is characterised by motion. In contrast to fossilised languages, living languages are in a continuous motion. They adapt to the social contexts in which they are used and they move with time, changing chronologically. A specific instance of chronological or diachronic change occurs on the individual level.

Language change in the individual results from adjusting one's language system(s) to one's communicative needs. If we look at the bilingual as an integrated whole, we can watch how changes in the language environment, and therefore in language needs, affect her/his linguistic competence in the one or the other language, not in her/his linguistic competence in general. Speakers may move from monolingualism to bilingualism, from bilingualism to trilingualism, that is different systems (LS1, LS2, LS3, etc.) are transitionally commanded by the same individual. According to the communicative needs, the native speaker has transitional command of different language systems over a period of time, resulting, for instance, in monolingualism, bilingualism, trilingualism, etc.

In the case of multilinguals we are frequently confronted with the phenomenon of language loss, language deterioration and/or attrition, a phenomenon frequently observed by sociolinguists. Generally, language loss has been investigated in terms of language death under pressure of a competing or dominant language, that is, in a fundamentally bilingual situation. This might create the inaccurate impression that language loss or unlearning only takes place in linguistic situations where languages are in sociolinguistic and therefore frequently also psycholinguistic competition. We may rather assume that language loss can take place in normal and healthy monolingual speakers as well, that is in the form of intrapersonal (systematic) variation (Mehotcheva, 2010).

Furthermore, this psycholinguistic model of multilingualism is learner-oriented and tries to explain individual learner differences in language acquisition. Herdina and Jessner are interested in various factors affecting (the learner's) language performance, for instance, attitude and motivation, anxiety, language aptitude. The approach taken in the

DMM thus focuses on the dynamic systems of language learners. This view implies that language itself is in constant flow, and so are the language systems in a multilingual, depending on the various factors involved in the language acquisition process.

DMM provides the necessary conceptual psycholinguistic framework for modelling multilingual proficiency. It describes the language systems of a bi- or multilingual reacting differently to identical input in different situations, that is, different languages commanded by the same speaker which are viewed as separate systems (LS1, LS2, LS3, etc.) exhibiting different properties. This model, taking the wholistic view of bilingualism into account, stresses the fact that an adequate description of multilingualism must comprise not only transfer phenomena including codeswitching, language mixing, language attrition, but also the positive cognitive consequences of multilingualism (e.g. enhanced metalinguistic and metacognitive abilities, divergent thinking), which become apparent if certain social and cognitive conditions are met. Multilingual proficiency is, therefore, to be considered as consisting of dynamically interacting linguistic subsystems which themselves do not necessarily represent any kind of constant but are subject to variation.

2.2.3 Summary

In this second part of chapter two an overview of research on Bilingual Language Acquisition and Third Language Acquisition (BLA and TLA henceforth) has been presented after an introduction of what Bilingualism implies.

Firstly, definitions and major typologies for classifying bilingualism have been presented, namely "Societal and individual bilingualism", "Natural and secondary bilingualism", "Balanced and dominant bilingualism", and "Additive and subtractive bilingualism". After this, theories about the cognitive effects of bilingualism have been discussed, specifically Cummins' Threshold and Interdependence Hypotheses. Next, the major individual variables associated with bilingualism have been explained: Language Use (frequency and context), and Biliteracy and bilingualism (balance). Finally, a summary about research related to the good performance of bilingual learners in academic contexts has suggested that bilingual education is often advantageous in comparison with a monolingual one. This is very interesting for the present dissertation since the subjects presented are

all bilinguals who develop in a third language, the topic treated in the last part of this section: Third Language Acquisition.

In this last part, TLA, the effects of bilingualism on TLA have been highlighted and the main conclusion to be drawn after the representation of some studies is that in a sociolinguistic situation that promotes biligualism, like the one for the students in the present study, Catalonia, bilingualism appears to exert a positive effect on third language learning (see 2.2.2.1). Finally, the section ends with the main models in TLA research specially focusing on the dynamic model of multilingualism (Herdina and Jessner, 2002). This model explains that multilingual proficiency is to be considered as consisting of dynamically interacting linguistic subsystems which themselves do not necessarily represent any kind of constant but are subject to variation.

After a review of Formal Instruction Contexts (2.1), and Bilingual and Trilingual Contexts (2.2), what follows now is a presentation of CLIL Contexts (2.3).

2.3 CLIL Contexts

This section is intended as a presentation of studies analysing language acquisition in a CLIL context. CLIL is an integrative new approach to education which encompasses both a concern for language and for content instruction, the former being the focus of the present study. The present section is divided into two parts. The first one includes an introduction with a definition of the different terms that have been used to refer to this approach, and it describes the context where CLIL was born within the background of the European Strategy towards Multilingualism. It also includes the CLIL dimensions, modalities, features, benefits and challenges that such a new educational paradigm entails. All these factors related to CLIL are going to be used to identify the type of programme and subjects which is the focus of this dissertation, in the chapter devoted to the presentation of the study (Chapter 3). The second part reviews research on CLIL in Canada and Europe with a special mention in the case of Spain and Catalonia.

2.3.1 CLIL Contexts of Acquisition

Being educated in a language other than one's mother tongue has been around for over 5000 years. The ability of a social group to impose its language, coupled with the desire to promote linguistic, cultural and

national conformity, have often resulted in situations where a second or a foreign language was used as the medium of instruction. The European context is no exception. Variegated forms of bilingual education date back over several millennia (Glyn Lewis, 1976). Luxembourg has had bilingual education since 1843 (Davis, 1994) and trilingual education since 1913 (Berg, 1993). According to the Eurydice Report (2006), Malta introduced bilingual education in the 19th century, Bulgaria in the 1950s, Estonia in the 1960s as well as the first French-German bilingual schools in Germany in 1969 and so on. Moreover, a multilingual European School network was started in 1953 (Swan, 1996) to take account of linguistic diversity for children of mobile European civil servants.

Within a bilingual approach to education, learners are dealing with the content matter in any chosen area/s of the curriculum by means of a language which they would otherwise have only learnt in conventional classes. This has been undoubtedly the general approach adopted by international schools, European schools and private schools with an international profile for many years. They have operated on the principle that education in general and language teaching/learning in particular are powerfully enriched by such a multilingual approach.

Different terms have been used to refer to the programmes related to this approach: immersion, content-based language teaching/learning, teaching content through a foreign language, bilingual education, language enriched learning, or more recently plurilingual education, and content and language integrated learning. According to Pérez-Vidal and Campanale (2006: 18), each name refers to a slightly different situation, four of them deserving an explanation:

Immersion is the term used in research conducted in Canada (presented in 2.3.2.1) aimed at analysing the effects of the different programmes set up in Quebec for the following reasons. In Quebec, English speaking children needed to learn French, the official language in Quebec. This prompted a group of parents to lobby their school board for improvements to the teaching of French. After consultation with McGill university scholars in bilingualism, they proposed what was going to be labelled an immersion program to the board. From the first day of school in Kindergarten children would be instructed entirely in French and taught to read in this language. Only later in Grade 2 -7 years of age- would they start with their L1, English, until little by

little, by Grade 6 -12 years of age- half the curriculum was taught in French and half in English.

Content-based language teaching/learning was the term identifying an approach to language teaching successfully disseminated in the US around the 80s[2].

Bilingual education was the term used both in the American continent and in Europe during the 80s and 90s and specifically in the White Paper entitled Teaching and learning: Towards the learning society (1995) (Christian & Genesee, 2001). The importance attached to effective language teaching and learning if every citizen was to benefit from the single market in a united Europe led to a series of debates within the European Union (EU). At the end of these, all member states agreed on a resolution seeking improvements in the quality and diversity of language teaching/learning throughout the EU (Grenfell, 2002: 23). Higher levels of language proficiency and greater cultural awareness were seen as the key to the construction of Europe. Bilingual education was thus recommended. It is also used as a generic term referring to any type of programme where a target language is used to deal with subject content.

2.3.1.1 European policies

More recently the umbrella term Content and Language Integrated Learning and its acronym CLIL have been adopted by various European researchers and agencies as a generic term for such programmes. Chronologically, the CLIL approach appeared with the experience and wealth of research carried out in Canada and the United States in perspective. The construct is characterized by an extension of the English formula 'language across the curriculum', which revolves around the idea of integration of mother tongue education in school, into 'languages across the curriculum' (Wolff, 1998: 26) to include educational, and social goals in the approach, as explained further

---

[2]This approach sought to offer alternatives to the classroom practices used with learners from immigrant communities (Snow, Met & Genesee, 1989; Brinton, Snow & Wesche, 1989). It was seen as an offspring of communicative language teaching, developed in the context of new perspectives in linguistics, sociolinguistics and philosophy (Wolff, 1998:21) a radical one in which 'meaningful syllabae' and 'purposeful realistic interaction' were achieved by introducing subject content in language courses within a linguistic framework (sic. Coyle, 1998: 62).

along this text (see 2.3.1.2). Hence, the term encompasses many different forms of learning in which a language carries a special role alongside the learning of any specific subject or content. It is this specifically European model which this dissertation presents.

Content and Language Integrated Learning (CLIL) was the term proposed for primary and secondary education, coexisting with Content-Based-Instruction (CBI), whereas the term Integrated Content and Language (ICL) was also used to refer to the same multilingual modality yet in tertiary education (Wilkinson, B., Zegers, V. and Van Leeuwen, Ch., 2006). Three different translations into Spanish have been proposed so far, in alphabetical order: AICLE (Adquisición Integrada de Contenidos y Lengua Extranjera), EICLE (Enfoque Integrado de Contenidos y Lenguas Extranjeras), and Semi-inmersión (Pérez-Vidal, 2008; Escobar-Urmeneta, 2007).

Still following Pérez-Vidal and Campanale (2006), it has been contended that the growth of multilingual education in Europe is perhaps the result of economic factors, the impetus of the Bologna declaration requiring European transparency and harmonization of qualifications irrespective of the language of instruction, and mobility policies (see also Van Leeuwen, 2006: 26). What is undeniable is that the concern for languages as an asset to be preserved and promoted within the construction of Europe has led European institutions to herald change in the domain of education in general and languages in particular.

The European Commission's (2003) Action Plan stating recommendations of multilingual policies had been preceded by the White Paper on education and Learning (1995), whereby citizens should be functionally proficient in their mother tongue and two other European Languages. The paper framed a whole strategy towards multilingualism which included the factors "interdisciplinarity" alongside "intensity of exposure", as two key factors in the strategy and strongly recommended policies to member states. It had the following objectives:

- Diversification of modern languages learned as early in life as possible.
- Promotion of democratic access to knowledge for all European citizens, whatever their first languages.

- Priority given to the development of capacities in very young children by:
  - The early acquisition of a second language beginning at nursery school.
  - The intensive and transdisciplinary teaching of this second language at primary school.
  - The intensive and transdisciplinary teaching of at least a third modern language at secondary school.

In addition, the Council of Europe presented the community with two instruments for the promotion of the highest levels of language education: the Common European Framework of Reference and the European Language Portfolio (2001)[3].

Thereafter, a specific European interdisciplinary approach represented by CLIL began to take shape, it was then handed over to the community to be refined and served as the background to a number of varied experiments being carried out in different European countries not unfamiliar with the wealth of good practice and research carried out in Canada and the United States. Accordingly the CLIL concept emerged, under the auspices of the European Council, but also within a large number of Commission funded projects. Indeed, the BILD and the DIESeLL projects, the Thematic network in Bilingual Education, the CLIL Compendium, the ALPME project, the TIE-CLIL, the TICCAL, the CDI-BIT, and two networks, the CLIL Cascade and MOLAN, which represent just a small sample of the work undertaken in the last decade, and which has resulted in a construct, a European construct, around which shared knowledge and expertise is already an asset to European language acquisition and language pedagogy research (Pérez-Vidal, 1997, 2009; Marsh and Marsland, 1999; Ruiz de Zarobe and Lasagabaster, 2010 on Spain; Van de Craen and Wolff, 1997 for a complete European technical report, Nikula, 1997; Nikula and Marsh, 1997 on terminological considerations, and finally Dalton-

---

[3] The Framework is addressed to the educational community. It establishes a common base for curriculum and materials design which particularly affects the assessment of linguistic competence in language learners. The Portfolio has a twofold objective. Firstly it is something to be given to each individual citizen. It has an *informative* objective (particularly in the first part, the *Passport*) of promoting his or her mobility on the basis of transparency in European accreditation systems (see Cassany et.al. 2004a, 2004b; Pérez-Vidal, 2007a for a thorough description in connection with the Spanish portfolio). Secondly, it can help the young to use those parts in it which promote individual autonomy, self-assessment and awareness of other cultures, with a *formative* objective.

Puffer, 2007 among others). Central to the development of this context are several associations and institutions such as the European Council in Strasbourg, the Centre for Moden Languages in Graz, or the Association ELL, the actual springboard for may of the projects just listed[4].

Many programmes around Europe have been set up throughout these years of industrious innovation and practice, at primary, secondary and tertiary level, following recommendations from European institutions. Belgium, Finland, Germany, Italy and Holland, France and Spain, Hungary, and other newly arrived countries to the European Union have witnessed the spread of initiatives in this direction taken either by schools, parents, associations or the administration (see Baetens Beadsmore, 1993; Klapper, 1996 as an example of a German report; Baetens Beadsmore, 1998; Van de Craen and Pérez-Vidal, 2001, for primary and secondary level accounts; the survey in Marsh *et al*., 2001, and Pérez-Vidal, 2001b, as a survey with a Europe wide scope; Grenfell, 2002; Van Leeuwen *et al.,* 2003; Wilkinson, Zegers and Leeuwen, 2006, for tertiary level; Van de Craen *et al.,* 2008, for a report on Belgium)[5]. As Pérez-Vidal (2001c) stresses, in the Catalan and Basque autonomous communities of Spain, those programmes benefited from the accumulated experience of the highly successful CLIL programmes for the normalization of their official languages, Catalan and Basque which, in a background of societal bilingualism, since the 1980s have set up different educational models geared to ensure additive bilingualism (Artigal, 1993 for a general presentation of programmes; Sierra, 1994 on the Basque model, Vila-Moreno, 2008 for the Catalan area programmes).

It is worth noting that the recent priority attached to languages is not only the consequence of a political decision taken by member states to act against poor standards in mainstream education, but rather the effect of the recommendations of higher level council of Europe's policies. European multilingualism as a goal, and CLIL as an instrument towards the goal, have been formally proposed by European decision bodies and have been developed mostly within European funded projects. In

---

[4] The ELC was launched in 1997 as an association formed by some 120 universities seeking to develop the multilingual construct in Europe and works towards its implementation (www.celelc.org)

[5] See Lauridsen, K. and Toudic, D. (2008) for a presentation of major current trends in multilingualism in Europe under the leadership of Professor Wolfgang Mackiewicz.

practice, many teaching schemes may have been developed totally unaware of such written briefs, as a result of good teaching practices encompassing new local needs and new instruments, yet it is useful to realise that there is a common background to our private endeavours.

Summarising Pérez-Vidal and Campanale (2006) and more recently Pérez-Vidal (2009), the momentum gathered in the past decades by CLIL may be explained not only by its timely emergence and multifaceted nature, as described before, but also by the sound rationale on which it is based, with Communicative Language Teaching (CLT) as its backbone. It would be useful to try and remember the facts. At the end of the 80s, the *Threshold Levels* were published by a group of prominent applied linguists working under the mandate of the Council of Europe. They contained the structures, functions and vocabulary for subsequent levels of language competence. Such seminal work was taking place at the same time as was developed a functional-semantic perspective to syllabus design and the communicative methodology for classroom teaching.

The view proposed by Pérez-Vidal (2009) seems extremely appropriate to understand that

> CLIL is essentially the natural development of communicative approaches, updated with the incorporation of the effects of several recent phenomena taking place either in the field of teaching and learning, as the new views on autonomous learning, or in the field of technological development, such as the impact of the widespread use of internet and ICT, or in social relations, such as the world's growing internationalisation and student's mobility schemes. Indeed, if we try to explain the emergence of CLIL, three factors seem to have taken the 'communicative paradigm' further ahead and are at the origin of the Content and Language Integrated perspective to education and linguistic pedagogy:
> - the European Union's political project and increasing globalisation and mobility, which the Union's policies themselves promote;
> - the new pedagogical insights such as the key role played by individual differences and in particular attitude and motivation in the development of autonomy in language learning and;
> - technological progress. (Pérez-Vidal, 2009: 6)

2.3.1.2 Rationale of a CLIL approach

The rationale for the approach rests on three main ideas corresponding to a socio-cultural dimension, an educational-content dimension, and a linguistic dimension (Pérez-Vidal and Campanale, 2006: 20-26, and subsequently Pérez-Vidal, 2008).

A Content and Language Integrated educational approach holds intrinsic values which can help us improve the standards of education in Europe. First, it can foster European citizenship, and pave the way towards the integration of citizens in multilingual and multicultural Europe. Second, it renews content teaching by emphasizing the importance of learners' responsibility in their own learning, and encouraging creative, critical thinking. Third, it can improve current foreign language programmes in at least two ways: by increasing exposure time to a target language and by adding meaningfulness to the actual teaching. Let us consider each of these dimensions in turn.

a. The European socio-cultural dimension: building European citizenship

The socio-cultural dimension to a CLIL approach to education is related to the role played by languages, multilingualism and multiculturalism in the construction of an ethos in the construction of Europe. It rests upon the concepts of Citizen Europe, the Europe of Knowledge and the Europe of Languages. Since the first two are beyond the scope of this research we are going to focus in more detail on the concept of the Europe of Languages.

Citizen Europe

The construction of Europe will only be possible if its linguistic and cultural diversity, Europe's unique originality and richness –together with our tradition of social values such as solidarity- is preserved.

The Europe of Knowledge

It establishes that real wealth creation is no longer linked to the production of physical goods, but to the dissemination of knowledge, consisting of a capacity for innovation, research, education and training. This process is directly related to the aim of developing *life-*

long learning and the *promotion of the highest level of knowledge through broad access to education and its permanent updating* of which languages are an integral part.

The Europe of Languages

This takes us straight into the third and last key concept of the Europe of Languages. A Europe of peoples and cultures will only be possible if a *Europe of languages is built preserved and developed.* Additionally, both at an educational and professional level, multilingualism is a guarantee for the success of any exchange programme. Mobility should not only be understood as the movement of citizens within Europe, but also towards Europe, thus including people from other non-European nations with their own languages. In consequence, it seems evident that the construction of Europe should not only be built on the languages of each European nation, but also on those languages from newly arrived communities. Within such a view, the corporate communicative needs are not reduced to English. 'English-only' is not a satisfactory policy, either for establishing a solid base for European citizenship or for successful achievement in the world of industry, commerce and services. In sum, in Europe all languages are valuable: one's own language; your neighbouring countries languages; English; and other nations' languages.

On the basis of these three tenets, the Union's highest bodies issued a number of recommended set of policies, language programmes and instruments such as the projects and initiatives listed above (2.3.1.1) to recognise the need for a greater cultural and linguistic open-mindedness in all the countries of the European Union and also for specific educational programmes and instruments to promote them both (see European Commission 1995, 2003, 2006, 2008, and European Council 2007). These ideas find in CLIL programmes an educational approach incorporating the promotion of values and skills conducive to multilingualism and multiculturalism in Europe.

b. The educational-content dimension

The educational-content dimensions are actually two dimensions in one. Firstly they can be presented on the basis of socio-constructivist ideas in what has been described as *the four Cs* curriculum –the four Cs standing for **C**ulture, **C**ontent, **C**ognition and **C**ommunication (Coyle, 2000). This author suggests that: 'it is through *progression* in the

knowledge, skills and understanding of the content, *engagement* in associated cognitive processing and *interaction* in the communicative context that learning takes place'. The stress on the need for communication in the CLIL classroom is perhaps the most significant of all. From a methodological point of view, no CLIL class can rely on frontal teaching, no content can be really dealt with unless interactive communication and oral activities take place in the classroom (see 2.2.2.1 for a discussion of the effects of CLIL classrooms' discourse).

In practice, we are dealing with a methodological paradigm consisting in a knowledge framework' in which
- Curricular content is related to cognitive processing
- Cognitive processes are in turn related to linguistic requirements which must be met (Klapper, 1996)

Klapper argues that cognitive processes can be classified as experiential or explicit. Experiential processes are contextualized whereas explicit processes are decontextualized. Linguistic demands are higher in decontextualized processes which draw on abstract thinking. In contrast, linguistic demands are lower in experiential processing, which draws from concrete experience. Consequently, from a cognitive-linguistic point of view curriculum design in CLIL lessons must sequence its demands on the learner and plan units accordingly (for example, when planning geography lessons, first ask learners to draw a map of a country with mountains and rivers on it, then ask them to describe the whole river and mountain system in the country). Furthermore, it has been argued that the cognitive processes which take place in CLIL classrooms guarantee deeper thinking processes and equally good if not better learning of content. Empirical research is needed to prove this point, yet, first hand impressionistic descriptions account for it.

c. The language dimension

From the point of view of language pedagogy, the concept can be characterized as an extension of the United Kingdom's educational formula '*Language* across the curriculum' into '*Languages* across the curriculum' (see for example Wolff, 1998: 26). 'Language across the curriculum' incorporates in all subjects of a school curriculum a concern for the development of mother tongue linguistic skills. As a multilingual extension of this concept, 'Languages across the curriculum' sets off to include other languages besides the mother

tongue with the goal of promoting multilingualism using a transdisciplinary view of language development in the school system and involving educational, and social goals in the approach. Hence, the term encompasses many different forms of learning in which a language carries a special role alongside the learning of any specific subject or content.

As has been explained above (see 1.1.2 Language Acquisition in Formal Instruction Contexts), from the point of view of successful language acquisition, current methodological approaches in formal instruction (FI) show certain weaknesses. We have summarised that if we analyse the kind of input they provide to be processed and the output generated by learners in classrooms from a psycholinguistic point of view, deficiencies stand out. As we are going to see now, CLIL seems to be a formula to overcome such obstacles.

Quantity and quality of input
As was stated by Krashen (1985) and applied linguistics has overall accepted (explained in section 1.1.2: Input Hypothesis), a first requisite for language acquisition is exposure to input in the target language in considerable amounts: in fact, massive amounts if we bear in mind the acquisition of first languages in multilingual contexts (Pérez-Vidal *et.al.*, 2008). In addition to such a quantity factor, a quality factor must also be guaranteed for successful language learning. As for the latter, it is not only number of hours of exposure which is required, but there is also a need for intensity in such exposure, particularly in foreign language learning (Muñoz, 2006, García-Mayo *et. al.*, 2003) which explains why, for example, a teaching programme consisting of two hours per week over a year will be less efficient than the same total number of hours distributed over fewer weeks.

Conventional FI falls short of providing massive and intense exposure to input, as even in communicative classrooms input from the teacher in a foreign language is limited (see Canale and Swain, 1980). Conventional FI also ill-represents the wealth of domains of use and functions of language of general communication, quite obviously, as communication seldom covers the domains of real communication. Finally, input is often non-authentic, in that it is directed to the presentation of the linguistic system of the language, with the exception of communicative activities which generally take up only a small percentage of class time. Only periods abroad, common in many settings and schools nowadays, a core feature of European mobility

programmes at school and university level, provide opportunities for larger amounts of exposure and further the possibility of hearing language in all those domains as they represent natural immersion contexts of acquisition (Pérez-Vidal, 2007b). We would like to argue that, following Pérez-Vidal (2011) within the framework of the SALA Project[6], as a step between FI and stay abroad periods stands CLIL, offering higher amounts of exposure, with language used for a wealth of topics if we compare it with FI and always for a meaningful purpose.

<u>Quantity and quality of output</u>
The second requisite, put forward by Swain (1995) through her Output Hypothesis, is the need for learners to produce in the target language, to a sufficient extent. There are several benefits in production also summarised in section 1.1.2.

Conventional FI allows for learners' output with negotiation of meaning when learners are carrying out communicative activities. However, it can only be restated that the percentage of speaking time per student in conventional programmes with 2-3 hours of class per work tends to be insufficient, and so is the amount of individual feedback received. A stay abroad context obviously offers greater opportunities for output from learners, and, in its absence, a CLIL approach stands as the second best option, as already argued.

Finally, it is worth stressing the role of feedback and accuracy in language acquisition. As we are going to explain in more detail in the next section related to research on the implementation of CLIL, a weakness in the productive skills, both spoken and written, in grammatical and sociolinguistic competence has been reported in the literature evaluating such programmes. The proposals geared to redressing such an unbalance have been in the direction of suggesting the introduction of a concern for focus on form, a lesson to be learnt when designing and implementing new CLIL programmes (Harley *et al*., 1990; Lyster, 2007; Muñoz, 2007, Pérez-Vidal, 2007c).

When dealing with the more practical side of CLIL its pedagogical modalities and features need to be presented. Pérez-Vidal (2009), sees them as follows:

---

[6] The SALA Project seeks to contrast the effects of SA, FI and CLIL (see 2.3.2.4).

2.3.1.3 Modalities and features in CLIL programmes

Modalities

Three different modalities of CLIL have been suggested in relation with the focus of instruction in the classroom (Pérez Vidal, 2004). Indeed, as can be seen in table 5 below, CLIL programmes can be classified along a continuum with content at one end and language at the other against which each particular programme can be categorised.

Table 5. CLIL modalities

CLIL

CONTENT (content teacher)	CONTENT AND LANGUAGE (CLIL teacher)	LANGUAGE (language teacher)

Source: La enseñanza bilingüe: El Enfoque Integrado de Contenidos y Lenguas en Europa (Pérez-Vidal, 2008: 14)

In the case of a content-CLIL programme, to the left of the table we assume that a content class, particularly in secondary education however not always in primary, is taught by a content teacher with a focus on content teaching, who has an additional command of the foreign language used as the medium of instruction.

In the language-CLIL programme, to the right of the table we assume that a language class, taught by a language teacher, whose main concern is language, yet, has introduced some kind of 'content' or 'theme' element in the syllabus.

In the middle stands the most recommendable situation, one in which both content and language are the goal of the teacher who has sound knowledge of the content and practical expertise on language teaching methodologies.

Often the yardstick to identify the 'real' focus of the lesson is to consider what the students are going to be tested and marked on. In our opinion the 'ideal' CLIL programme, is the one just described which:
   a) combines a dual focus/objective, so that both content and language are taken care of, and both are tested and marked, following the rationale spelled out in the previous section

and the lessons to be drawn in relation of the need to 'focus-on-form';

b) adds hours of exposure to the target language, to the already existing conventional language lessons; that is, content lessons which would otherwise be taught through the L1 are now taught through the TL.

Table 6 shows those 3 possible strands organised in columns, with examples for each type of focus. In strand a) when the programme has a content focus and uses a target language different from the one generally used, either the linguistic knowledge necessary to follow the lesson is taken for granted (1.) or it requires extra curricular support (2.). In strand b) when the focus is on language, two possibilities result in terms of who handles the teaching of the content area: either a specialist in subject content uses a CLIL approach to introduce the students to the specific terminology of the subject in a different target language from the language generally used as the medium of instruction, while revising content, or different 'themes' are used to introduce meaning in the curriculum of conventional language lessons. In strand c) represented by the CLIL approach to education, both content and language are included in planning the curriculum. Either a language specialist working with the content subject is in charge of the CLIL programme, or a teacher with high levels of competence in a target language undertakes the CLIL programme, ideally after proper training in the approach, or finally, there is coordinated work between the content specialists and the languages specialists, either outside classrooms only, or both outside and inside, that is peer-teaching.

Table 6. CLIL programmes according to focus

a. CONTENT FOCUS	b. LANGUAGE FOCUS	c. CONTENT AND LANGUAGE FOCUS
(1) Language competence is given for granted.	Language competence is in focus.	Curriculum and syllabus planning includes explicit content objectives and explicit language objectives.
Students are tested on content.	Students are tested on content.	Students are tested on content with an explicit language component (often vocabulary).
	It usually requires involvement from both content and language teachers.	It usually requires a school programme. Teachers are language specialists. There is often peer teaching with a content teacher.
Example: Content teaching in Modern Language departments.	*Example: Revision modules to introduce terminology in a different language in a specific subject.*	*Example: CLIL taught in primary education.*
2) Language content is catered for in extracurricular classes.	Themes are a way to develop language competence with meaningful activities..	Curriculum and syllabus planning includes explicit content objectives and implicit language objectives.
Students are tested on content.	Students are tested on language.	Students are tested on content with a language component.
		It usually requires a school programme. Teachers are content specialists.
Example: International schools with language support units for new arrivals.	*Example: Content Modules in English language classrooms.*	*Example: CLIL programmes.*

Source: Content and Language Integrated Learning (CLIL). Teaching materials for use in the secondary school classroom. (Pérez-Vidal and Campanale, 2006: 29)

Features

The features of CLIL programmes can be identified along two different categories, as has been done for Canadian immersion programmes: core features and variable features (Johnson and Swain, 1997). Core features identify the essential traits which ensue from a CLIL educational approach and rationale, that is the one summarised in the preceding sections. Variable features identify the specific programmes developed by each particular school, or school consortiums, and are locally conditioned. Core features are listed in table 7 and specify the fact that the TL is not an environmental language, is not the L1 of the majority of teachers, who tend to be locally recruited, not internationally, and the fact that the programme increases the hours of exposure. A very interesting, key point concerns the 'culture' of classrooms, which is not the 'culture' of the TL country. That is, by and large we teach following the pedagogical tradition and the classroom rules of the local country, on a principle of subsidiarity, in contrast, say, with conventional international programmes which adopt the culture of the TL, the classroom practices and traditions of the TL country, and possibly use books published in that country. CLIL is based on the local curriculum and specifically designed classroom materials[7].

---

[7] Notwithstanding it has been noted that when CLIL approaches are put to practice, a great amount of variation can be found. Spanish programmes in secondary schools in Germany do claim that they are taking into account Spanish heritage culture in their CLIL programmes.

Table 7. Core features in CLIL programmes

	CORE FEATURES IN CLIL PROGRAMMES
1	The TL is not an environmental language so exposure is limited to the classroom
2	The L2 is the medium of instruction
3	The classroom is the only context of learning
4	Learners have limited competence in the TL
5	The rest of the languages in the curriculum are also given attention
6	Teachers share the learners' L1 and are competent in the TL
7	The curriculum taught through the TL is the same as it is in the L1
8	The culture of the content-based classroom is the same as the culture in the L1 classroom
9	The hours of exposure to the TL are increased
10	There is a European approach to multilingualism

Source: La enseñanza bilingüe: El Enfoque Integrado de Contenidos y Lenguas en Europa (Pérez-Vidal, 2008: 13)

Table 8 lists the variable features of programmes, those which define them and differentiate them. Beginning with curriculum and syllabus design, and a plan for the implementation and stabilisation of the programme, both made possible by proper budgeting and training of staff, the development of teaching materials, and the general support from the administration and the families, not to forget the status of the TL.

Table 8. Variable features in CLIL programmes

	VARIABLE FEATURES IN CLIL PROGRAMMES
1	The CLIL programme and curriculum planning (involving students and staff)
2	Continuity of the programme
3	Requirements and support for students
4	Teacher training programme for teachers

5	Financial resources
6	Material resources
7	Institutional support
8	Political external support
9	Social support (granted by families)
10	The status of the TL

Source: La enseñanza bilingüe: El Enfoque Integrado de Contenidos y Lenguas en Europa (Pérez-Vidal, 2008: 13)

When pondering on the variable features of a CLIL programme, it is immediately evident that CLIL teaching is about innovation and coordination. CLIL requires a project, a plan, linguistic and educational planning, which must come from the management of the school. CLIL programmes should be part of the school project, and must be accepted by families, and staff, who need to be prepared to face the challenge, in order to make a difficult path easy for the learners (Pérez-Vidal, 2009: 10-12)

2.3.1.4 The benefits of content and language integrated approaches

To finish with the presentation of the CLIL approach and before focusing on the preliminary research results of its implementation, the list of the expected benefits of CLIL follows. As Pérez-Vidal (2009) has suggested, they cover nearly all areas in which deficiencies have been found for other types of programmes and can be grouped in the three different categories used to describe the rationale behind CLIL (see 2.3.1.2) as linguistic, educational and pedagogical, and social as follows[8]:

---

[8] In recent years research has focused on the neurological benefits of CLIL approaches. In this sense, CLIL is an approach which can accommodate the edgy amygdale, the contemplating pre-frontal cortex and even the motivating median forebrain bundle in the brains of our [Content: Language]-Learners (Ting, T., 2011: 96). (Van de Craen *et al.*, 2008: 146, states that it is clear that learning in a CLIL environment results in discrete brain activity, which seems to echo the results of the cognitive aspects. These effects are the embodiment of brain plasticity in young learners and are as such not the results of CLIL itself. However, the aforementioned results show that CLIL exploits this plasticity and as such helps create *better* brains.)

Linguistic benefits

Undoubtedly CLIL programmes are often aimed at improving language skills above any other objective. In mainstream education, often CLIL programmes are organised alongside conventional language programmes. This offers the best opportunity to overcome the deficiencies outlined above and argue for a beneficial impact of the programme. As indeed:

a) By adding a CLIL programme to a conventional foreign language programme we are **increasing the number of hours of exposure to the target language.**
b) By using the language as a medium of instruction for content subjects, **communication is authentic**, not contrived, with the purpose of discussing the content matter of courses.
c) By focusing on content matter we are also **extending the number of domains and functions of language being used** (although the range of communicative situations and interlocutors will always be more restricted than in natural contexts such as periods abroad) adding on the general sociolinguistic repertoire.
d) By using a foreign language to deal with content subjects we are **stimulating interaction**.
e) By using the language for learning content **communication becomes meaningful as language is the means to communicate** not an end in itself.

Educational and pedagogical benefits

However much one of the objectives of CLIL programmes may be language, it tends to be clear that content teaching does not loose, except with respect to time. CLIL requires a slower pace, which often results in a need to prioritize and work through syllabae at a different rhythm albeit doing very sound work in classrooms. Pedagogically it also seems clear that CLIL has the following positive effects over teaching and learning. Both teachers' efforts to be 'clearer' than usually, and learners' efforts to follow lessons taught through the TL, not the L1 seem to exert a positive impact:

a) **A cross-sectional approach to language learning** is enforced as not only language teachers will care for languages.
b) Using language to communicate content is sensed to be more

> **motivating** than using language to reflect on it.
> c) By using a foreign language to deal with content subjects learning processes become important, something which **spurs didactics.**
> d) When using a foreign language to teach content subjects **studying skills** are required to overcome difficulties.
> e) The use of the foreign language is based on an **intrinsic motivation to communicate,** hence unconscious or implicit learning may take place.

Social benefits

Finally, as stated at the beginning of this chapter, there is a social slant to CLIL which makes it ideal for enhancing a broad view of the world, promote linguistic and cultural diversity in the classroom, and a European perspective to education. This can be enhanced if the other European instruments, the Portfolio and the European Framework of Reference are being part of the school project.

> a) Using a foreign language to teach subjects promotes linguistic diversity.
> b) It promotes intercultural approaches to education.
> c) It promotes European citizenship.

Having listed the expected benefits, it must be born in mind that this is an armchair list which requires confirmation by means of studies with empirical data. Several areas are already receiving attention: pragmatics and discourse in the classroom (see for example Dalton-Puffer and Nikula, 2006; Dalton-Puffer and Smit, 2007) and linguistic benefits (Lasagabaster, 2008; Van de Craen, 2008; Moore, 2009; Lorenzo, Casal and Moore, 2010) as summarised in the section 2.3.2. Evidence is beginning to show the direction in which success is moving, both intrinsic linguistic and educational benefits, and external benefits in the way of an increase in the school's profiles, etc (Marsh *et al.,* 2001, in the CLIL COMPENDIUM identify and develop 5 kinds of benefits: linguistic, content, cultural, learning, and internationalisation) (Pérez-Vidal, 2009: 12-13).

2.3.1.5 Challenges

The impact of the CLIL European construct goes beyond the change that other approaches to language teaching made. Not only language

but also, and probably most centrally, both content teaching and education as a whole are targeted with the proposal. Hence if important benefits should accrue, the challenges are of equal magnitude, quantitatively and qualitatively important.

Challenges are of different types, pedagogical, material, yet also personal. Pedagogically, teachers must train themselves into the rationale, and the technicalities to conduct a highly interactive lesson, while dealing with content and helping out with language. Methodologies shall have to cater for the development of learning strategies and autonomous learning. Teachers probably have to equip themselves and the learners with suitable materials, make use of internet and develop a work plan with short, medium and long term objectives. Individual learning styles and linguistic levels of competence also have to be taken into account, and, in multilingual groups, different L1s will also have to be considered. The issues related to evaluation and assessment become even more complex than they normally are.

All of this, requires, no doubt, effort and will. In the case of the school where the data for the current study are collected, the efforts made by the school have been summarised in Escobar and Pérez-Vidal (2004). Yet, the rewards are many, and we can spell them out by listing some of the mottos in the educational approach that we are advocating: broadening the mind to international European perspectives, focusing on learning strategies and on the development of cognitive strategies, too, designing and practicing interactive activities, contrasting languages, listening and speaking through pair-work and group-work, using ICT (Pérez-Vidal, 2009: 14).

At this point a summary of the results of research looking scientifically into the impact of CLIL programmes is required:

2.3.2 CLIL Research

In this section the empirical research analysing European CLIL with special emphasis on Spain and Catalonia is presented and discussed. Studies are contrasted with the Canadian results published to date. For such a purpose, the latter are presented first (2.3.2.1 below)

2.3.2.1 Canadian Research

The findings regarding the Canadian immersion programmes are extremely useful for the European CLIL approach and therefore Europeans have drawn on research carried out in Canada (see Johnson and Swain, 1997; Wechse, 2002). This means that Europeans have benefited from earlier methodological lessons in Canada, in particular regarding the need to take into account communication and to "focus on form" in the CLIL classroom as will be explained more in detail at the end of this section.

Two main domains have been on focus in the research analysing immersion programmes in Canada. In the first place, as Genesee puts it, the *product* or *summative* research (Genesee, 1987: 184). The work has concentrated on the investigation of the *language and content dimension* of different types of programmes. Analyses include extensive enquiries into the quality of language learned as well as its relationship to academic and social skills in the native language. Through measurements of receptive and productive abilities in programmes, the effects of programme variables on student achievement are identified. On the other hand, research has focused on the *qualitative study of bilingual education* through the analyses of different programmes, such as Johnson and Swain (1997) or Bernhardt (1992). This is *process* research, oriented to probing immersion teachers' beliefs, behaviours and strategies used in bilingual programmes, alongside students' behaviours. The relative scarcity of research in this second strand has been noted by Genesee who expressed his concern over: "The virtual absence of information concerning the *pedagogical and linguistic strategies* used by immersion teachers. Lacking such information, we are poorly prepared to train teachers in the most effective instructional strategies, a programme of research to investigate how teachers integrate academic and language instruction is called for" (1987: 18).

In the first strand of research, and since the days of the initial programmes, several studies have given a comprehensive evaluative picture of the outcomes of such programmes. Lambert and Tucker (1972) evaluated the first model with a high success rate. In subsequent publications Canale and Swain (1980), then Cummins and Swain (1986), Genesee (1987), Harley *et al.* (1990) and Bialystok (1991) in several synthesis reports have come to an overall conclusion that indicates the need for introducing specific changes to the programmes. While students in highly communicative immersion classes achieve higher levels of fluency and communicative ability, these do not lead to

higher levels of accuracy or more refined sociolinguistic skills. As Lyster's (1987) study revealed, there was a weakness in the learners' productive skills, spoken and written. Harley *et al.* (1990), and more recently Lyster (2007) have proposed balancing the experiential and analytical approaches, that is introducing approaches that focus on form.

More specifically this author points out that research has clearly demonstrated that immersion students, regardless of program type, develop much higher levels of second language proficiency than do non-immersion students studying the second language as a regular subject (i.e., for one period per school day). In comparison to non-immersion students, immersion students develop (a) almost nativelike comprehension skills as measured by tests of listening and reading comprehension; and (b) high levels of fluency and confidence in using the second language. However, production skills are considered non-nativelike in terms of grammatical accuracy, lexical variety, and sociolinguistic appropriateness.

Several studies paved the way to Lyster's conclusion. Harley *et al.* (1990) conducted a large-scale study of the second language proficiency of immersion students, operationalizing proficiency in terms of grammatical, sociolinguistic, and discourse traits. In comparison to native speakers of French of the same age (i.e., 11-12 years old), immersion students performed as well on measures of discourse competence, but "were clearly less proficient on most grammar variables, and especially on verbs in the oral grammar test" (p. 16). They also performed significantly differently on all sociolinguistic measures. Specifically, immersion students used significantly fewer instances of singular *vous* and conditional verb forms to express politeness. With respect to strategic competence, prior research had confirmed that immersion students were highly successful at using communication strategies enabling them to get their message across through recourse to their first language and the use of gestures, general all-purpose terms, or circumlocutions (Harley, 1984).

Concerning lexical variety, Harley (1992) documented a tendency for immersion students to use a restricted vocabulary limited to domains experienced in school, and to overuse simple high-coverage verbs at the expense of morphologically or syntactically complex verbs, such as pronominal and derived verbs. Allen *et al.* (1990) found generally that immersion students' first language significantly influenced their second

language lexical proficiency (see Jiang, 2000). Other studies of the interlanguage development of immersion students revealed non-targetlike uses of grammatical and sociolinguistic features that include, but are not limited to, the following:

- prepositions (Harley *et al.*, 1990)
- object pronouns (Harley, 1980)
- word order (Selinker, Swain, & Dumas, 1975)
- grammatical gender (Harley, 1979, 1998; Lyster, 2004a)
- features of the verb system such as the use of imperfective aspect, conditionals and third person agreement rules (Harley, 1986)
- productive use of derivational morphology (Harley & King, 1989)
- use of verbs with syntactic frames incongruent with the learner's first language (Harley, 1992)
- singular *vous* and mitigating conditionals (Harley *et al.*, 1990; Lyster, 1994; Swain & Lapkin, 1990)
- vernacular features and other informal variants (Mougeon & Rehner, 2001; Rehner & Mougeon, 1999)

What emerges from these studies is that immersion students are second language speakers who are relatively fluent and effective communicators, but non-targetlike in terms of grammatical structure and non-idiomatic in their lexical choices and pragmatic expression –in comparison to native speakers of the same age.

Genesee (1994a) described the productive skills of immersion students as

> linguistically truncated, albeit functionally effective (Genesee, 1994a: 5).

but also stressed that immersion students' second language proficiency does not limit their academic development

> The documented effectiveness of the immersion programs indicates that an approach in which second language instruction is integrated with academic instruction is an effective way to teach the language skills needed for educational purposes (Genesee, 1987: 176).

But would it also be possible for immersion students to develop a wider range of skills to enable them to use the second language for social purposes, with some degree of communicative effectiveness, as well as for educational purposes? Such would be more in keeping with the overall objectives of Canadian and other immersion and content-based programs which, in addition to ensuring normal first language development and academic achievement, aim to develop functional competence in both speaking and writing the target language, as well as an understanding and appreciation of target language speakers and their culture (Genesee, 1987; Met, 1994; Rebuffot, 1993).

Integrated programmes: the desirable future

The findings of Canadian research are extremely revealing for the design and implementation of programmes in Europe. A word of advice should be taken under consideration in the sense that an additional explicit analytic component in otherwise communicative modes of instruction, as CLIL approaches are, can only prove beneficial. Indeed, second language acquisition research has shown the limitations of implicit instructions unless in substantial amounts, particularly as far as adults learners, who can already benefit from an explicit focus on form, are concerned (DeKeyser, 2002, 2007). That meaning and form oriented instruction is simply superior to either one of the two individually seem undeniable nowadays (Hulstjin, 1989; Robinson, 1995). Communicative forms of instruction are characterised by a focus on meaning and communication, which is established by genuine interaction between the teachers and the learners generally through pair/group work interaction, a creative non-restrictive use of language, and via opportunities for the negotiation of task topics, activities which feature prominently in CLIL approaches. In turn, as thoroughly explained in the first part of this chapter (see 2.1.3), focus-on-form is characterised by the fact that attention is drawn towards language forms in order to develop linguistic awareness which may result in uptake and subsequently intake (DeKeyser, 2002). We now turn to a review of CLIL research conducted in Europe

2.3.2.2 European Research

This study is framed within the European context of CLIL research, a brief review of which here follows. For that purpose, the main research concerning the impact of CLIL approaches specially focusing on content and language outcomes will be first hereby presented. After

this, the most important research outcomes in Spain will be summarised focusing on the description and analysis of CLIL programmes and research outcomes derived from them in the different autonomous communities. Since Catalonia is the community where this study takes place, more detailed emphasis will be devoted to this region.

As regards research conducted in the European continent, several recent publications attest of the enormous impact of CLIL approaches (Pérez-Vidal and Escobar Urmeneta, 2002; Escobar, 2006; Hellekjaer, 2006; Ackerl, 2007; Dalton-Puffer, 2007, 2008; Lasagabaster, 2008; Moore, 2009, Lorenzo *et al*., 2008, 2010; Ruiz de Zarobe and Jiménez Catalán, 2009; Lasagabaster and Ruiz de Zarobe, 2010).

However, Moore (2009) has pointed out that research is fraught with serious difficulties:

> i) the narrow focus of many studies which do not include control groups of students learning in FI contexts;
> ii) the 'true' comparability of research and control groups as often CLIL groups include self-selected 'higher level' learners;
> iii) lack of adequate instruments to measure what is not a typical language learning context, in which content and language conflate;
> iv) the fact that many CLIL classrooms are experimental may mean that results should not be extrapolated (Moore, 2009: 121, 122).

Today, still little is known of the variables and the teaching and learning *processes* to which success or failure of the bilingual/CLIL programmes can be attributed, both educationally and linguistically. Similarly, the real *product* or benefits and gains are only begining to be quantified. Researchers seem to be working in two directions. One reflects on general policy issues, programme design, teacher education, teachers', students' and programme evaluation for CLIL. The second explores language acquisition, while also attending to content acquisition in some cases, with special attention given to exploring the impact on proficiency of different CLIL task designs (Escobar, 2006).

All in all, as pointed out by Pérez-Vidal (2009b), evidence has been accumulating in the past years showing that CLIL learners achieve higher levels of competence than their FI peers (Ackerl, 2007; Dalton-Puffer, 2007; Escobar, 2006; Hellekjaer, 2006; Lasagabaster, 2008;

Moore, 2009). European research has substantiated North American findings with regards the absence of negative effects on the L1 and promising evidence from the perspective of the European 1+2 goal suggesting that CLIL learners also become more motivated to study subsequent languages (Lasagabaster and Sierra, 2009; Merisuo-Storm, 2007) and are more successful when they do so (Cenoz and Valencia, 1994; Jessner, 1999).

As Moore (2009) explains, in his study of secondary school students' academic reading abilities Hellekjaer (2006) included a group of CLIL students in comparison with mainstream secondary school students. He found that the CLIL students outperformed general EFL students in the academic reading test. He claims several reading advantages for CLIL students:

- CLIL students quickly learn to read for overall meaning, to tolerate uncertainty and vagueness.
- They are better able to adjust reading strategy to reading purpose.
- They are particularly skilled at guessing or deducing word meaning from context and this results in an expanded vocabulary.

Still following Moore (2009), the author explains how Ackerl (2007) compares CLIL and mainstream writing abilities in final year examinations in an Austrian secondary school. Applying a form of error analysis to the two sets of data, she found that while the number of errors did not differ significantly, the nature of the errors does. With regards verb forms, for example, while the mainstream texts kept largely to present (simple and progressive) and past (simple) forms, the CLIL learners employed a wide variety of forms. She also calculated the ratio of verbs relative to the number of different verbs and again the difference seems significant: CLIL 57%, mainstream 29%. This suggests that CLIL learners' writing might be overall at a more sophisticated level than their mainstream counterparts.

Dalton-Puffer (2008) makes the point that while there are, of course, linguistically gifted students in the mainstream whose scores parallel those of CLIL students, overall scores tend to be better for CLIL which

> significantly enhances the language skills of the broad group of students whose foreign language talents or interest are average (Dalton-Puffer, 2008: 5).

The author distinguishes between content and language outcomes, a distinction adopted below to present the existing research outcomes.

Content outcomes

Before dealing in detail with the language learning effects of CLIL education, this author briefly comments how the learning of content is a continuous concern for educators and parents, as the issue of how being taught in the foreign language may affect the subject knowledge and skills of the learners. It is feared that since the medium of learning is less perfectly known than the L1, this will lead to reduced subject competence either through imperfect understanding or through the fact that teachers pre-empt this problem and simplify contents beforehand (Hajer, 2000). Generally speaking research results are, however positive, with most studies making the observation that CLIL learners possess the same amount of content knowledge as their peers who were taught in the L1. CLIL students have even been shown to outperform peer controls when tested in the L1 (Day/Shapson, 1996; Van de Craen *et al.*, 2006). This, it has been claimed, may have to do with the fact that CLIL students work more persistently on tasks, showing higher tolerance of frustration, thus acquiring a higher degree of procedural competence in the subject (Vollmer *et al.*, 2006). Vollmer and associates have also argued that linguistic problems, rather than leading to task abandonment, often prompt intensified mental construction activity (through elaborating and relating details, discovering contradictions) so that deeper semantic processing and better understanding of curricular concepts can occur. This argument finds supportive evidence in the research of Bonnet (2004), who found that students may well switch to the mother tongue when a conceptual problem has occurred but this does not normally lead to the solution of this problem. It certainly seems, therefore, that rather than being a hindrance, L2 processing actually has a strong potential also and in particular for the learning of subject-specific concepts. It must be mentioned, however, that there is also contrary evidence showing CLIL students to be at a disadvantage when tested on various school subjects (Washburn, 1997; Nyholm, 2002).

Language outcomes

General statements on the effect of CLIL on students' language learning outcomes are unsurprisingly positive. It is often observed that,

by way of CLIL, students can reach significantly higher levels of L2 than by conventional foreign language classes (e.g. Wesche, 2002) and that positive effects on communicative competence are visible (e.g. Wode, 1994; Klieme, 2006). A comparison of the performance of CLIL students and their non-CLIL peers on a standardised placement test in an Austrian secondary school showed that a higher percentage of students from the CLIL group reached the required B2 (CEFR) level than from the group who had followed only the conventional FL curriculum (Haunold, 2006). However, this does not mean that the non-CLIL group was without top scores. Rather, it is the case that the CLIL classes have a significantly broader band of students just below the top level. In other words, people with special linguistic gifts reach very good results, even high proficiency, also via normal EFL classes, but CLIL significantly enhances the language skills of the broad group of students whose foreign language talents or interest are average, as Daltons Puffer's quote above stresses. This is an effect which has been observed repeatedly (e.g. Mewald, 2004; Eder, 1998). Zydatiß (2006) has argued that school grades do not adequately reflect that CLIL classes have a higher average level of foreign language competence. The grades (in Germany but also in many other countries) are norm-referenced in the sense that they are usually given relative to the level which obtains in that particular group of learners (the class). This means that the actual grades or marks given tend towards a normal distribution even though an average grade in a CLIL class expresses a higher absolute level of language competence than in a regular class. In a system where university entrance for instance is dependent on school grades, this effectively puts CLIL students at a disadvantage and has led to the thinning out of bilingual streams in Berlin in the upper secondary years. (Zydatiß, 2006).

In any case, in her state of the art Dalton-Puffer (2008) stresses an important fact regarding outcomes on CLIL approaches:

> Under CLIL conditions certain aspects of language competence are developed more than others (Dalton-Puffer, 2008: 5).

Table 9 contrasts areas where clear gains are observable with areas where they are not. As we can see in it, the skills which are favourably affected are: a) receptive skills; b) vocabulary; c) morphology; d) fluency; e) creativity, risk-taking, fluency, quantity; and f) emotive-affective factors. On the other hand, those aspects which are either

unaffected or for which research is inexistent or inconclusive are: a) syntax; b) writing; c) informal/non-technical language; d) pronunciation; and e) pragmatics.

Table 9. Language competences favourably affected or unaffected by CLIL

Favourably affected	Unaffected or Indefinite
Receptive skills	Syntax
Vocabulary	Writing
Morphology	Informal/non-technical language
Creativity, risk-taking, fluency, quantity	Pronunciation
	Pragmatics
Emotive/affective outcomes	

Source: Outcomes and processes in content and language integrated learning (CLIL) (Dalton-Puffer, 2008: 5)

According to this author, while CLIL programmes rarely offer the same amount of contact with the language as actual immersion programmes, they do contribute to the passive language skills by enlarging the number of different speakers which learners are confronted with face-to-face and by (potentially) offering additional reasons for reading. On the side of the productive skills, it can be said with regard to speaking (e.g. Mewald, 2004; Rieder and Hüttner, 2007) that CLIL students often display greater fluency, quantity and creativity and show the kind of higher risk-taking inclination often associated with good language learners (Naiman, 1995). This presumably stands in direct association with the frequently observed positive affective effects of CLIL: after a certain amount of time spent in CLIL lessons the learners seem to lose their inhibitions to use the foreign language spontaneously for face-to-face interaction. A parallel effect of time and quantity can be observed on some aspects of English morphology (personal observation, cf. Zydatiß, 2006): particularly low-level processes like the third person 's or irregular past tenses but also the modals have been shown to gain a higher degree of automatization and appropriacy of use. The greatest gain in terms of the language system, however, is undoubtedly produced in the lexicon: through studying content subjects in the foreign language CLIL learners possess larger vocabularies of technical and semi-technical terms and possibly also of general academic language which gives them a clear advantage over their EFL-peers. Two footnotes deserve to be made here, however. In general the study

of vocabulary learning in instructed settings has shown that gains are particularly great if vocabulary is worked on explicitly: Interestingly, vocabulary is usually the only linguistic aspect which is explicitly treated in CLIL lessons (Matiasek, 2005). Furthermore the causality of CLIL in the enhanced vocabulary size of students has also been relativized by Sylvén's (2004) study with results showing that in a comparison of CLIL students and peers, reading habits are a stronger predictor than participation in CLIL for student performance in a vocabulary test.

The second column in table 9 displays areas that either seem unaffected by the extra foreign language exposure offered by a CLIL programme or have not been systematically examined so far. Observationally, the pronunciation of CLIL pupils does not seem different from that of their peers, but the issue has not been explicitly studied. An interesting issue in this respect is the question of how far the pragmatics acquired in the classroom translate into pragmatically adequate behaviour outside school and in how far CLIL students differ or do not differ from their peers with regard to their pragmatic learning. The conditions of language use in the classroom seem responsible for the fact that CLIL students do not outperform their peers in terms of the syntactic complexity of their utterances. This is an issue that will be further dealt with in the section on classroom language. While it was said above that the lexicon is the clear winner in CLIL, this advantage is largely constrained to technical language while the general and informal registers do not profit at all or not to the same extent (Sylvén, 2004).

The most important issue according to Dalton-Puffer's view (2008), however, is writing. Two studies conducted in Germany and Spain (Vollmer *et al*., 2006; Llinares and Whittaker, 2006) have recently investigated the written performances of secondary school students on post-teaching writing tasks in social science subjects. In both cases a significant share of the texts produced remained off target on a number of criteria, ranging from fulfilment of the required discourse function, via cohesion and coherence to grammar and appropriate style. The explanations of these deficiencies in academic literacy take recourse to the kind of pre-scientific understanding of the subject which is visible from these texts, but also to the fact that the general writing competence of the learners is in need of development, particularly since parallel results were obtained on writing tasks completed in the mother tongue. What is at issue here clearly is the role of writing in content-teaching in general, irrespective of the language it is conducted in.

Since very little explicit language teaching happens in CLIL lessons, it must be assumed that what the learners learn or do not learn is directly connected to the conditions of language use that hold during content teaching. In other words, according to Dalton-Puffer (2008), the structures of classroom discourse are the key to explaining the learning outcomes as they have been presented in this section.

The language of CLIL classrooms

CLIL lessons are lessons and as such reflect the conditions which make up the situative context of institutional education. It is widely known that educational discourse is determined by certain spatial and temporal conditions (buildings, classrooms, timetables) as well as the goals of the institution and the roles of the participants (cf. Edwars and Westgate, 1994; Walsh, 2006; Dalton-Puffer, 2008).

It follows, by simple power of logic, that CLIL students are listeners most of the time. What they listen to are the utterances of their teachers and peers. In list form the sources of spoken input are the following:

! teacher questions
! teacher feedback
! student answers
! student presentations
! reading aloud

Whether whole class discussion is much less teacher-centred is an issue worthy of discussion but will not be developed further here. On the linguistic level, the absence of lecturing is due to the fact that the prevailing style of CLIL lessons is interactional (as stressed in 2.1.4.b). As a consequence of this, in the students' input there is an absence of longer pieces which set out facts, concepts and the semantic relations holding between them in a coherent discourse of some syntactic and textual complexity. Instead, subject content is introduced by a sequence of teacher questions and students responses that follow the teacher's internal script (e.g. Ehlich and Rehbein, 1986). This script, however, is not accessible to the students and often remains inexplicit. In terms of language production this means that CLIL students most frequently employ their active language skills in answering teacher questions.

Another issue that is frequently topicalized in relation to CLIL and has received comparatively broad research attention (Lyster, 1997, 2004; Dalton-Puffer, 2008) is the topic of correction. It is often mentioned as an advantage of CLIL classrooms that learners feel more relaxed in using the foreign language because the focus of attention is on the meanings and not on linguistic form, something which is thought to be much closer to how conversations are conducted outside classrooms and therefore more natural.

Overall, language problems are not attended to with the same likelyhood as content problems. In so far, research results in Europe support common perceptions about CLIL classrooms as being places where linguistic form is focused on significantly less than in EFL lessons (see Pérez-Vidal 2007c for a study on this issue).

It is evident, however, that individual teachers differ in their attitude towards language problems in the CLIL classroom.

Having summarised the main views on the effects of CLIL within the general European research, we now turn specifically to research conducted in Spain with the new CLIL programmes in the different autonomous communities at primary and secondary level.

2.3.2.3 Spain

The richness of Spain's cultural and linguistic diversity has led to a wide variety of CLIL policies and practices which provide us with many examples of CLIL in different stages of development that are applicable to contexts both within and beyond Spain. In this section the details of this rapid development of CLIL in the Spanish scenario are explained.

As we have mentioned, interest in CLIL has spread exponentially through Spain during the last years, and CLIL programmes have received support from educational authorities and have been implemented in mainstream schools. The degree and characteristics of this implementation, however, vary greatly from one region to another, since the political structure of Spain comprises 19 autonomous communities that were granted political and administrative power by the 1978 democratic constitution. It must be borne in mind that Spain went through important changes in language policies after the dictatorhip during the decade of the 80s. In the bilingual communities

(Balearic Islands, Basque Country, Catalonia, Galicia, Navarre, and Valencia) the regional languages were granted official status in 1980 and have thus found their place in the educational system.

Following Ruiz de Zarobe and Lasagabaster (2010), the legislative frameworks guiding the Spanish education system are the Spanish Constitution (1978), the Organic Act on the Right to Education (LODE, 1978) and the Organic Law of Education 2/2006, 3rd May (Ley Orgánica de Educación LOE 2006) which develop the principles and rights established in Spain. Even though the Organic Law of Education offers the legal framework to provide and assure the right to education at national level, the autonomous communities regulate the adaptation of this Law to their territories. This fact allows them to have the power to administer the educational system within each region although the Organic Act of Education offers the core frame for the whole country.

Due to this diversity, there are as many models as regions, but they are developed with the same main objective: to achieve communicative competence in second and foreign languages across the curriculum. The different models vary significantly from one region to another, but can be divided into two main contexts:

- Monolingual communities, where Spanish is the official language. In these communities, education is partly done in Spanish and also in one or two foreign languages, when CLIL is implemented.

- Bilingual communities, where Spanish is the official language together with another co-official regional language, namely Basque, Catalan, Galician and Valencian, both of which are mandatory at nonuniversity levels and regulated at university level. In these communities, education is undertaken in both co-official languages, plus in one or two foreign languages, when CLIL comes into force. The result is then that for the bilingual communities CLIL implies a third or foreign language, as is the case of the present study.

In the case of bilingual communities, the support granted to regional languages since the 1980s through the mainstream education systems as the medium of instruction of content subjects has had a double influence on education. On the one hand, the expertise gathered after years of practice in bilingual communities has provided an excellent example for the design and implementation of programmes in monolingual communities. This knowhow has allowed different

regions across the country to transfer their experience and by doing so, monolingual communities have been able to keep pace with bilingual communities. On the other hand, in bilingual communities CLIL has evolved as the best approach to incorporate foreign languages in a system where already two languages need to be accommodated in the curriculum. Moving from regional to foreign languages has proved to be a natural way to generalise the use of more than one language as the medium of instruction. Therefore, increasing priority has been given to CLIL as the best way to foster multilingualism and language diversity, one of the aims of European policies in the last decade (Ruiz de Zarobe and Lasagabaster, 2010).

However, although the variety of CLIL-type provision models has increased over the last decade, not all the autonomous regions have implemented the programme in the same way. For example, as will be explained in more detail in the following sections, the Basque Country, Catalonia, Valencia or Galicia are all bilingual communities fostering multilingualism, but following different approaches and models. In Madrid or the Balearic Islands on the other hand, the Spanish Ministry of Education, Culture and Sports and the British Council signed the 'MEC/ British Council Agreement', to implement *the Bilingual and Bicultural Project* in 1995. This project aims at raising English language levels of children in state schools by following an official bilingual and bicultural curriculum. Other regions such as Andalusia are implementing *the Plan de Fomento del Plurilinguismo*, while in La Rioja several policy lines such as *Proyectos de Innovación Linguística en Centros* (School Language Innovation Projects) and Bilingual Sections are being set up by the regional ministry of the community of La Rioja. Last but not least, the Extremaduran Educational Authority (*Consejería de Educación. Dirección de Calidad y Equidad Educativa*) is promoting the so-called *Proyectos de Sección Bilingue* (Bilingual Sections Projects) in order to set up CLIL experiences in Primary and Secondary schools.

Due to the scope of this dissertation, it is not possible to analyse all 19 autonomous communities in detail. In the following sections, following an alphabetic order, the research carried out in Andalusia, the Basque Country, Galicia, and Madrid is analysed because these are representative examples of monolingual and bilingual communities which have implemented CLIL following different plans and projects. Since the Catalan language area comprises Catalonia (the context of the present study), the Balearic Islands and most of the Valencian

Community, these communities are analysed together at the end of the chapter more in deep.

Andalusia

Andalusia stands out in Spain as the regional community with a decided policy for the implementation of CLIL. In 2005, the Andalusian government launched the Plurilingualism Promotion Plan (*Plan de Fomento del Plurilingüismo en Andalucía. Una política lingüística para la sociedad andaluza*). The Plan represents a concerted effort to adhere to European policy and is built around five programmes incorporating seventy-four distinct strategies to be implemented over the period 2005–9. Its ultimate aim is to engender a radical shift from social monolingualism to multilingualism through education, under the European ethos that 'Europe will be multilingual or it will not be'. In Andalusia, it should be pointed out, possibilities for extra-mural exposure to and use of educational L2s are scarce and this reinforces the need for multilingualism through schooling.

In September 2007, figures showed four hundred and two CLIL schools participating in this project, of which two hundred and two were primary and two hundred secondary. Three hundred and forty three are developing programmes in English, fifty in French and eight in German. There are also twenty eight vocational branches offering CLIL modules (Moore, 2009; Lorenzo, Casal and Moore, 2010).

After one and a half years of implementation of the *Plan*, the Universidad Pablo de Olavide conducted an evaluation on behalf of the local administrations. More specifically, this research was carried out within the Andalusian Bilingual Sections programme, one of the cornerstones of the government´s Plurilingualism Promotion Plan.

Participant selection was organised in line with three major variables: urban/rural; primary/secondary education; and L2 (English, French and German). In the academic year 2007-2008, when the fact-finding component of the research was conducted, there were 403 schools across the region running bilingual sections. A two-stage sampling was employed in data gathering. In the first stage a sample of 61 institutions was randomly chosen across the eight provinces of the area of the study ensuring that each particular zone was evenly represented through a stratified sample approach. In the second stage, fourth year primary (aged 9-10) and second year secondary (aged 13-14) students were

identified as target respondents. This population was selected because, at the time in question, taking all three L2s into account, these were the learners who had had the longest possible experience of the bilingual programme within the Andalusian project. Control groups were evaluated alonside the bilingual sections. This was facilitated by the school organisation system itself.

The results here presented narrow the focus to four key research questions: Linguistic outcomes and competence levels; acquisitional routes and individual differences; L2 use in CLIL classrooms; and educational effects beyond the L2.

As for the first question, related to linguistic outcomes and competence levels, the study concludes a confirmation that CLIL learners show greater gains than their monolingual peers. It also suggests that the advantage extends to structural variety and pragmatic efficiency, hence encompassing language growth at lexico-grammatical and discourse levels. Evidence regarding incidental learning and positive transfer through content-focused instruction is another result within this question:

> CLIL learners L2 output features rhetorical moves and discourse patterns such as hedging and tentative language, hypothesising, impersonal structures and metaphorical grammar, typical of academic discourse but not addressed within primary or early secondary L2 syllabi. This suggests a considerable degree of positive transfer in the manipulation and maintenance of cohesion and coherence (Lorenzo, Casal and Moore, 2010: 10).

As far as acquisitional routes and individual differences are concerned, later start learners demonstrate competences comparable with early start learners. These results appear to imply that, in CLIL programmes, middle or late introduction can result in competences similar to those obtained in early introduction.

The third research question, L2 use in CLIL classrooms, shows that in tandem, and further evidenced by results detailing the types of materials and activities that the content and language teachers each use more frequently, each is dealing with a specific area of language expertise: the language assistants (native speakers that help language teachers in class) foster conversational style language, the language

teachers focus on sentence-level grammar and the content teachers work at the textual level. If this observation holds, it means that team teaching between content and language specialists is providing for a wider range of discourse input and therefore CLIL has the potential to provide an extremely rich language learning environment.

Finally, regarding educational effects beyond the L2, the results show that there is widespread agreement among bilingual section teaching staff (including L1 teachers and coordinators) that CLIL is beneficial to the educational process in general, an opinion echoed by parents and learners alike. First of all, there is a general consensus that CLIL enhances cohesion within schools. The teaching body as a whole considers that interdepartmental cooperation and cohesion is improved in bilingual sections. Coordinator interviews and teacher questionnaires have revealed that teacher involvement in CLIL planning is high and characterised by engaged collaboration between content and L2 teachers and language assistants (Lorenzo, Casal and Moore, 2010: 16). CLIL also appears to contribute to new forms of language awareness among both content and language teachers. Content teachers are leading to an acknowledgement both of the ubiquitous nature of language and to the fact that the successful transmission of subject matter content relies heavily on its linguistic selection and grading. In turn, language teachers are becoming aware that planning for advanced literacy is just as important as basic communicative L2. The gains reported in content focus, content learning and subject area objectives can be attributed to this increase in coherence. Nevertheless, it became apparent that many language teachers are still attempting to align language structures with content in a somewhat erratic manner (no doubt a legacy of their structurally biased professional development) and this area remains fuzzy.

From a language learning perspective, there is a wide consensus regarding the benefits which CLIL implies for L2 learning. In addition, according to Lorenzo, Casal and Moore (2010), in the context of this research, CLIL implies a new language model and it both coincides with and has contributed to a move away from the *ars gramatica* and towards a genre-based approach to language study. This conflates with the concept of Languages Across the Curriculum referred to above (see 2.3.1.2). As the authors of this evaluation project report claim,

> in conjunction, however, these results suggest that CLIL is an
> approach which may hold significant potential for European

education planning. Not only does it promote the integration of content and language, CLIL also fosters greater interdepartmental collaboration and conflates with other language development initiatives such as Language Across the Curriculum, the genre-based approach and multi-disciplinary curricula (Lorenzo, Casal and Moore, 2010: 19).

The Basque Country

CLIL experiences implemented in the Basque Autonomous Community (BAC) both in the private and the public sectors are presented here. The REAL Group (*Research in English Applied Linguistics*) covers a great deal of the research done in the Basque Country. As a case in point, Lasagabaster and Ruiz de Zarobe (2010) explain that the BAC is a bilingual community in which both Basque and Spanish are official languages and as such they are taught at school from the age of four and throughout compulsory education. It means that English is the L3 for Basque students. In the past years, CLIL programmes have started to blossom in some Basque schools. Although research in CLIL is quite young in the BAC, a number of studies reviewing both linguistic and non-linguistic outcomes have recently been undertaken.

Concerning language competence, Ruiz de Zarobe and Jiménez Catalán (2009) include a compendium of studies with data from the BAC based on the different areas and skills of language competence. Such research has provided empirical evidence supporting the view that there is a mismatch between receptive and productive skills, with better results on the side of the receptive skills in the target language as a result of a CLIL approach to education.

Ruiz de Zarobe (2008) compared the longitudinal oral and written competence of 89 secondary learners divided into three groups (a non-CLIL group, a CLIL group with one curricular subject in English, and a CLIL group with two curricular subjects in English). They were analysed longitudinally in the third and fourth year of Secondary Education and in the second year of post-compulsory Secondary Education. For the purpose of the analysis on speech production, five categories were used: pronunciation, vocabulary, grammar, fluency and content, each with a maximum score of ten points. The results were significantly better for the five categories analysed in the case of CLIL students with two subjects in English followed by CLIL students with

one subject in English. As for written competence, differences were significant in relation to content and vocabulary in favour of students with more hours of CLIL. These results show that the CLIL approach has a clear impact on both communicative tasks, but preferably on oral communicative competence and that written skills need to be trained on their own right, as Vollmer *et al.* (2006) and Llinares and Whitaker (2006) claim.

A second research study in the BAC was conducted by Lasagabaster (2008) and again CLIL and traditional EFL subjects were compared. The sample (198 secondary education students) was divided into three groups: a non-CLIL group aged 15-16 (NCLIL 15), a same age group which had been following a CLIL programme for two years (CLIL 15), and a 14-15 group with only one year of a CLIL programme (CLIL 14).

The results obtained in CLIL 15 indicated that the CLIL groups significantly outstripped their non-CLIL counterparts in the three language skills under scrutiny (listening, writing, speaking) and the grammar test, as well as in the overall English competence score (the mean of the aforementioned four different tests). Consequently, the conclusion to be drawn was that the CLIL approach had a clear impact on all the language skills and the grammar test analyzed in this study when students enrolled in the same grade were compared.

According to one of the working hypothesis, the students enrolled in CLIL 14 programmes would catch up with the NCLIL 15 students, despite the fact that the former were a year younger. The results not only bore out this hypothesis, but also showed that there were statistically significant differences in favour of the CLIL 14 students in some of the tests. As a matter of fact, the CLIL 14 groups scored higher than the NCLIL 15 in all tests but the listening. In addition, the differences turned out to be significant in the grammar and overall English competence tests. Therefore, not only did the CLIL 14 catch up with their NCLIL 15 counterparts, but they also surpassed them in overall foreign language competence.

As regards content, in order to confirm that content learning is not diminished in the languages taught through the foreign language, a longitudinal study was carried out during the years 2004 to 2006 in Secondary Education in 6 out of 12 schools involved in a *Plurilingual Experience*. The study involved a minimum of 50 students per level in three different grades (Secondary 1, Secondary 2 and 1st Post-

Compulsory) and a control group of ten students in each experimental group. On the one hand, a longitudinal comparative evaluation of the linguistic competence in English of both cohorts, the experimental group and the control group, was carried out with the help of the Cambridge ESOL competence tests. On the other hand, they evaluated the results of the content taught in English as a third language and the position of the other languages of the curriculum, Basque and Spanish. The results showed that students in the experimental group obtained better marks than the control groups in all tests and the differences increased after two years. Thus, the results suggest that the plurilingual experience increases the rate of learning the vehicular language, in this case English, and improve the linguistic and communicative competence of the third language. Apart from that, according to Lasagabaster and Ruiz de Zarobe (2010), learning through English does not create any obstacles in content learning.

Another study focusing on the impact of EFL CLIL programmes on the learning of Basque and Spanish was carried out by Egiguren (2006). This author compared the early implementation of an EFL approach with CLIL groups in the BAC and examined wether the type of approach (EFL versus CLIL) exerted any effect on the normal development of Spanish and Basque. This author compared two groups of students, the first one made up of students who started to learn English at the age of 4, and the second one at 8, but the latter also had two hours per week of Arts taught in English. In this case no differences were found when the participants' proficiency in Basque and Spanish was compared at the age of 10, which leads Egiguren to conclude that CLIL does not seem to hinder the acquisition of the two co-official languages (no mention was found as regards results in English). Furthermore, in the questionnaires and interviews carried out as part of the qualitative component of the study, both teachers and students emphasized the personal enrichment which involves learning various languages. They also seemed to be more motivated about their work, more conscious of the benefits to be obtained by learning several languages and more willing to make an additional effort.

Finally, moved by the positive results reported by many private schools, the Department of Education, University and Investigation (DEUI) of the Basque Government has also encouraged the teaching of curricular subjects through foreign languages, mainly through English, beginning in 2003-2004. In order to test the effectiveness of CLIL programmes, the DEUI compared CLIL and non-CLIL groups at two

points in time in compulsory (ages 13-16) and optional (16-18) secondary education (ISEI-IVEI, 2007). The findings show that the CLIL groups obtained overall better results than the non-CLIL groups in oral and written production and comprehension. Besides, the study claims that the content knowledge acquired does not decrease when the teaching is carried out through the English language, and that the level is similar to that obtained by the non-CLIL groups who are taught either through Basque or Spanish, depending on the linguistic model they are in.

Galicia

This section deals with the main courses of action and the preliminary results related to CLIL implementation in Galicia, a Spain's north-western region originally bilingual in Spanish and Galician (San Isidro, 2009b).

In Galicia, CLIL pioneering experiences started in 1999 as pilot projects in some secondary schools. These pilot projects resulted in the formal regulation of CLIL provision through several legal directives. The CLIL Galician model consists of teaching non-linguistic subjects by means of integrating additional languages in a progressive way. The name attributed to these classes is Bilingual Sections. So what the Galician educational system has is subjects taught on a bilingual basis (Additional Language-Galician or Additional Language-Spanish), not bilingual schools with an integrated curriculum. The educational regulation of CLIL teaching in Galicia has run parallel to the gradually increasing number of schools taking part in CLIL programmes - from the initial 12 secondary schools to the present-day 200 primary and secondary schools with 600 bilingual sections.

This increase has been brought about by the "Plan de Linguas" (Plan of action aiming at fostering the learning of environmental and additional languages, with a real substantial investment on the part of the government), which has obviously boosted CLIL implementation. What is clear is that different actions have contributed to improving the linguistic competence of both teachers and students and, above all, motivating them to understand additional languages as something instrumental in their life-long learning experience (language assistants, summer immersion programmes, teacher training, special attention to foreign students, leaves and awards as motivating strategies, etc.)

As for the first results related to CLIL implementation in this community, in terms of motivation and the increasing number of schools and teachers involved, the results of recent studies show a general case for we success. However, much still remains to be done. General objective testing analysing linguistic results is still necessary to check the 'purported benefits of CLIL in Galicia' (Lasagabaster, 2008).

In June 2008 a pilot self-reported 30-item questionnaire was sent to CLIL teachers in 114 schools with the following CLIL-related items focusing on student's results. The CLIL-related items were:

1. Students' motivation towards additional languages has increased.
2. Broadly speaking, students have improved their oral comprehension and expression in the additional languages.
3. Broadly speaking, students have improved their written comprehension and expression in the additional languages.
4. Students have improved their performance in CLIL subjects.
5. Broadly speaking, students have improved their linguistic competence in the two local languages.
6. Students' interest in other cultures has increased.

Results seem to be quite positive, above all in items related to motivation and additional languages improvement (items 1 to 3). Evidence for items 4 to 6 related to CLIL subjects, environmental languages and foreign culture show more varied results, although positive as well.

Effort, involvement and investment on the part of the Galician Educational Department, the 'Plan de Linguas' has made it possible to see the first successful results regarding CLIL implementation and the revitalisation of additional language learning.

Madrid

As regards CLIL programmes in the Madrid area, there are two main programmes in operation currently. The first one, already mentioned, is the result of an agreement signed between the Ministry of Education and The British Council in 1996, and is known as the MEC/British Council Project. The second one was launched in 2004 by the

Comunidad Autónoma de Madrid (CAM) and it works under the name of CAM Bilingual Project.

Following Llinares and Dafouz (2010), the general features of the MEC/British Council Project in primary and secondary education can be summarised as follows:

- Teaching through authentic materials across the subject areas with a strong focus on literacy skills (reading and writing) from a very early age (infant education).
- Exposure to natural language so that learners soon develop syntactic structures and meta-linguistic awareness, as happens in real communicative contexts.
- Implementating a whole school project, where the Spanish and English teachers, as well as other staff, work in a team effort to achieve the best results. In addition, an active implication on the part of parents and students themselves seems essential to achieve success.
- Establishing a twin-school network between Spanish schools and their English counterparts, whereby teachers and pupils are encouraged to develop joint curricular projects in all subject areas as well as to organise staff and student exchanges.

Some of the general outcomes of the MEC/British Council Project are as follows: students display high concentration skills in all subjects, higher-order cognitive skills such as questioning, summarising, predicting and hypothesizing are also developed. In addition, students show more willingness to work collaboratively, higher personal confidence, ability to confront challenges and awareness of cultural differences (Reilly and Medrano, 2009: 64; Halbach, 2009: 20-21).

As for the CAM Bilingual Project, in 2009-10 there were a total of 206 bilingual schools involved, which means approximately 40,000 learners and around 1400 teachers. The first cohort of learners involved in this project reach secondary education in 2010-2011. In terms of organization and legislation, a minimum of 30% of the syllabus needs to be taught in English, and a maximum of 50%.

According to Llinares and Dafouz (2010), in terms of oucomes, there are higher levels of motivation from all stakeholders involved (students, teachers and parents) as well as more self-esteem and confidence. Academically, students within the CAM Bilingual Project obtain higher results in foreign language competence, especially in the

receptive skills (listening and reading), although evidence is still not so clear for non-linguistic areas.

As a result of the implementation of the MEC/British Council Project in 1996 a research group in the Universidad Autónoma de Madrid began collecting a corpus of classroom data in two schools following this programme. The UAMLESC (Universidad Autónoma de Madrid Learner English Corpus) started in 1998 recording and transcribing pre-primary classes (five-year-olds) and followed the same children throughout primary education. In her analysis of five-year-old children's realisation of communicative functions in different bilingual schools, Llinares (2006, 2007a) found that learners exposed to a lower quantity of input outperformed other learners in higher-immersion contexts in the amount and functional variety of the language produced, when the lower-immersion learners took part in tasks designed to encourage their participation in self-initiated interactions. These results show the relevance of the quality of exposure over quantity of exposure, as far as functional features of foreign language learning are concerned.

Another interesting research project is PROCLIL (Providing Guidelines for CLIL Implementation). It is a European research Comenius Project coordinated by the University of Cyprus, which centers on CLIL as a recommended pedagogical procedure. As far as data collection in the Madrid area, the study aims at identifying good practice in CLIL programmes and to investigate its effectiveness as regards the learning of language and content at infant and primary schools. Findings so far indicate that primary teachers tend to focus more on content (64%) than on language (36%) and that their input is mainly in English (90%). As for didactic units, teachers devote more time to practical activities (53%) than to reasoning/exploratory talk (47%). Concerning skills, teachers devote more time to listening (40%) than to talking (26%), reading (18%) and writing (16%). As far as pupils are concerned, their talk gradually progresses from Spanish to English across the three cycles of Primary. In the first cycle, for instance, the percentage for Spanish is 70% while English is used about 30%. As regards materials, teachers mainly use textbooks (56%), their own materials mostly based on English materials (36%), and, to a much lesser extent, digital resources (8%). Finally, regarding evaluation, teachers tend to focus more on the assessment of content knowledge (65%) rather than on language (35%) (Llinares and Dafouz, 2010).

Concerning research on CLIL in secondary education in Madrid, the UAM-CLIL project began in the academic year 2005-06, with the collection of data in two CLIL classes from two schools participating in the MEC/British Council Project. Students were followed from 1st to 4th year of Obligatory Secondary Education in the subject of Social Science, taught in English. Three types of data were collected: whole class discussions, written compositions and oral interviews. Using Systemic-Functional Linguistics as the framework for the analysis, the main objective of this project was to describe the features of the language of CLIL students in relation to the language needs of the discipline they were studying, comparing the degree of lexico-grammatical development of English in the students, and the functional realisations found in their spoken and in written discourse. Both longitudinal and cross-sectional studies have been carried out. Most studies have been cross-sectional, comparing the same students' written and spoken performance (Whittaker and Llinares, 2006), and comparing CLIL students with parallel students of the same topic in their L1 (Llinares and Whittaker, 2009). These studies show that there is little difference between the spoken and written modes in CLIL learners and that, in general, CLIL students use appropriate lexis to express content-specific ideas. When compared to their Spanish peers' performance on the same topic in the L1, differences such as CLIL students' use of more clauses and fewer phrases to express circumstances were noticed. Other studies have focused on development over two academic years focusing on the complexity of students' noun phrases (McCabe, Llinares and Whittaker, 2008) and L1 transfer (Vázquez and Llinares, 2009). The first study concludes that CLIL students do not control the systems that signal given and new information while the second argues for the possible effect of teachers' interactional styles on Spanish CLIL students' transfer errors in subject dropping. Although most of the analysis has focused on the way the students deal with the representation of the content of the different topics (the ideational function), the UAM-CLIL research group has briefly studied how they intervene in that representation, using expressions of modality (the interpersonal function). Preliminary results indicate that CLIL students have a limited repertoire in their use of modality, with an overuse of a multifunctional "can".

Finally, as a way to conclude this section and before focusing in greater detail on Catalonia, the Balearic Islands, and Valencia, it is worth noticing that, whereas a significant number of studies completed in

Spain have tackled oral interactional patterns, those focusing on content-related results are very scarce.

> Although foreign language skills as a whole seem to improve, the benefits of CLIL are not as clear-cut for non-linguistic areas. In our opinion, assessment concerning both language proficiency and content knowledge is urgently needed (Lasagabaster and Ruiz de Zarobe, 2010: 286).

2.3.2.4 Catalonia

The following sections focus on the description and analysis of the impact of CLIL programmes specifically in Catalonia, the autonomous community in which our data are collected. However, the case of the Balearic Islands and the Valencian Community are also going to be briefly presented since they are part of the Catalan language area[9]. Following a summary of the legislation, the general linguistic policies at primary, secondary and tertiary levels, and the nature of the CLIL programmes, we seek to report on the numbers of students, the languages, and measurable outcomes of the programmes. For such a purpose, we are mainly adopting Pérez-Vidal and Juan-Garau's perspective.

The Catalan territories: Catalonia, the Balearic Islands and Valencia

Today the Catalan language area in Spain involves three autonomous communities in which intercultural encounters with two official languages, Catalan and Spanish, coexist with a myriad of other tongues. Vila (2008: 34) stresses the fact that:

> Contemporary language policies in Spain have become much more pluralistic than they used to be until 1978. Of course, Castilian remains the State's sole official language, but Catalan has become official also in almost all its historical territories, namely Catalonia, the Balearic Islands, and Valencia. All school children–i.e. including non-Catalan L1–living in these territories must learn both official languages.

---

[9] It must also be noted that the UBI is a partner in the SALA project within the ALLENCAM group.

A historical overview shows that, starting back in the seventeenth century up to the 1970s, Catalan went through a long period in which it was losing ground. However, language policies were at that point reversed, coinciding with a general change in ideology. We would like to suggest that the change affecting the Catalan language was spurred by the European strategy towards multilingualism which reinforced the normalization processes undertaken in the Catalan language area–just as they occurred in the two other autonomies in Spain, the Basque country and Galicia–. Hence, multilingualism has favoured minority languages in general and Catalan in particular, not only because of its explicit defence of minority languages, but also because of its plea for diversity. Vila (2008: 34) describes the situation as one in which "the current distribution of power" leads to a dual responsibility on language policies affecting compulsory education: since the late 1970s language policy has been distributed between the central and the autonomous governments. This author further argues that it is because of the fact that Catalan is the language of instruction (i.e. an immersion programme similar to CLIL) that Catalan sociolinguistic competence among pupils in schools has risen, whereas in the Valencian Community, where for the most part schools do not use Valencian, as described below (only 24% of the schools do) pupils do not always become productive bilinguals.

Within such a background of linguistic revitalization as far as Catalan is concerned, the autonomous governments of Catalonia, the Balearic Islands and Valencia have duly and responsibly applied the European agenda on multilingualism. We will now focus on the CLIL initiatives undertaken under such policies.

Background to the research: CLIL programmes

In this section the case of Catalonia is presented, with a reference to the Balearic Islands and the Valencian Community. These autonomous communities represent the two largest Catalan speaking societies in Spain. It must be emphasized that the linguistic revitalization of autochthonous languages, which took place during the final decades of the 20th and early 21st century, following the first democratic Spanish constitution (1978) and the publication of statutes of autonomy and linguistic normalization laws for the three communities seems to have established a solid ground on which to build plurilingual policies encompassing third and fourth languages.

Following a brief presentation, due to the fact that we are referring to the context of the present research work, Catalonia is first described considering legislation aspects, programme characteristics and in-service teacher training, at the primary, secondary and tertiary level. Second, against this background, current research assessing the linguistic outcomes of such initiatives is presented.

a. The case of Catalonia

PRIMARY AND SECONDARY EDUCATION
Catalonia lies on the north-eastern coast of Spain. It comprises four provinces and has a population of 7 million inhabitants. According to the last official survey (Pons & Sorolla, 2009: 31), 97.5 % of the population reportedly understands Catalan, the community's autochthonous language, while 85 % can speak it, 90.5% can read it and 62.3% can write it. Secondary education students rate their competence in Spanish as being higher than in Catalan (*Consell Superior d'Avaluació del Sistema Educatiu,* 2008b).

If we now turn to Engish as a target language, figures concerning English in the last official survey on results at the end of compulsory education, at age 16, and after approximately 840 hours of formal instruction (FI) reveal that 54.1% of the students have reached an A2 level of the CEFR in oral comprehension, while 51.8 % have done so in reading comprehension. When gauging productive skills, however, only 41% have reached an A2 level in English. The official survey conducted with data from students at the end of post-compulsory education, that is after completing roughly 1050 of FI, shows that 68% students have reached an A2 level and an average of 38% a B1 level in all four skills and grammatical competence tests (*Consell Superior d'Avaluació del Sistema Educatiu,* 2005: 67-69).

Programme description

In 2005, the Catalan Ministry of Education launched the current Plan of Action for the Promotion of Third Languages (*Pla d'impuls a les terceres llengües*) including several strands. The most prominent one is the Experimental Foreign Language Plan (*Pla Experimental de Llengües Estrangeres*, PELE)*,* which seeks to promote integrated school projects (CLIL plus project-based orally focused modalities). This plan is a natural continuation of the 1999 CLIL ORATOR scheme, with English as the main language which by 2005 had resulted in a total

of 32 projects mostly in the area of Science and Technology (Pérez-Vidal & Escobar-Urmeneta, 2002). Earlier accounts of CLIL in Catalonia can be found in Navés and Muñoz (1999), Nussbaum (1990), Pérez-Vidal, (1997), and Pérez-Vidal and Escobar Urmeneta (2002).

As table 10 below shows, the number of schools and students participating in the Plan of Action are by all means high. However, the funds allocated to each school are not, as only 2,000 €are received for the whole three-year project. Other benefits are also available such as the fact that participant schools are eligible for teacher education schemes, native teaching assistants and pedagogical assessment. To cover teachers' English language needs, 21,700 teachers, both English specialists and non-specialists are receiving English language training in local official language schools

Table 10. Learners and institutions participating in the Plan of Action (2005-09)

	2005-06	**2006-07**	**2007-08**	**2008-09**	**TOTAL**
Institutions	100	307	521	785	1,713
Learners	17,986	85,598	164,232	261,109	528,945

Source: Data provided by the Education Department of the Catalan Autonomous Government

b. The case of the Balearic Islands

Following a decade of implementation of the British Council MEC project in this autonomous community, CLIL in the Balearic Islands (Spain) has been mostly implemented through the European Sections Programme, which was first launched in the academic year 2004-2005 and has grown exponentially since then. In 2008-2009, a total of 119 European Sections were approved in 100 primary and secondary schools in the Balearic Islands (*Conselleria d'Educació i Cultura*, 2008). The number of learners and teachers taking part in the programme is higher in primary than in secondary education. Schools taking part in the programme are state-funded. The first foreign language chosen has been English. A new European Section can be started to teach any non-linguistic area, subject or module of the curriculum totally or partially in the foreign language chosen (see Pérez-Vidal and Juan-Garau 2010 for a more detailed account of CLIL programmes in the Balearic Islands).

c. The case of the Valencian Community

Early education in foreign languages, or plurilingual education, has been introduced by means of the Enriched Bilingual Education Programme (Programa d'Educació Bilingüe Enriquit, PEBE). This programme allows for the introduction of English as the medium of instruction for content subjects at age 6, when children start primary education, for 1:30 hours weekly. Schools adopt PEBE on a voluntary basis with no extra funding or human resources allocated to them. In the year 1999, 55 schools adopted the PEBE. Since then numbers have steadily increased up to the current figure of 291 primary schools in the whole of the Valencian Community.

The expansion of the PEBE coincides with new legislation establishing organic regulations to implement plurilingual education in the last two years of nursery education. The new decree permits the early introduction of English at the age of four in nursery education for 1:30 hours per week distributed in two or three sessions. In 2008-09, a total of 240 schools incorporated English at this level. Plurilingual policies have recently received further impulse thanks to yet another new regulation allowing for an experimental programme for nursery and primary education to be launched. On a voluntary basis, schools can establish their curriculum as follows: 80% taught through English, 10% through Spanish, and 10% through Catalan. Three schools in the Valencian Community will implement this programme, largely modelled on international schools, in the academic year 2009-10 (see Pérez-Vidal and Juan-Garau 2010 for a more detailed account of CLIL programmes in the Valencian Community).

Research on CLIL in the Catalan language area

In the past two decades, a number of language research groups have seen the light in Catalonia. Those analysing linguistic outcomes of CLIL programmes are in alphabetical order: the consolidated group ALLENCAM (*Adquisició de Llengües des de la Catalunya Multilingüe*) within which the SALA project has been conducted, among others, coordinated by Carmen Pérez-Vidal at Pompeu Fabra University (UPF) and in coordination with the SALA project at UIB coordinated by Maria Juan Garau[10]; the GRAL research group (*Grup*

---

[10] The ALLENCAM group includes a third state-funded research project, *Periferias*, coordinated by Aurora Bel.

de Recerca en Adquisició de Llengües) within which the BAF project has been developed, among others coordinated by Carmen Muñoz at University of Barcelona (UB); and the GREIP research group (*Grup de Recerca en Ensenyament i Interacció Plurilingües*), coordinated by Lucila Nussbaum Capdevila at the Autonoma University of Barcelona (UAB) which has developed the ARTICLE project, among other work related to CLIL, undertaken by Cristina Escobar.

Investigations in the Universitat Jaume I, in Valencia, carried out by the research group LAELA, coordinated by Eva Alcón, have focused on the acquisition of pragmatics in bilingual and multilingual acquisition contexts.

Following such alphabetical order, the study of the linguistic benefits of CLIL programmes in Catalonia and the Balearic Islands is being carried out within the framework of the SALA-COLE Project, a coordinated effort, led by Pompeu Fabra University (UPF) involving the University of the Balearic Islands (UIB) which began in 2004. Pérez-Vidal (2010) reports on the main goals of the project: the aim of the SALA-COLE Project has been to measure the effects on the acquisition of L3 English by Catalan/Spanish bilinguals of two learning contexts with undeniable interest, due to their linguistic, social, and even political impact: a Study Abroad (SA) context spent in the target-language country and a CLIL context of acquisition. These contexts have been, in turn, compared with the regular Formal Instruction (FI) context in English that students receive in their home institution, for both secondary and tertiary EFL learners.

This author further develops the hypothesis underlying the study, *the gradation hypothesis*, which establishes a continuous line moving from the conventional, eminently formal, classroom context (FI), through the CLIL classroom, which approximates natural contexts of acquisition more closely, to the Study Abroad (SA) or fully natural immersion context. To explore this hypothesis, the Project has focused on the contact opportunities and learning conditions provided by these different contexts and their impact on learning outcomes and attitudes at different stages in the acquisition of English by two age groups, adolescents and adults. The latter consists of a sample of university students with advanced levels of English and the former a sample of secondary school learners with intermediate levels. Such a mosaic of comparisons will hopefully allow for generalizations to be drawn concerning the factors and circumstances that promote and impede the

acquisition of languages. Furthermore, quantitative data have allowed to examine the gains with respect to initial competence levels for each context and sample groups, while the qualitative data have helped explain individual differences in gains in productive and receptive abilities of CLIL programmes.

The first results of the effects of the CLIL context are being issued. With data from secondary education in the Balearic Islands, Juan-Garau *et al*. (2008) and Salazar-Noguera and Juan-Garau, (2008) concluded that content-based approaches were very positively assessed by both teachers and learners, although they were not sufficiently well known or extended as yet. They were viewed as an excellent alternative to foster multilingualism (see also Juan-Garau, 2008).

Another issue on which SALA has focused is input and interaction in CLIL lessons. Teachers' reactions to learners' interlanguage in primary and secondary CLIL lessons in Catalonia (Barcelona, Lleida) have shown that no focus-on-form is found in teachers' input addressed to learners, a strategy recommended in the Canadian literature which discusses results in immersion settings in that country. More specifically, a representative sample of integrated lessons conducted in Catalonia were analysed by Pérez-Vidal (2007) with the objective of exploring teachers' input and interaction strategies in relation to the communicative nature of teaching and the presence of focus-on-form episodes. The study was conducted on a small sample of primary and secondary education content lessons in 3 different school programmes in Catalonia, selected so as to include content lessons taught through the medium of English from different geographical locations and including both primary and secondary levels. The results reveal how:

> Focus-on-form input was inexistent in the classrooms, and that three quarters of the moves in teachers' discourse were geared to eliciting output from learners, checking understanding and referring to classroom content matter (Pérez-Vidal, 2007c: 49).

Pérez-Vidal concludes by explaining how extensive research carried out in the Canadian immersion programmes shows the key role played by focus-on-form in the communicative interaction taking place in integrated pedagogy. She highlights that

> the CLIL lessons analysed show high concern for meaning, but not for form. If our small sample is representative of more

extensive practices, there seems to be a need for introducing FoF approaches to complement current practices in CLIL teaching, as Lyster (2007) has emphasised (Pérez-Vidal, 2007c: 50).

Further, preliminary SALA results contrasting CLIL programmes and FI in the secondary learners' group point to the presence of significant differences in several oral fluency measures between CLIL and FI students, to the advantage of the former, especially in rate of speech, as measured in words per minute produced (Juan-Garau, submitted). The advantage of CLIL learners tends to increase over time. The data contrasting FI and SA in this group are currently being collected. More specifically, the study addresses the following research questions: 1. Do EFL CLIL learners speak more fluently than EFL learners who exclusively follow a FI programme in English? 2. Do EFL CLIL and FI learners make significant oral fluency gains in the course of one academic year? Participants were 27 secondary school EFL learners in Year 2 (ages 13 to 14). All the subjects in the CLIL group (N=16) and the control group (N=11) were Catalan/Spanish bilinguals. As regards treatment, the CLIL learners studied Social Science (History and Geography) in English for three hours weekly, in addition to receiving formal instruction (FI) in this language in their English class for another three hours per week. It was their first year in the CLIL programme. The control group was formed by learners also taking part in a European Section through the medium of French so as to ensure comparability with the CLIL group. They only received FI in English. Thus, the CLIL and control groups were exposed to approximately 180h (90h CLIL + 90h FI) and 90h (FI) conducted in English respectively (Juan-Garau, submitted: 4).

Concerning the first research question, results indicate that CLIL learners speak more fluently than learners who exclusively benefit from EFL education. The fact that there is an initial advantage of the CLIL group at T1 in terms of speech rate and pause duration must be explained by intervening factors other than the impact of the CLIL programme itself, as learners have just started it. These factors include learner attitude and motivation as well as admission to the programme, which in the case of the school under study is done on a voluntary basis with the parents' consent and taking into account the learner's academic record. It should be noted, nevertheless, that participants in the control group are also in the school's CLIL programme, only in French, which makes them comparable to the CLIL group. By T2, after

an academic year, the two groups have significantly grown apart in most of the measures considered, confirming a higher oral fluency for CLIL subjects.

As for the second research question, fluency gains between T1 and T2 are only significant in speech rate and only for the CLIL group. Still, a tendency towards improvement is generally apparent in the CLIL group, a developmental pattern unparalleled in the control group. It might be argued that an academic year, given the exposure provided, was not sufficient for learners to register more substantial benefits.

> In sum, the European Sections Programme appears to have a positive effect on learners' oral fluency, especially concerning their speech rate (Juan-Garau, submitted: 4).

Results contrasting CLIL and FI for the group of university learners are currently being collected. When contrasting FI and SA, findings show clear advantages in listening, oral fluency and accuracy, as well as in written fluency, accuracy and lexical repertoire after the SA three-month period, in contrast with FI. The use of formulas increases. In contrast, phonology improves significantly more after FI at home and grammar similarly in both contexts (see Pérez-Vidal, 2009b, for an overview of results).

The second group within which the linguistic benefits of CLIL programmes are being analysed is the GRAL group at the University of Barcelona (UB). Parallel to the ALLENCAM-SALA studies, in terms of language proficiency gains Victori[11] and Vallbona (2008) conducted the following study in a semi-private primary and secondary school, which was interested in implementing CLIL-based subjects in Grades 3 to 6. Students who had never received CLIL instruction before were compared with students of the same level who had received one hour a week of CLIL instruction in the subject of natural science over the period of a year. Data were gathered on their productive and receptive skills by using a battery of instruments (oral test, listening comprehension test, a dictation, a cloze test, and a written composition in both their L1 and in English). While this is still an ongoing project, part of the data has already been analyzed in Vallbona (2009), focussing on 5th and 6th Grade levels. According to the results, despite

---

[11] We would like to express our profound sorrow for Mia Victori's sudden decease over the course of the past year.

the fact that the CLIL students only received a total of 35 hours a year of additional exposure to English, significant differences were identified in some of the variables analyzed and in all cases 6th Grade students showed better skills at dictation, and 5th Grade students outperformed in lexical complexity, fluency and accuracy when their written skills were compared. The questionnaires revealed a generally enthusiastic assessment of the CLIL experience by students and teachers alike. However, a number of problems and challenges for the teachers were also identified, most of which related to a lack of training, inadequate language proficiency, and lack of time and resources.

These authors state that CLIL learners outperform their counterparts non-CLIL learners:

> In other words, when learners who had taken a CLIL course in addition to EFL classes were compared with their peers from the same grade, it was the group of students who received a greater amount of instruction in English coming from CLIL classes that significantly outperformed their peers. An interesting question to ask, then, was how far can CLIL learners go? If CLIL learners do better than their peers, is it possible that CLIL learners can manage to be as good as older learners one, two or three grades ahead? (Navés and Victori, 2010: 39).

In this vein, the two studies presented below were intended to throw some light on such a question. They are a follow-up to previous research undertaken as part of the BAF project at the University of Barcelona (see Muñoz, 2006), whose main goal was to examine the effects of onset age in the acquisition of English as a foreign language.

These two studies compare CLIL and non-CLIL learners from different Catalan schools and grades: For the first one, which focused on the subjects' general language proficiency, a total of 837 students from 5th, 7th, 8th and 9th CLIL (n=218) and EFL (n=619) Grades were compared. For the second study, which analyzed learners' writing skills, a total of 695 students from Grades 5 to 12 were compared (CLIL n=200, EFL n=495), including those students who had in addition completed the timed-composition task. The instruments used in both studies were a subset of the battery of tests developed and used by the BAF project.

To analyze the subjects' overall language proficiency (study I), a 25-item multiple choice listening test, a 30-item cloze, a standardised 50-item multiple choice grammar test and a 50 word dictation in English were used. In the second study, a timed written composition was administered, requesting learners to write about themselves, their past and future expectations. This was analysed for accuracy, fluency, and syntactic and lexical complexity, following the work of Wolfe-Quintero *et al*. (1998).

As for the results, it is clearly shown in both studies that CLIL learners outperformed their non-CLIL counterparts in most of the tests administered. As explained above, researchers were particularly interested in exploring whether CLIL learners from lower grades might do as well as non-CLIL learners some grades ahead, which grades seemed to benefit most and in which domains. They conclude:

> In the four overall proficiency tests administered it was found that 7th Grade CLIL learners obtained results similar to those of non-CLIL learners one and two grades ahead for each of the measures analyzed: dictation, reading comprehension, grammar and listening skills. In addition, Grade 8 CLIL learners not only did as well as but actually outperformed non-CLIL learners one grade ahead (9th Grade) in each of the tests administered. (Navés and Victori, 2010: 47).

In the second study, CLIL learners' writing at lower grades was observed to be as good as or even better than that of older learners a few grades ahead. CLIL learners from Grade 5 wrote as much as non-CLIL learners two grades ahead (Grade 7). Interestingly enough, the oldest learners from Grades 11 and 12 did not write longer compositions than CLIL learners at Grades 7, 9 and 10. CLIL 7th Grade learners also obtained results comparable to those of non-CLIL learners at Grade 9 in syntactic and lexical complexity. CLIL Grade 9 learners, on the other hand, wrote more complex sentences than 11th Grade non-CLIL learners, proved to be as fluent as 12th Grade non-CLIL learners, and were among the most accurate of all, along with 11 and 12th Grade students, but did not obtain better results than older non-CLIL learners in lexical variety.

From both studies it can be concluded that when learners are at Grades 7 and 9 and have received CLIL instruction they achieve a level

equivalent to or even higher than learners a couple of grades ahead in several domains of language. These promising results have, nevertheless, to be analysed with caution because as noted before, "the amount of hours of instruction of the CLIL groups was not kept constant, because of the different types of schools involved and hours of instruction received" (Navés and Victori, 2010: 47).

Another type of research studies has mainly focused on the opinions expressed by students and teachers after having participated in CLIL programmes. Feixas *et al.* (2009) examined the perceptions of both university teachers and students. Despite the benefits ascribed to the implementation of these courses, especially in terms of proficiency gains and positive motivation, the authors also noted a number of challenges and difficulties that should be seriously taken into account, among them, students' lack of sufficient proficiency and a capacity to express curricular contents in a second language.

A great many of the studies reviewed center on the analysis of students' oral and written output resulting from class interactions or written production, respectively. This is also the case of the GREIP research group at the University Autonoma in Barcelona where researchers (Escobar and Nussbaum, 2008; Evnitskaya and Aceros, 2008; Escobar and Sánchez, 2009, among others) mainly explored these aspects related to CLIL: interactive tasks in a CLIL class, in which Science content was taught through the medium of English, and CLIL teacher training. These topics were analysed in an interrelated manner since the studies were always conducted in the occasion of pre- or in-service CLIL teacher training programmes. Escobar and Nussbaum (2008) study focused on the role and characteristics of interactive tasks in a CLIL Science class, by students in their 3rd year of secondary education, having different learning profiles and proficiency levels. The aim of the study was to assess task versatility. The analysis of conversational data allowed the authors to conclude that learners interpreted and tackled the task in unexpected ways, using diverse and complex collaborative strategies until they succeeded in turning the initially common task into a tailor-made one, which matched their linguistic and communicative needs, as well as their personal interests. In other words, results showed that students reinterpreted the tasks given and displayed a variety of collaborative strategies to adapt to particular communicative and learning needs.

In a subsequent publication, Escobar and Sánchez (2009) presented the overall results of the study about language learning in inclusive Content and Language Integrated Learning (CLIL) classrooms in secondary school. The pedagogical experience lasted four weeks. The main pedagogical approach followed was the use of learning tasks carried out by students working in dyads. Through the use of different indicators extensively employed in the field, the study measured the gains that students showed in fluency and lexical repertoire in a pretest / treatment / post-test research design. The post-test demonstrated significant progress in the assessed indicators. The study also showed that improvements transcended purely formal linguistic aspects and that it was necessary to find other measurement tools which may help to capture the extent of those improvements.

Still at the University Autonoma, Evnitskaya and Aceros (2008) tackled interactions in a CLIL classroom from a conversational perspective. This study is also based in the same didactic sequence of communicative tasks created by the research group CLIL-SI/GREIP and used in the previous two studies presented above. In this particular case, researchers examined a particular type of a relation, called a didactic contract, established between students while they carried out in-class communicative activities. More specifically, they observed two aspects of the contract expected to take place between pupils who perform pair work activities. First, it was its characterization (appearing, establishing and functioning conditions). Second, it was its role in students' foreign language learning. The analysis showed the didactic contract between peers to be a flexible one and it revealed the temporary phenomenona that can favor target language learning.

To finish with CLIL research in the Catalan language area, regarding studies on the effects of multilingual programmes in the Valencian community, a great bulk of research has been conducted within the framework of the ELA research group in Jaume I University in Castelló. Alcón and Safont (2001) are representative of pioneer work in Spain in the domain of pragmatics and multilingual language acquisition. More specifically, Safont (2005) approached trilingualism, with the combination of Spanish, Catalan and English as its focus of analysis. The book aims at providing a bridge between two applied linguistics subfields, namely those of interlanguage pragmatics and third language acquisition. It examines the production and identification of request acts formulas on the part of bilingual learners of English in the Valencian Community.

2.3.3 Summary

The third part of chapter 2, CLIL Contexts, has dealt with the context in which a content and language integrated view of education appears specially in Europe: the new multilingual strategy put forward by European institutions, including mobility programmes, new views concerning the role of learner autonomy in language learning, and new technological developments, which are a step forward from communicative views of language teaching and education in general. Having previous experiences in Canada and North-America in the background, it has identified the different dimensions of the CLIL approach, a social European dimension, a pedagogical and content dimension and a language acquisition dimension, as well as the modalities and features of integrated programmes. After this, it has briefly focused on its expected benefits, claiming for a need for more empirical research. The challenges teachers and institutions face when trying to implement CLIL approaches to education have also been identified.

Finally, in the second part of the chapter, research on the implementation of immersion programmes in Canada first and then research on CLIL in Europe with a special focus on Spain and Catalonia has been presented.

As the previous sections have shown, research activities in the area of CLIL have gained momentum over the last three or four years producing the first interesting results. Studies on learning outcomes are beginning to show which areas of foreign language competence are most likely to profit from CLIL instruction and which seem to do less so.

In this line of research, we have seen that what emerges from studies in Canada is that in comparison to non-immersion students, immersion students develop (a) almost nativelike comprehension skills as measured by tests of listening and reading comprehension; and (b) high levels of fluency and confidence in using the second language, while production skills seem to be non-nativelike in terms of grammatical accuracy, lexical variety, and sociolinguistic appropriateness (Lyster, 2007). In sum, immersion students are second language speakers who are relatively fluent and effective communicators, but non-targetlike in terms of grammatical structure and non-idiomatic in their lexical

choices and pragmatic expression –in comparison to native speakers of the same age (see section 2.3.2.1).

European research has substantiated North American and Canadian findings with regards to the absence of negative effects on the L1 and with the observation that CLIL learners possess the same amount of content knowledge as their peers who were taught in the L1, and sometimes they even outperformed them (see sections 2.3.2.2).

Concerning language outcomes, we have seen that general statements on the effect of CLIL on students´ language learning outcomes are very positive. In this sense it is interesting to highlight here the effect which has been observed repeatedly and that Dalton-Puffer stated in this way:

> People with special linguistic gifts reach very good results, even high proficiency, also via normal EFL classes, but CLIL significantly enhances the language skills of the broad group of students whose foreign language talents or interest are average (Dalton-Puffer, 2008: 5).

According to this author, the skills which are favourably affected by CLIL educational approaches are: a) receptive skills; b) vocabulary; c) morphology; d) fluency; e) creativity, risk-taking, fluency, quantity; and f) emotive-affective factors. On the other hand, those aspects which are either unaffected or for which research is inexistent or inconclusive are: a) syntax; b) writing; c) informal/non-technical language; d) pronunciation; and e) pragmatics. Studies on CLIL classrooms are producing evidence which can serve as an explanation for those learning outcomes. One important example in connection with this is the finding that content teaching is conducted almost completely without writing activities, a fact which, according to Dalton-Puffer (2008), stands in direct relation to those outcome studies that find the advantage of CLIL students in writing to be small.

If we compare the findings of Canadian research with the results in Europe we observe that there is a weakness in the learners' productive skills, spoken and written, in grammatical and sociolinguistic competence in both contexts. It is as if there were common perceptions about immersion and CLIL classes as being contexts where linguistic form is focused on significantly less than in EFL lessons. The proposals for change in the two continents are in the direction of suggesting the overall general experiential approach to learning to be balanced with

more analytical approaches, that is, introducing approaches that focus on form in an otherwise communicative classroom (Harley *et al.*, 1990).

It should be noted, however, that while some of the results obtained in the research here discussed coincide with claims made for CLIL at other latitudes in the world, it is still too early to infer any generalised outcomes for European CLIL. As we have explained in section 2.3.2.3 (Spain), in Andalusia an evaluation project related to the implementation of CLIL (among other programmes aimed at fostering multilingualism) confirms that CLIL learners show greater gains than their monolingual peers. CLIL learners were clearly outperforming their mainstream peers in all linguistic skills. The study also suggests that the advantage extends to structural variety and pragmatic efficiency, hence encompassing language growth at lexico-grammatical and discourse levels.

In the Basque Country (see section 2.3.2.3) some research studies also add further evidence to the idea that the CLIL exerts a significant positive effect on students´ language proficency in both receptive and productive skills. Lasagabaster (2008) claims that the review of studies completed by Dalton-Puffer showed CLIL advantages concerning receptive skills, vocabulary and fluency, but their results also show an improvement of other aspects such as writing and pronunciation which seem to remain unaffected in the German context and so do the Department of Education, University and Investigation (DEUI) of the Basque Government. Therefore, in the Basque country the advantages seem to be even more obvious than those revealed in studies undertaken in Germany (Lasagabaser, 2008: 39).

Also in Catalonia (see section 2.3.2.4) it was found that CLIL students outperformed non-CLIL learners in lexical complexity, fluency and accuracy when their written skills were compared.

In this vein of research, it will be very interesting to see wether the current research study may help to shed some light on these different results. The learning of foreign languages through the CLIL approach is an attractive challenge, and this is especially so in bilingual contexts in which two co-official languages (Catalan and Spanish) coexist such as the one presented in the following chapters.

In conclusion, we have seen contradictory results in some domains of linguistic competence and an important question remains to be answered: Will CLIL programmes remain attractive and generally supported by all the stakeholders once they become the norm in the future? Time only will tell…

A very important realization arising from classroom studies is that despite the differences between EFL lessons and CLIL lessons, they both are specimens of educational interaction, conditioned by all the factors institutional education involves. As a consequence we need to state in no uncertain terms that not only EFL classrooms are limited language learning environments but so are CLIL classrooms, even though in subtly different ways. Conversely, each of the two offers unique opportunities for students to learn and use the target language to the extent that they are difficult to reproduce in the other. Ideally then, EFL and the language dimension of CLIL ought to be integrated into one foreign language curriculum.

This is precisely the vision presented by Pérez-Vidal (forthcoming) in relation with wether CLIL education only suffices to learn a foreign language. This author argues that with CLIL programmes learners seem to make substantial linguistic progress, however not at all levels of language competence.

> Even if some Spanish CLIL studies seem to show benefits at all levels, others report contrary results, or results which are surprising, such as the fact that CLIL learners are better than their non-CLIL counterparts from the start. This may be due to the fact that CLIL learners are often 'selected'. In addition, CLIL programmes are mostly pilot programmes; hence results obtained from them cannot be extrapolated to other mainstream situations. In sum, CLIL programmes seem to improve learners' receptive skills, their vocabulary and their creativity by and large.
>
> SA (Stay Abroad) is another 'specific' context of learning which has been recently scrutinized in Europe, and even more recently in Spain. Research results from learners enjoying a sojourn in the target language country reveal that the greatest benefits obtained are at the level of oral fluency, listening, pragmatics and writing. These domains are complementary with the domains in which CLIL seemed to offer the greatest benefits.

This seems to indicate that our learners should have access to both CLIL programmes and SA programmes if we want to uphold the objective of multilingual education. A combination of these two specific contexts of learning, in addition to the conventional FI context should help them develop as young multilinguals able to communicate in an increasingly changing international society in which knowledge of languages and positive attitudes toward multiculturalism are an important requisite for success (Pérez-Vidal, forthcoming: 21).

In sum, in the light of the results of research on SA and in particular the work undertaken within the Barcelona SALA project, Pérez-Vidal (forthcoming) argues that other contexts of learning in addition to Formal Instruction (FI), such as Study Abroad (SA) might be the necessary complement to CLIL.

Finally, another important idea is the one put forward by Dalton-Puffer (2007) about the fact that the learning goals which are formulated for CLIL tend to be unspecific at present. This author argues that in order to avoid stagnation of the CLIL enterprise it will be necessary in the future to state more explicitly which language learning aims are pursued through the practice of CLIL (and by implication, therefore, which are not or cannot be pursued but must be taken care of by EFL lessons or altogether different learning environments such as study abroad as argued above). Once these more concrete language learning goals come clearly into view, it should be easier for CLIL teachers to align their didactic/pedagogical decisions about teaching the content in such a way that their classrooms can be content- and language-rich. Very often what is good for language (such as having to actively encode new concepts for a specific audience) is also good for content.

On the basis of the research findings presented in this chapter, the current study seeks to contribute with further empirical data to a better understanding of the impact of CLIL programmes on EFL learners' linguistic progress.

II. THE STUDY

3. RESEARCH QUESTIONS AND METHOD

In this chapter the research questions and the method used to carry out the present research study are described. The organisation of the chapter is as follows: Section 3.1 presents the objectives of the present investigation. In section 3.2 the research questions and the hypotheses used to address this research are described. After this, a section with the method is offered. In it the context and participants of the study, the design, treatment, and instruments used, and the data collection procedure are explained in detail. The last part within the method section (3.3) is a large description of the quantitative and qualitative measures used to analyse the data.

3.1 Objectives

In the study presented here the effect of a CLIL programme on EFL linguistic progress is examined. More specifically, in order to do that, the study measures the differential linguistic gains achieved by two groups of participants. The first group, (Group A), follows formal instruction and in parallel CLIL instruction, (FI + CLIL). That is the group receives some 'extra' hours which are CLIL hours. The second group, (Group B), follows a FI only programme. Likewise, the study focuses on the degree of inf1uence of individual factors such as age, or gender on the level of competence attained in each different context.

3.2 Research Question and Hypotheses

The present dissertation will try to answer the following research question:

Research Question 1:

RQ1: How does context of learning affect the linguistic development of young bilingual secondary education EFL leaners when contrasting a group experiencing FI only and a group experiencing FI in combination with CLIL? Namely,

RQ1a) When contrasting the differential effects of the two different programmes, a FI only and a FI+CLIL, that is with an additional CLIL component, which programme results in linguistic benefits if any and which skills benefit the most if any?

RQ1b) Does age have an impact on the potential benefits learners may achieve in either of the two contexts compared: FI and FI+CLIL?

RQ1c) Does gender have an impact in the benefits obtained with either a FI context of learning, or a FI+CLIL context?

Hypotheses:

H1:   When contrasting the linguistic development of two groups of bilingual secondary education EFL learners experiencing FI only and FI in combination with CLIL respectively, the CLIL context of learning will affect in several different domains of language competence and forms.

H1a)   When contrasting the differential effects on learners' linguistic progress of two programmes a FI programme, and a FI+CLIL with the additional hours, the group in the FI+CLIL will improve significantly more than the other especially in receptive skills.

H1b)   At different ages and with a similar number of hours, younger learners receiving FI+CLIL will benefit more than elder learners only receiving FI because 'the earlier the better' for CLIL instruction. On the other hand, at the same age and with a different number of hours, learners with a higher total number of hours receiving FI+CLIL will benefit more than learners with fewer hours only receiving FI.

H1c)   The female participants in the FI+CLIL group will not significantly outperform their male peers in the linguistic progress achieved after a one year CLIL treatment as they reportedly do in FI.

3.3 Method

3.3.1 Context and Participants

School Context

This study is conducted within the Catalan/Spanish bilingual educational context. The setting in which the subjects are immersed has

been defined as additive trilingualism (Cenoz and Valencia, 1994). Catalan, the language of instruction, together with Spanish are the majority languages, and English is taught as a foreign language. Learners are also studying a second foreign language in the school at later ages.

Data from Catalan/Spanish bilingual learners of English as an L3 in two different types of exposure contexts have been examined: FI (formal instruction English as a foreign language) and CLIL (content and language integrated learning). One group experiencing FI and CLIL and another one experiencing FI only are analysed. They have been selected because they respectively represent the first group undertaking a CLIL programme (in addition to FI) and the last group only undertaking FI before CLIL started in a state-run (ordinary governtment-supported) school in Barcelona (Spain) offering courses from infant school to post-compulsory secondary education. In the year 2001, the school considered the possibility of introducing a third language, English, as the medium of instruction in the school curriculum.

In fact, the school council of this school decided to adopt a CLIL approach after carefully taking into consideration and evaluating the different sides of such a kind of project. They saught to guarantee adequate exit levels in English as a foreign language and a good preparation for a university degree where languages are an asset. As a consequence of the programme they were also going to achieve two further objectives, namely, enhancing the school's profile, and motivating and promoting its teaching staff.

The model adopted by the school in order to bring about change had various facets. It first sought assessment from educational leadership for general planning. It then identified two sets of agents, the families and the teachers. They would be treated separately in line with their roles in the process, in the understanding that during the preparatory period the learners would be assigned no role, as they were left the responsibility of coping with change once it would be underway. Finally, it organised the adequate course of action in order to prepare for the implementation of CLIL in the school. Given that of the agents involved in bringing about change, teachers were going to be the key to the success of the programme, a thirty-hour inset teacher training programme was devised. The programme offered a rationale for CLIL

instruction and revised teaching techniques and strategies that would help teachers design and implement adapted syllabi for Science and English. As a result, the roles of teachers in the CLIL project was redefined as table 11 shows:

Table 11. The roles of teachers in the CLIL project

I.	Former role	CLIL new role
Primary class teachers (Generalist teachers)	• In charge of nearly all subjects in the curriculum, including Science. • Responsible for the learners' welfare.	• No longer teaching Science, but advisors to new Science teachers. • Responsible for the learners' welfare in the CLIL context.
Primary English teachers (Teachers with an EFL component in their Degree)	• English teachers in primary.	• English and Science teachers.
Secondary English teachers (Highly specialised English teachers)	• Advisors to primary English teachers.	• Advisors to primary English-Science teachers.
Secondary Science teachers (Highly specialised Science teachers)	• Advisors to primary class teachers.	• Supervisors of the standards.

Source: Teacher education for the implementation of CLIL in the school system (Pérez-Vidal and Escobar, 2004)

The programme also catered for other aspects, such us preconceptions and self-regulatory abilities. This was facilitated by providing room for open discussion, but, most importantly, by leaving aside prescriptive approaches when presenting concepts and techniques. Finally, the programme acted as a forum for discussion –why do we do things this way?- and promoted a salutary revision of educational practices within the school. This proves that innovative programmes such as the introduction of multilingual education provide a priceless opportunity for bringing about educational change and help teams of teachers rethink their teaching practices (Pérez-Vidal and Escobar, 2004).

As regards the design of the implementation of the CLIL programme, it was decided that it would start at grade 3 and grade 5 (8 and 10 year olds respectively) with CLIL lessons in Science. The nursery school (3 year olds) was also to initiate a programme of so-called 'language showers', of which no further mention will be made in this dissertation for reasons of space.

In sum, the school designed an extremely robust programme in order to ensure its success. In appendix 2 the programme's chart and its progressive implementation is presented (see appendix 2).

Participants

For the purpose of this study the linguistic production of 100 Catalan/Spanish bilingual EFL learners in the school just described was analysed. Participants were following their secondary education at Grades 7, 8 and 9 (that is at their first, second and third year of *ESO*, if we use the Spanish terms). They were at ages 12, 13 and 14 respectively. In order to proceed to analyse the differential effects of the variable 'learning context', two sample groups were identified, each group representing a different learning context under examination. Group A (n= 50) was the experimental group experiencing a FICLIL (FI + CLIL) context, and Group B was the control group (n= 50) experiencing a FI context, the 'control' context. There were a 50% of males and females in both groups.

As table 12 summarises, Group A, the experimental group, had received conventional formal instruction in the foreign language classroom, and, in addition, being part of the school's CLIL programme, they had studied Science with English as its medium of instruction two hours per week since Grade 5 (10 years old). The group included learners who were doing Grade 7 and 8 and hence were 12 and 13 years old at the beginning of each academic year. On the other hand, Group B, the control group, acquired English following conventional formal instruction in the foreign language classroom. The group included learners doing Grade 8 and 9 who were 13 and 14 years old at the beginning of each academic year. Having been placed together in the same school since nursery, they had all started learning English at the age of 6 (Grade 1), following the official curriculum in Catalonia, so both groups shared a common age of onset of exposure to English as their L3. So by the time data collection started they had had

6 years of FI. The experimental group had started the CLIL programme at age 10. So by the time data collection started they had had 3 years of CLIL instruction. The majority of these learners were reportedly following extra-curricular classes.

Table 12. Participants (N=50)

Onset Age	T1 (2005)	T2 (2006)
GROUP A: FI + **CLIL** FI: 5 yrs**.** **CLIL: 10 yrs.**	Grade7 /1st ESO (12 yrs.)	Grade8/2nd ESO (13 yrs.)
GROUP B: FI FI: 5 yrs.	Grade8/2nd ESO (13 yrs.)	Grade9/3rd ESO (14 yrs.)

Source: personal

3.3.2 Design

The design of the study was longitudinal. The results obtained along two consecutive academic years (2004-2005, 2005-2006) were analysed. Table 13 below shows its longitudinal pre-test, post-test design. Group A and Group B learners were measured respectively before and after one academic year in order to tap on gains obtained in each context over the course of a year. Then the difference in gains obtained by each group was calculated.

Data collection took place two years after the start of the CLIL programme in the school. As explained above, Group A had started CLIL at Grade 5 (10 years old), so it had had three years of CLIL, by the time of the first data collection, that is 210 hours (70/year). As table 13 displays, Group A had had an accumulated number of 1,120 hours of FI (aproximately 140/year) and, in addition, 210 of CLIL instruction at T1 (Grade 7, 12 years old), and, one year later, at T2, 1,260 hours of FI and 280 hours of CLIL (Grade 8, 13 years old). Group A gains between T1 and T2 are compared with gains by Group B, the control, with an accumulated number of 1,260 hours of FI at T1 (Grade 8, 13 years old) and 1,400 at T2 (Grade 9,14 years old). **The design allows for a between-groups comparison of the effect of 210 hours (140 FI + 70 CLIL) in Group A versus 140 (FI) in Group B.** Hence, the difference in the gains obtained by each group over a year treatment.

Table 13a[12]. Design

	T1 (2005)	**T2 (2006)**
A: FI + **CLIL**	Grade7 / 1st ESO (12 yrs.) FI: 1120 h + **CLIL: 210h** = 1330	Grade8 / 2nd ESO (13 yrs.) FI: 1260 h + **CLIL: 280h** = 1540 (+ 210h)
B: FI	Grade8 / 2nd ESO (13 yrs.) FI: 1260 h CLIL: 0h	Grade9 / 3rd ESO (14 yrs.) FI: 1400 h CLIL: 0h (+ 140h)

Source: personal

The independent variable in the study is contact with English in two learners' populations: one with FI, and one with FI+CLIL. On the other hand, the dependent variables are gains in skills / levels: written production, written comprehension, oral comprehension, and lexico-grammatical ability. Finally, since the age and gender of the sample can be controlled, the impact that these two variables have on skill development in relation to level will be measured, thus becoming the control variables.

In order to address the age hypothesis another between-groups comparison will be established between Group A at T1 and Group B at T2 when they share a similar number of hours at different ages (1330 h at 12 years old versus 1400 h at 14 years old). Finally, in order to address the gender hypothesis a within-group comparison will be established between the subsample of female versus male participants in each group.

3.3.3 Treatment

Both the CLIL and the Formal Instruction treatment were characterised by a methodology that follows a communicative approach. Communicative forms of instruction focus on meaning and communication, which is established by genuine interaction between the teachers and the learners generally through pair/group work interaction, a creative non-restrictive use of language, and via opportunities for the negotiation of task topics (not to be more detailed here for reasons of space).

---

[12] This table is reproduced in section 4.2 in order to help readers to follow the presentation of results.

Although both treatments shared this communicative approach, the Formal Instruction treatment was characterised by the fact that attention was very often drawn towards language forms and grammar in order to develop linguistic awareness. That is, a more focus-on-form view typical of conventional EFL classes. On the other hand, in the CLIL treatment, grammar was hardly ever studied, and it was reading and specially speaking abilities that were highlighted and practised most of the time. As explained before, teachers were specially trained with a robust tailor-made course in order to ensure success. Finally, an adapted Science book from a well-known publisher was adopted in order to ensure the teaching of all contents in the curriculum for the specific academic years[13].

More specifically, each of the two contexts included in their syllabus practice with particular patterns of language skills. Table 14 below displays the skills analysed in the present research study. The boxes of the column in the middle and to the right state the amount of practice learners experience in the corresponding skill (left-hand side column) according to each context of learning (CLIL or FI). On the one hand, in our FI context the writing and reading skills are very often practiced, at least once a week. Listening is practiced in a limited way, through teacher talk, and lexico-grammatical abilities are very often practiced, at least once a week and very often in every single session. Oral production is limited. On the other hand, in the CLIL context, whereas reading is highly practiced in every class session and with an amount of authentic texts which is unusual in FI, listening and writing abilities are limited to teacher talk and very short exercises respectively. In addition, lexico-grammatical abilities are hardly ever practiced[14].

---

[13] This information about the methodology used in relation with these two groups treatment was collected through a personal communication with the FL teachers and the FL coordinator in a systematic manner.

[14] Oral production, although not contemplated in our study, is also important. It takes place to a higher degree than in FI classes as learners always use the target language to interact with the teacher. Peer work, however, seldom involves it. In contrast, one might say that in FI peer work is controlled for the language being used when carrying out tasks, but teacher to learners conversations may show high amounts of code-switching from English to Catalan/Spanish.

Table 14: Skill Practice

SKILL	CLIL Context	FI Context
Writing	**Seldom practiced** (short exercises)	**Often practiced** (at least once a week)
Reading	**Highly practiced** (every class session)	**Often practiced** (at least once a week)
Listening	**Seldom practiced** (teacher talk)	**Seldom practiced** (teacher talk)
Grammar (Lexico-grammatical ability)	**Very seldom practiced** (once a month)	**Very often practiced** (at least once a week)

Source: personal

3.3.4 Instruments and Data Collection Procedure

Data were elicited both for productive and receptive skills. Production was elicited in writing, and reception both in writing and orally. In addition, lexico-grammatical abilities were also tapped on.

As can be seen in table 15, the instruments used to obtain the data were the following:

In order to gauge production, students were administered a written task in which they were asked to write a composition on a given topic. In order to analyse comprehension, a reading task (cloze) and a dictation were administered. Finally, lexico-grammatical ability was measured by means of sentences with progressive difficulty tasks with a multiple choice format (see appendix 1). Learners were given two hours to complete the whole battery of tests (with some free time after each test), and the order and time allotted to each activity was as follows:

Table 15. Instruments

Production	• Written ability	- Composition (dialogue and two very short narratives)
Comprehension	• Reading ability	- Cloze (texts with progressive difficulty)
	• Oral ability	- Dictation
Lexico-grammatical ability	• Grammar test	- 30 sentences with progressive difficulty in a multiple choice format

	• Grammaticality judgement test	- 20 sentences with progressive difficulty.

Source: personal

First hour (lexico-grammatical ability and written production skills):

A. Grammar (sentences with progressive difficulty): 15 minutes
B. Grammaticality judgement test (sentences with progressive difficulty): 10 minutes
C. Writing (composition): 20 minutes

Second hour (comprehension skills):

D. Reading (cloze): 15 minutes
E. Dictation: 15 minutes

3.3.4.1 Production (written ability)

In order to gauge production, students were administered a writing activity whereby they had to write a dialogue on the basis of a picture (see appendix 1C). Participants had 20 minutes to complete the task. It showed two policemen, a mother and a boy at the entrance door of a home. Learners were shown the picture and then they had 20 lines to answer the following two questions (10 lines per question):

Why did this happen?
How do you think the situation will end?

These picture and questions were chosen because it was thought that the young boy in the picture would allow for a process of identification which should be inspiring (Foster & Skehan, 1996; Tavakoli & Foster, 2008). In addition, since they were asked to write a dialogue and the answer to the two questions (two very short narratives), different genres had to be used. Finally, the choice of a composition is based on the number of subskills that come to play when learners write a piece of text and it is also a task that is practised in the classroom.

3.3.4.2 Comprehension

- Reading skills

In order to collect the data for the measurement of gains in comprehension skills, learners were asked to read short texts with progressive difficulty in which blanks were left following a cloze procedure (see appendix 1D). Participants had fifteen minutes to complete the cloze. These texts dealt with Tsunamis, a topic already studied in the Science subject and therefore familiar to the learners as far as their vocabulary and structures. This activity not only measured comprehension skills but also lexico-grammatical abilities since some of the blanks had syntactic and lexical items, something which will be taken into account in our discussion of results.

In its original form the cloze procedure reduces redundancy by deleting a number of words in a passage, leaving blanks, and requiring the person taking the test to attempt to replace the original words. After a short unmutilated 'lead-in', it is usually between every 5th and 11th word that is deleted (Hughes, 2003).

One of the reasons to have chosen a cloze format to test reading comprehension is that in more open-ended formats, e.g., short answer questions, the student has to deploy the skill of writing. The extent to which this affects accurate measurement of the trait being assessed has not being established. Cloze tests, as multiple-choice tests, avoid this particular difficulty (Weir, 1998). However, if a comparison between cloze and multiple-choice is carried out, it has been established that the two techniques measure different aspects of the reading activity – namely that a timed cloze measures <u>the process</u> of reading, i.e., the reader's ability to understand the text while she/he is actually reading it; multiple choice, on the other hand, measures <u>the product</u> of reading, namely the reader's ability to interpret the abstracted information for its meaning value. There is a good deal of supportive evidence in the literature for using the cloze format (Weir, 1998):

> (…)The reader comprehends the mutilated sentence as a whole and completes the pattern (…) This procedure came to be highly regarded as a measure of testing reading comprehension and even as a measure of overall language proficiency. (…) Cloze tests are valid and uniform measures of reading comprehension

abilities (…) Cloze tests measure the reader's ability to decode interrupted or mutilated messages by making the most acceptable substitutions from all the contextual clues available (Weir, 1998: 50).

A cloze procedure is considered an integrative method because it draws on the student's ability to process lengthy passages of language: in order to replace the missing word, candidates have to make use of the abilities that underlay all their language performance. Therefore it provides a measure of those underlying abilities, its content validity deriving from the fact that the deletion of every *nth* word meant that a representative sample of the linguistic features of the text was obtained. Support for this view also came in the form of relatively high correlations between scores on cloze passages and total scores on much longer, more complex tests, such as the University of California Los Angeles (UCLA) English as a Second Language Placement Test (ESLPE), as well as the individual components of such tests (such as reading and listening). In sum, the cloze procedure is very attractive as a measure of reading and overall ability and reports of early research seemed to suggest that the result would be a reliable and valid test of candidates' underlying language abilities. In the literature, cloze tests are often considered valid and uniform measures of reading comprehension. In addition, cloze tests are quite easy to construct, administer and score (Hughes, 2003: 189).

- Oral skills

A dictation was administered (see appendix 1E) to measure oral abilities/skills. It consisted of a fifteen minute dictation dealing with the Antarctica. Again it was a familiar topic related to the Science subject.

It was chosen because it was meant to provide a 'rough and ready' measure of listening ability. Certainly dictations are also very easy to create. In the 1960s it was usual, at least in some parts of the world, to decry dictation testing as hopelessly misguided.

> After all, since the order of words was given, it did not test word order; since the words themselves were given, it did not test vocabulary; since it was possible to identify words from the context, it did not test aural perception. While it might test punctuation and spelling, there were clearly more economical ways of doing this (Hughes, 2003: 195).

Some decades later this orthodoxy was challenged. Research revealed high correlations between scores on dictation tests and scores on much longer and more complex tests (such as again the University of California Los Angeles -UCLA- English as a Second Language Placement Examination -ESLPE). Examination of performance on dictation tests made it clear that words and word order were not really given; the candidate heard only a stream of sounds which had to be decoded into a succession of words, stored, and recreated on paper. The ability to identify words from context was now seen as a very desirable ability, one that distinguished between learners at different levels.

In addition as being easy to create, dictations are also relatively easy to administer. Another reason to choose a dictation to measure listening ability was the difficulty of focusing on specific listening points whilst students are exposed to ongoing discourse, a problem which dictations solve. According to Weir (1998) it is often adviseable to include a more discrete format than listing when testing oral comprehension with the possibility this gives of including a greater number of specific items. A dictation can provide this discreteness as well as being valid in content terms for certain groups of candidates. Furthermore, a dictation includes a format that provides reliability through the large number of items that can be generated. Like cloze, dictations are often employed as a useful measure of general proficiency because as a testing device they measure many different language features to be effective in providing a means of assessing any particular skill. Oller (1979) argued that a dictation was an adequate test of listening comprehension because it tested a broad range of integrative skills. This author claimed that a dynamic process of analysis by synthesis was involved. Dictation draws on the learner's ability to use all the systems of the language in conjunction with knowledge of the world, context, etc., to predict what will be said (synthesis of message) and after the message has been uttered to scrutinise this via the short term memory in order to see if it fits with what had been predicted (analysis).

Dictation for Oller tests not only a student's ability to discriminate phonological units but also his/her ability to make decisions about word boundaries; in this way a testee discovers sequences of words and phrases that make sense and from these he/she reconstructs a message. The identification of words from context explained before as well as from perceived sounds is seen by Oller as a positive advantage of

dictation in that this ability is crucial in the functioning of language. The success with which the candidate reconstructs the message is said to depend on the degree to which his internalised 'expectancy grammar' replicates that of the native speaker. Fluent native speakers nearly always score 100% on a well administered dictation while non native learners make errors of omission, insertion, word order, inversion, etc., indicating that their internalised grammars are, to some extent, inaccurate and incomplete; they do not fully understand what they hear and what they reencode is correspondingly different from the original (Weir, 1998: 60).

Participants were not informed about any of the data collections, this way their performance was not affected by prior preparation or rehearsal of the tests they had to undergo. The tests were administered by their teachers in a conventional classroom session in an exam-like situation.

3.3.4.3 Lexico-grammatical ability

As can be seen in appendix 1A/B, to obtain data on lexico-grammatical ability, participants were administered a grammar test which included 30 questions with progressive difficulty, in multiple choice/cloze format. Learners had to select the answer to each question from a number of given options, only one of which was correct,

EXAMPLE: *Sally often _____ to do her homework - forgetting / is forgetting / forgets*).

In addition, participants were also shown a list of 20 sentences whose correction had to be judged:

	Correct	Incorrect

EXAMPLE: *She play the piano very well.*
She's fourteen years old.

They were given 25 minutes to complete these two parts.

3.3.5 Analysis / Measures (analytic qualitative, quantitative)

Different procedures were used for the analysis of the data gathered through the instruments described in the previous sections. On the one hand, the reading task, the dictation, the grammar and grammaticality

judgement tests were straightforward marked following objective criteria. A correcting matrix was used with the right answers.

However, dictations are certainly not easy to score. Marking may well be problematic if one wishes to take into account seriousness of error. Oller (1979, 1980), who was a leading researcher into both cloze and dictation, recommends that the score should be the number of words appearing in their original sequence (mispelled words being regarded as correct as long as no phonological rule is broken). According to Oller (1993), this works quite well when performance is reasonably accurate, but is still time-consuming. With poorer students, scoring becomes very tedious.

In the present research study performance in the dictations was quite accurate. Following Oller (1993), correct spelling was not required for a response to be scored as correct since it was listening that was meant to be tested. However, it was not enough for our participants to attempt a representation of the sounds that they heard, without making sense of those sounds. To be scored as correct, a word had to provide strong evidence of the learner's having heard and recognised the missing word, even if he/she could not spell it. It has to be admitted that this caused some scoring problems (see section 5.1).

On the other hand, the written task was corrected on the basis of standard objective and subjective procedures further explained below.

Then the frequency figure accounting for correct/incorrect items was calculated per task. A final figure representing a general score was thus obtained for each task in order to calculate linguistic progress for each specific subdomain of competence analysed: written ability, reading ability, oral perception ability, and grammatical ability.

As can be seen in table 16, the data obtained from the writing test is analysed quantitatively following an adapted matrix which profiles lexical and syntactic complexity, fluency and accuracy features (Wolfe-Quintero *et al*., 1998). The data is also analysed qualitatively following a rating scale (Friedl & Auer, 2007) whereby task fulfilment, organisation, grammar and vocabulary aspects are measured (see appendix 3).

Table 16. Measures used to analyse written development

Quantitative measures:	Syntactic complexity	Lexical complexity	Accuracy	Fluency
Qualitative measures:	Task fulfilment	Organisation	Grammar	Vocabulary

Source: personal

In the following lines the reasons for choosing both a quantitative and a qualitative analysis are explained. Written production can be analyzed according to two different approaches: analytic (quantitative research), or holistic (qualitative research).

The analytic approach uses objective measures (Casanave, 1994). This means that certain quantitative measures are adopted to give account of the participants' written proficiency which do not depend on the researchers' opinion.

The holistic approach is the most common way of measuring written proficiency among EFL teachers and is done by means of scales. This procedure is thought to be subjective because written compositions are given a global mark usually according to their content, organisation and correctness, and the score depends, to some extent, on the corrector.

Dichotomies such as the ones displayed in table 17 below show how, often, in applied linguistics qualitative research is contrasted with quantitative research:

Table 17. Qualitative research versus quantitative research

Qualitative research	**Quantitative research**
naturalistic	Controlled
observational	experimental
subjective	objective
descriptive	inferential
process-orientated	outcome-orientated
valid	reliable
holistic	particularistic
'real', 'rich', 'deep' data	hard, replicable data
ungeneralisable single case analysis	generalisable aggregate analysis

Source: An introduction to Second Language Acquisition research. (Larsen-Freeman and Long, 1991: 12)

Since most of the research that has focused on written production has taken the analytic approach (Celaya *et al*., 2001) and analytic measures have proved to be reliable, this approach was first chosen and adopted in the present study to examine the participants' complexity, accuracy, and fluency in writing. Analytic measures were thought to be highly reliable in order to operationalize the aforementioned areas. However, as the analysis was being carried out, we realised that counting the number of mistakes or the presence or absence of certain elements, as quantitative measures do, did not account for some features of our compositions (specially in the dialogues), particularly those showing communicative effectiveness. In this last case, it was impossible to measure the appropriateness of language used for the expression of functional meaning using quantitative measures. We also felt that we needed more descriptive information in different subdomains of written competence. For example, it was important to measure if there was a good choice of words, a clear overall structure, if the task was fully achieved with a variety of ideas and arguments and the text was entirely relevant to the topic with an adequate register, etc. Therefore, as explained before, a qualitative assessment rating scale to measure linguistic and communicative competence in writing (Friedl & Auer, 2007) was also adopted whereby task fulfilment, organisation, grammar and vocabulary aspects are measured (see appendix 3).

Notwithstanding, in order to avoid one of the most serious weaknesses of holistic methods, subjectivity, following McNamara (2009, personal communication) the evaluation with holistic measures was carried out by two specialists following the same procedure (one of the evaluators being the researcher of the present dissertation): the items assessed (explained in detail in the following lines) were vocabulary, grammar, task fulfilment, and organisation and each of them were sub-divided into 6 behavioural levels on a scale of 0 (not enough to evaluate) to 5 (very good) obtaining a qualification from 0 to 20 (see appendix 3). At level 5 it was felt that a student was likely to have no problems in coping with the writing tasks demanded of him/her. At level 4 they were found to have very few problems. At a level 3 and 2 a limited number of problems should appear in relation to the criterion and remedial help would be advisable. A level 1 indicated that quite a fair amount of help was necessary with respect to this particular criterion. A level 0 indicated total incompetence in respect of the criterion in question. Following McNamara (2009) the two specialists started evaluating the compositions together in order to find common criteria to set up the level. Once an agreement had been reached, they

proceeded in with the evaluation separately and they met once a week to check the results. At the beginning it was established that if the difference in the qualifications for one same student was higher than 2 (i.e one evaluator gave a 4 for vocabulary and the other one a 1) the evaluation would have to be repeated. However, it was never the case since results were always the same or with only one point difference. In this last case, the final result was the average value (i.e. if one evaluator gave a 4 for organisation and the other one a 2, the result was a 3).

In the following lines, first qualitative measures used are described. These are followed by the description of the quantitative measures used to analyse written lexical and syntactic complexity, accuracy and fluency

3.3.5.1 Qualitative measures

As explained in the preceding lines, a qualitative assessment rating scale intended to measure linguistic and communicative competence in writing (Friedl & Auer, 2007) was also adopted whereby task fulfilment, organisation, grammar and vocabulary aspects are measured. The items rated within the aforementioned areas are detailed below following Friedl & Auer (2007).

Task fulfilment

Under task fulfilment content and relevance, text format, length and register were measured. In other words, assessment focused on wether the task was fully achieved, if the content was entirely relevant, and wether the length, format and register were appropriate. Another aspect of high importance was the appropriateness of the language used for the expression of functional meaning in order to take into account the communicative effect (specially in the dialogues). Frequency of gaps or redundant information was also important.

Organisation

In this area the structure, paragraphing, cohesion and coherence, and editing and punctuation were considered relevant. That is to say, if the overall structuring was clear or poor, the paragraphing was meaningful or not, wether the use of connectives was limited or good, if students

made frequent editing mistakes, and/or wether conventions of punctuation were observed.

Grammar

Concerning grammar, students' errors and variety of structures, as well as readiness to use complex structures were considered. For example, the maximum score (5) would be given to a participant whose composition showed an accurate use of grammar and structures, hardly any errors of agreement, tense, word order, articles, pronouns, etc., a frequent use of complex structures and great variety of structures.

Vocabulary

Finally, within this area the range and choice of words, spelling, comprehensibility, and accurate form and usage were measured. A student with poor range of vocabulary and choice of words, highly repetitive, with numerous errors of form and usage, numerous spelling mistakes, with a very often not clear meaning, and a lot of translation from mother tongue would receive the lowest score: 1 (a 0 was given when we did not have enough information to evaluate).

3.3.5.2 Quantitative measures

Wolfe-Quintero, Iganaki and Kim (1998) provide a detailed review of the more commonly used analytic measures in writing development. The aim of this review was to find out what measures explained writing development better. As the quantitative measures adopted in this research were based on Wolfe-Quintero *et al.*'s study, this study will be used as a reference.

The measures proposed by Wolfe-Quintero *et al.* (1998) are distributed among frequencies, ratios and indices. All the units under analysis in this investigation are ratios because these relate units to other reference units (number of words per clause, number of errors per T-unit, etc.) and are easy to compute. Consequently, they are believed to be more complex and in turn more reliable, and for this reason they are more common than frequencies in the SLA literature.

Syntactic complexity: coordination index

Wolfe-Quintero *et al.* (1998) provide the readers with a variety of ways of measuring syntactic complexity suggesting the coordination index as an alternative to distinguish between learners who use coordination versus subordination to increase sentence complexity.

Thus, following these authors' recommendation, in the present investigation syntactic complexity has been measured by means of the coordination index, which is calculated by dividing the independent clause coordinations by the number of combined clauses (coordinated clauses + subordinated clauses). Since it is supposed that participants with a higher English level produce more subordinate clauses than coordinate clauses, the higher the coordination index is, the lower the level of ability will be. Therefore, a lower ratio is expected in the post-test.

Lexical complexity: Guiraud's index

One of the ways to measure lexical richness is by calculating the type-token ratio, which is obtained by dividing the total number of types (lexical words) by the total numer of tokens (total number of words). However, this measure has been widely critised due to its sensitivity to length (i.e. it decreases for longer samples where words are more likely to be repeated) and for this reason it was decided not to adopt this measure.

Instead, in the present study written lexical complexity is obtained with the Guiraud's index of lexical richness. It is calculated by dividing the total number of types by the square root of the total number of tokens, as it has proved to be an efficient measure to counteract the effect of text length. A higher value of this measure in the post-test will indicate that subjects' vocabulary is richer.

Accuracy: errors per word

Measuring accuracy can be more problematic than measuring fluency or complexity in that deciding what counts as error or error levels can be liable to subjective interpretation. However, since it is rather difficult to come up with objective criteria to establish error levels, they have not been considered in this study and "error types have been

preferred. We have classified errors as grammatical, lexical and spelling errors.

There is also a variety of ways of measuring accuracy, depending on the researcher's interests. If the goal of the research is to analyze accuracy in a general way, the ratios that have been more commonly used are: error-free T-units per T-unit, error-free T-units per sentence, error-free sentences per sentence. Ratios regarding the specific type of error that the participants have made may be desirable, in which case the following ratios can be calculated: morphological errors per clause, sentence or T-unit; syntactic errors per clause, sentence or T-unit, etc. Another type of ratio that can account per accuracy if the researcher is more focused on correctness is obtained by calculating the correct number of a specific category of words and dividing it by the total number of words (i.e. correct number of adjectives per total number of words, correct number of verbs per total number of words, etc).

The measure adopted to account for accuracy in this study is total number of errors per word. We believe that this measure is a specific measure that takes into account the total number of errors and thus it provides valuable information to account for to what extent a composition is accurate enough. It is calculated by dividing the total number of errors by the total number of words. A lower value is expected in the post-test if an improvement has occurred and this will therefore mean an error reduction.

Fluency: total number of words

Compared to oral fluency, written fluency has not received much attention from SLA researchers. One of the definitions of fluency is provided by Wolfe-Quintero *et al*. (1998), who, in reference to written fluency, state that

> fluency means that more words and more structures are accessed in a limited time, whereas a lack of fluency means that only a few words or structures are accessed (Wolfe-Quintero *et al*., 1998: 14).

Following Wolfe-Quintero *et al*. (1998), written fluency in the present dissertation will be measured as total number of words because, since the time to write the composition was limited (20 minutes), we believe that this measure is like a ratio (total number of words per minutes) and

that the higher the number of words is the more fluent the participant will be. For such a measure, a higher value is expected in the post-test.

After the transcriptions of the written tests elicited from students the CLAN program was run to analyze the frequencies for all the measures adopted for the present investigation (total number of words, number of types, number of tokens, number of error-free T-units, number of errors and number of clauses). The data from the frequencies obtained from running the CLAN program were introduced to an Stats Graphic matrix, and the formulae for each ratio were calculated. Finally, mean results of all the measures per group have been calculated and have been compared with an anova statistical analysis, the significance level set at <0.05.

4. RESULTS

This chapter presents the results of the analyses performed in order to answer the research question formulated in Chapter 3, which enquired how context of learning affects the linguistic development of young bilingual secondary education EFL learners when contrasting a group experiencing FI only and a group experiencing FI in combination with CLIL. Three specific issues are of interest in this field of research when contrasting the two contexts. The first one, related to general language development, is whether all linguistic abilities develop accordingly or differently. The subjects examined are measured as far as their writing abilities, their reading abilities, their listening comprehension abilities, and their lexico-grammatical abilities. The second issue, related to individual differences, is whether changes occur irrespective of age differences. The third and last issue relates to another individual variable, gender, which has caught the interest of research in recent years.

The description of the results obtained with the analyses performed on the basis of the design presented in the previous chapter will offer empirical data with which to confirm or disconfirm the hypotheses wich derive from our research question.

The results have been organized into different sections each dealing with one of the three research subquestions established in relation to the main question. The sections in turn contain several subsections. Firstly, section 4.1 deals with the development of general language abilities, and includes the results concerning it as referred to by Research Question 1a and the hypothesis derived from it:

RQ1a) When contrasting the differential effects of two different programmes, a FI only and a FI+CLIL, that is with an additional CLIL component, which programme results in linguistic benefits if any and which skills benefit the most if any?

H1a) When contrasting the differential effects of two programmes on learners' linguistic progress, a FI programme and a FI+CLIL with the additional hours, the group in the FI+CLIL will improve significantly more than the other, especially in receptive skills.

Subsection 4.1.1 includes results related to production in writing skills (composition), which in turn is divided into writing tests results obtained with quantitative measures (4.1.1.1) and writing tests results obtained with qualitative measures (4.1.1.2). After this, subsection 4.1.2 refers to results related to comprehension in reading skills (cloze test). Then, subsection 4.1.3 deals with results related to comprehension in oral skills (listening test: a dictation). Finally, subsection 4.1.4 concerns results in relation to lexico-grammatical ability.

Secondly, section 4.2 deals with whether changes occur irrespective of age differences, and includes the results concerning it as referred to by Research Question 1b and the hypothesis derived from it:

RQ1b) Does age have an impact on the potential benefits learners may achieve in either of the two contexts compared: FI and FI+CLIL?

H1b) At different ages and with a similar number of hours, younger learners receiving FI+CLIL will benefit more than elder learners only receiving FI because 'the earlier the better' for CLIL instruction. On the other hand, at the same age and with a different number of hours, learners with a higher total number of hours receiving FI+CLIL will benefit more than learners with fewer hours only receiving FI.

Concerning subsections within 4.2, they follow exactly the same order as subsections within 4.1 but in this case they refer to the issue of the impact of age differences in a FI and a FI + CLIL context in relation to each of the different skills measured in this study. Therefore, subsection 4.2.1 refers to results related to production in writing skills (composition), which in turn is divided into writing tests results with quantitative measures (4.2.1.1) and writing tests results with qualitative measures (4.2.1.2). After this, subsection 4.2.2 refers to results related to reading comprehension skills (cloze test). Then, subsection 4.2.3 deals with results related to comprehension in oral skills (listening test: dictation). Finally, subsection 4.2.4 concerns results in relation to lexico-grammatical ability.

In the third place, section 4.3 deals with the issue of the impact of gender differences in a FI and a FI + CLIL context in relation to each of the different skills measured in this study. These are results concerning Research Question 1c and the hypothesis derived from it:

RQ1c) Does gender have an impact in the benefits obtained with either a FI context of learning, or a FI+CLIL context?

H1c) FI + CLIL female participants will not significantly outperform their male peers in the linguistic progress achieved after a one year CLIL treatment as they reportedly do in FI.

Subsections within 4.3 deal with differences in progress in each of the two contexts of learning for three linguistic abilities: writing (subsection 4.3.1), reading (subsection 4.3.2), and listening (subsection 4.3.3).

We now turn to the presentation of these results.

4.1 Context Effects on Skill Development: Research Question 1a

In order to reject or confirm the hypothesis that when contrasting the differential effects of two programmes on learners' linguistic progress, a FI programme and a FI+CLIL with the additional hours, the group in the FI+CLIL will improve significantly more than the other especially in receptive skills (that is reading and listening, as opposed to writing and grammar) comparisons between two groups of participants, A and B, are established. That is Group A progress (T2 – T1) versus Group B progress (T2 – T1) were contrasted for that purpose. As explained in the design section, that is a comparison about the effect of the number of hours and treatment in Group A, 210 hours (140 FI + 70 CLIL), versus that of Group B (140 FI).

In the following subsections the results related to the differential effects of a FI+CLIL programme and a FI programme on the learners' different tests are described.

4.1.1 Skill Development: Writing

This subsection begins with results in the domain of written abilities because nowadays writing is a core element in institutional education in general and EFL in particular. As explained, writing was measured by means of a composition whereby they had to write a dialogue on the basis of a picture (see appendix 1C). It showed two policemen, a mother and a boy at the entrance door of a home. Learners were shown the picture and then they had 20 lines to answer the following two questions (10 lines per question): Why did this happen? How do you

think the situation will end? The individual compositions were analysed quantitatively and qualitatively in order to capture gains and lack thereof in each learning context.

Firstly, the analyses of the data with quantitative measures (4.1.1.1) is displayed. In the second place writing results measured qualitatively are presented (4.1.1.2). The reason to choose both quantitative and qualitative measures is explained in section 3.3.5.

4.1.1.1 Writing: quantitative measures

Syntactic complexity, lexical complexity, accuracy, and fluency are the areas tapped on in this study, as is in fact conventional in previous research.

Syntactic Complexity

An ANOVA statistical analysis with the significance level set at < 0.05 was performed in order to measure the use of subordination in the subjects' compositions. The results of this analysis were run with the data obtained with the syntactic complexity measure used in the compositions: the coordination index. The comparison between Group A's progress (T2-T1) and Group B's progress (T2-T1) shows greater progress is made by Group A than by Group B. However, this difference turned out **not to be statistically significant** (F[1,196]=0.25, p=0.6201).

As explained above, the coordination index was obtained by dividing the independent clause coordinations by the number of combined clauses (coordinated clauses + subordinated clauses). Since it is assumed that participants with a higher English level produce more subordinate clauses than coordinate clauses, the higher the coordination index, the lower the level of ability. Therefore, a lower ratio is expected in the post-test. The results were as follows: as appears in figure 1, Group A (FI+CLIL) obtained a coordination index of 0.40 at T1, that is column AT1, versus a coordination index of 0.39 at T2, that is column AT2. This results in a 0.01 improvement. In contrast, Group B (FI), that is column BT1, obtained an average of 0.47 at T1 versus an average of 0.49 at T2, that is column BT2. This results in a loss of 0.02. Therefore, there was a minor increase in the level of ability in Group A (FI+CLIL) versus a decrease in Group B's (FI) level of ability once the marks were

averaged. Such a difference between both groups turned out not to be statistically significant, as already mentioned.

Figure 1. Average performance in the syntactic complexity measure (coordination index) at T1 and T2
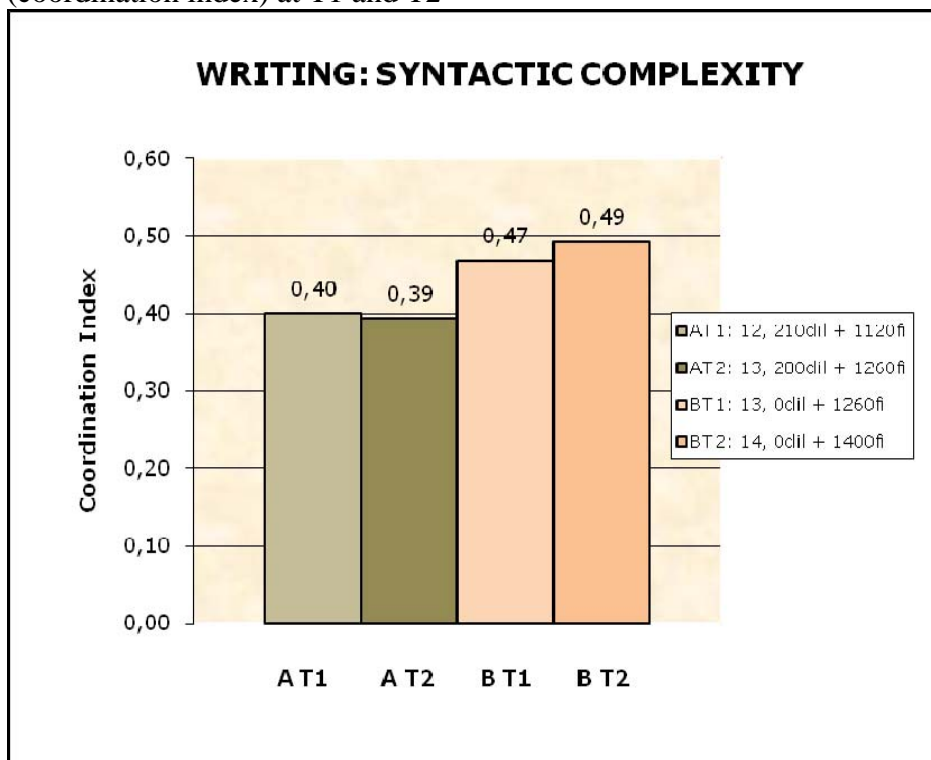
Source: personal

The previous results can be further represented in the following figure (2). It clearly shows how Group A (FI+CLIL) makes a small progress since there is a 0.01 decrease in the coordination index and therefore an increase in the use of subordination. In contrast, such a progress can not be seen in Group B (FI) as the decrease in the coordination index reflects the fact that it produces more coordinate clauses at T2 than at T1. Hence, in a FI context our subjects make use of higher levels of coordination than in a FI+CLIL context. These gains actually occur at the expense of subordination.

In addition, figures 1 and 2 clearly show how Group A (FI+CLIL) subjects start with a lower coordination index, hence with higher levels of subordination. This fact must be emphasized as it places Group A (FI+CLIL) at a different onset level than Group B (FI), something worth taking into account for a discussion of these results.

Figure 2. Progress in one year in Group A and Group B syntactic complexity (coordination index) measure
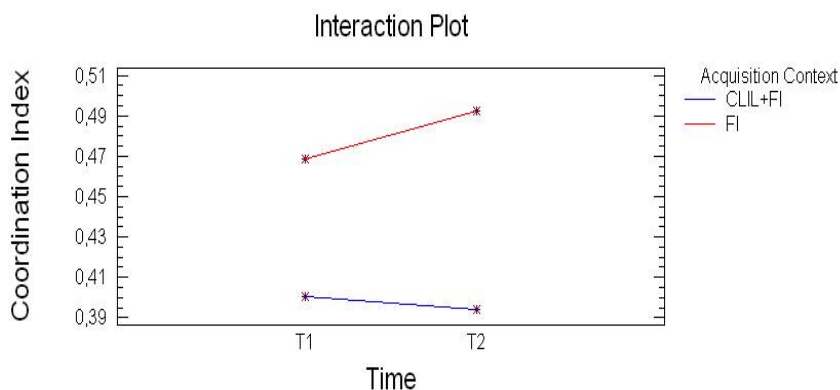
Source: personal

Lexical Complexity

The Guiraud's index was used to measure changes in the vocabulary used by our two groups of subjects after a FI+CLIL treatment and a purely FI treatment.

In this case the ANOVA revealed that the differential effect between Group A's progress (T2-T1) and Group B's progress (T2-T1) as regards lexical complexity in the compositions analysed showed that greater progress was made by Group B. However, the difference **was not statistically significant** ($F[1,196]=0.69$, $p=0.406$).

It is important to remember that the Guiraud's index was obtained by dividing the total number of lexical types by the square root of the total number of lexical tokens. Therefore, a higher value in the index indicated that the participants' vocabulary was richer. Figure 3 shows the average performance in this area of lexical richness and their graphic representation. On the one hand, Group A (FI+CLIL) obtained a Guiraud's index of 6.49 at T1, that is column AT1, and 6.7 at T2, that is column AT2. This is an improvement of 0.21. On the other hand, Group B's (FI) Guiraud's index was 6.3 at T1, that is column BT1, versus 6.7 at T2, that is column BT2. This is a 0.41 gain. Hence the rate of progress was higher in Group B (FI). Such a difference in progress between both groups was not statistically significant.

Figure 3. Average performance in the lexical complexity measure (Guiraud's index) at T1 and T2

WRITING: LEXICAL COMPLEXITY
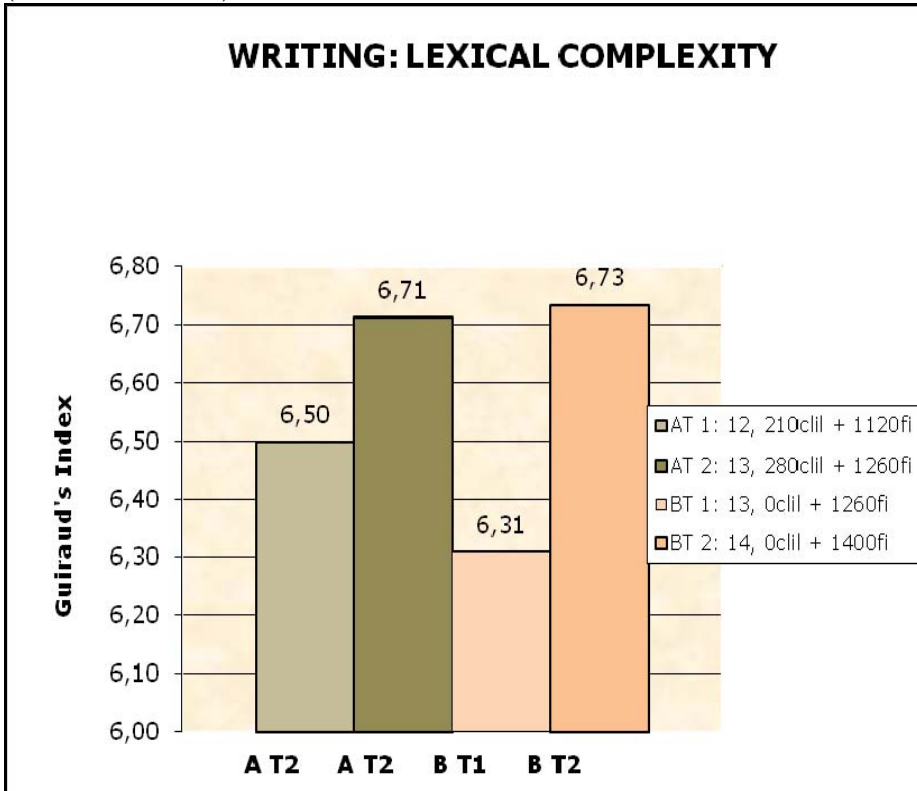
Source: personal

As in the previous area, Group A (FI+CLIL) starts the treatment with a higher onset level. This is not paired with higher gains though. In figure 4 below the three phenomena are clearly shown: the one year progress in lexical richness both in Group A (FI+CLIL) and Group B (FI), the faster rate of progress made by Group B (FI), and the higher starting level of the FI+CLIL group.

Figure 4. Progress in one year in Group A and Group B lexical complexity (Guiraud's index) measure
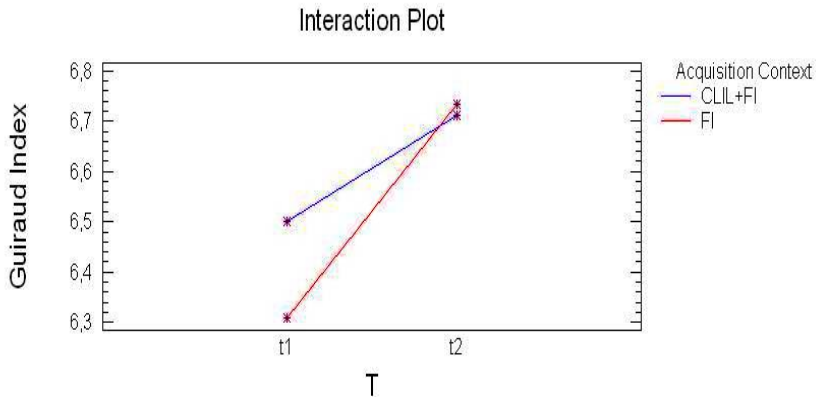
Source: personal

Accuracy

The third domain tapped in our analysis on writing is accuracy. Accuracy was measured by means of errors per word. In this domain Group A's progress (T2-T1) **was significantly higher** than gains made by Group B (T2-T1) ($F[1,196]=4.41$, $p=0.037$).

As shown in figure 5, Group A (FI+CLIL) obtained an average performance of 0.12 errors per word at T1, that is column AT1, versus a result of 0.078 errors per word at T2, that is column AT2. Therefore, it improved 0.042. In contrast, progress was not so high in Group B (FI) as the group obtained 0.092 errors per word at T1, that is column BT1, versus 0.086 errors per word at T2, that is column BT2. Hence it only reached a 0.006 improvement. Interestingly enough, in this domain Group A (FI+CLIL) starts with an initial disadvantage with respect to Group B (FI), something which had not happened with the previous measures.

Figure 5. Average performance in the accuracy measure (errors per word) at T1 and T2

WRITING: ACCURACY

Legend:
- AT 1: 12, 210dil + 1120fi
- AT 2: 13, 280dil + 1260fi
- BT 1: 13, 0dil + 1260fi
- BT 2: 14, 0dil + 1400fi
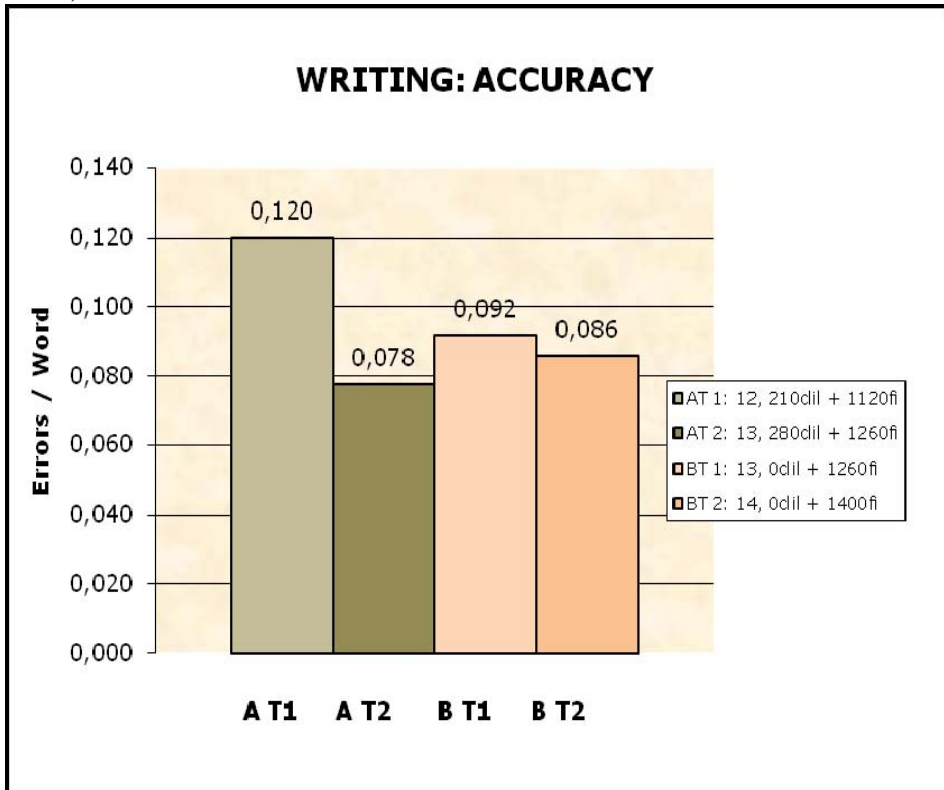
Source: personal

In figure 6 Group A's (FI+CLIL) higher advantage at T2 is evident if we compare it with the progress made by Group B (FI) in the same year. A higher decrease in the number of errors per word in Group A (FI+CLIL) is clearly shown. This, combined with the fact that the group had started with a higher number of mistakes at T1, makes their improvement in accuracy stand out.

Figure 6. Progress in one year in Group A and Group B accuracy (errors per word) measure

Source: personal

Fluency

The last domain scrutinised in this study as far as writing is concerned is fluency. The results of the ANOVA statistical analysis run with the data obtained from the quantitative measure total number of words when comparing Group A's (FI+CLIL) progress (T2-T1) and Group B's (FI) progress (T2-T1) showed that both groups decreased in fluency but Group B more than Grop A. However, this difference turned out **not to be statistically significant** ($F[1,196]=0.08$, $p=0.7801$).

Figure 7 below shows the average performance of each group in the domain of fluency. At T1 Group A (FI+CLIL) produced an average of 146.2 words in the compositions analysed, that is column AT1, but, surprisingly, after one year treatment the total number of words produced decreased to 145.1, that is column AT2. It decreased 1.1. Similarly, Group B's (FI) total number of words decreased from 149.1 at T1, that is column BT1, to 144.7 at T2, that is column BT2. This is a 4.4 decrease.

Figure 7. Average performance in the fluency measure (total number of words) at T1 and T2

WRITING: FLUENCY

Legend:
- AT 1: 12, 210clil + 1120fi
- AT 2: 13, 280clil + 1260fi
- BT 1: 13, 0clil + 1260fi
- BT 2: 14, 0clil + 1400fi

Values: A T1 = 146,2; A T2 = 145,1; B T1 = 149,1; B T2 = 144,7
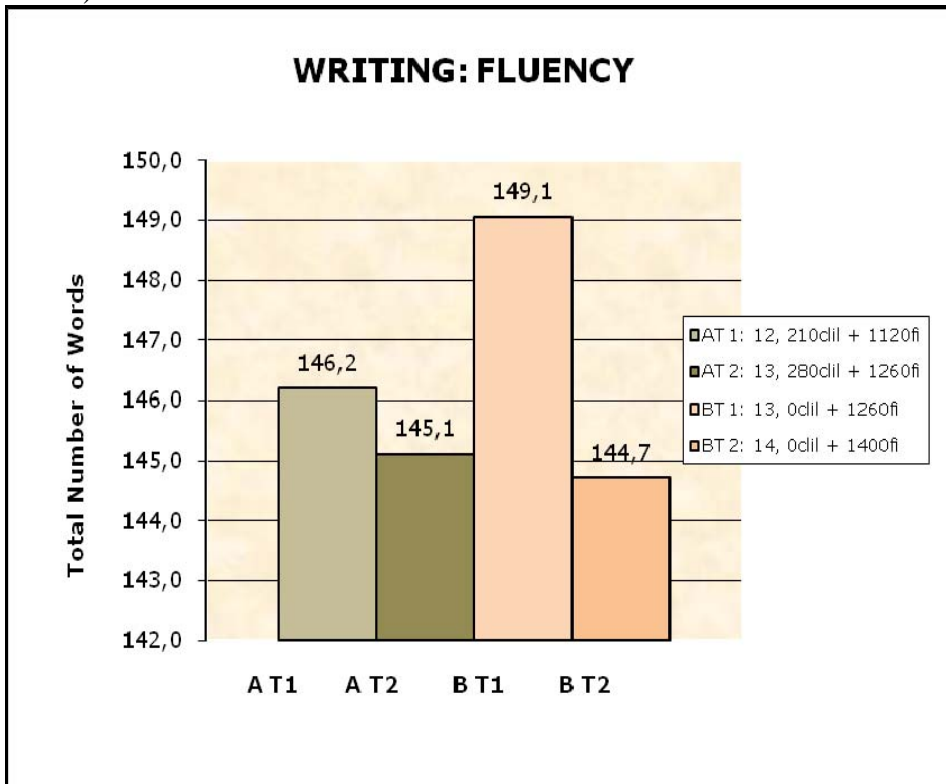
Source: personal

In figure 8 a faster decrease in the total number of words used in Group B (FI) from T1 to T2 is evident if we compare it with the decrease in Group A (FI+CLIL) in the same year. Surprisingly, both groups are less fluent after one year treatment as measured by total number of words. It is also remarkable that Group B (FI) started with a higher degree of fluency than Group A, just as had happened with the number of errors.

Figure 8. Progress in one year in Group A and Group B fluency (total number of words) measure
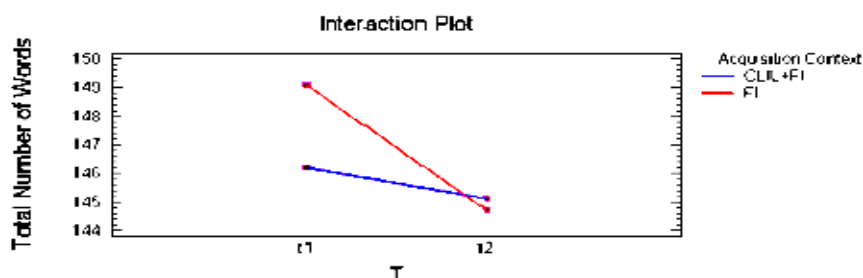
Source: personal

So far the quantitative results allow us to identify the following trends. Firstly, Group A (FI+CLIL) outperforms Group B (FI) only significantly in the domain of accuracy. However, they showed a tendency towards surpassing them as far as syntactic complexity goes. Another relevant result is the fact that in the domain of lexical and syntactic complexity Group A (FI+CLIL) has a higher onset level than Group B, whereas in the domain of accuracy and fluency, on the contrary, a lower onset level.

It is interesting to gain an overall appraisal of each context in turn and how each of the groups progresses. On the one hand, Group A (FI+CLIL) after the CLIL treatment writes shorter texts, which nonetheless are significantly more accurate, lexically richer and syntactically more complex. On the other hand, after the treatment, Group B (FI) also writes shorter texts which are more accurate and lexically richer, however less syntactically complex. Hence they both make some progress except in fluency but Group A's progress in accuracy significantly outrates that of Group B.

4.1.1.2 Writing: qualitative measures

As explained before, the compositions were also examined by means of qualitative measures. Four further subsections report on task fulfilment, organisation, grammar, and vocabulary in the compositions written by the participants.

Task Fulfilment

Task fulfilment is a measure related with the fulfilment of the task in the compositions analysed. In relation with this variable Group A's progress (T2-T1) was higher than Group B's progress (T2-T1). The difference between both groups was **not statistically significant** (F[1,96]=0.20, p=0.6572).

When measuring the compositions according to 6 behavioural levels on a scale of 0 (not enough to evaluate) to 5 (very good), as figure 9 shows, Group A (FI+CLIL) obtained an average performance of 2.92 at T1, that is column AT1, versus a 3.29 at T2, that is column AT2. This is an improvement of 0.37. On the other hand, progress in Group B (FI) was not so high since it obtained an average result of 2.63 at T1, that is column BT1, versus a 2.87 at T2, that is column BT2. This is a 0.24 improvement.

Figure 9. Average performance in the task fulfilment measure at T1 and T2
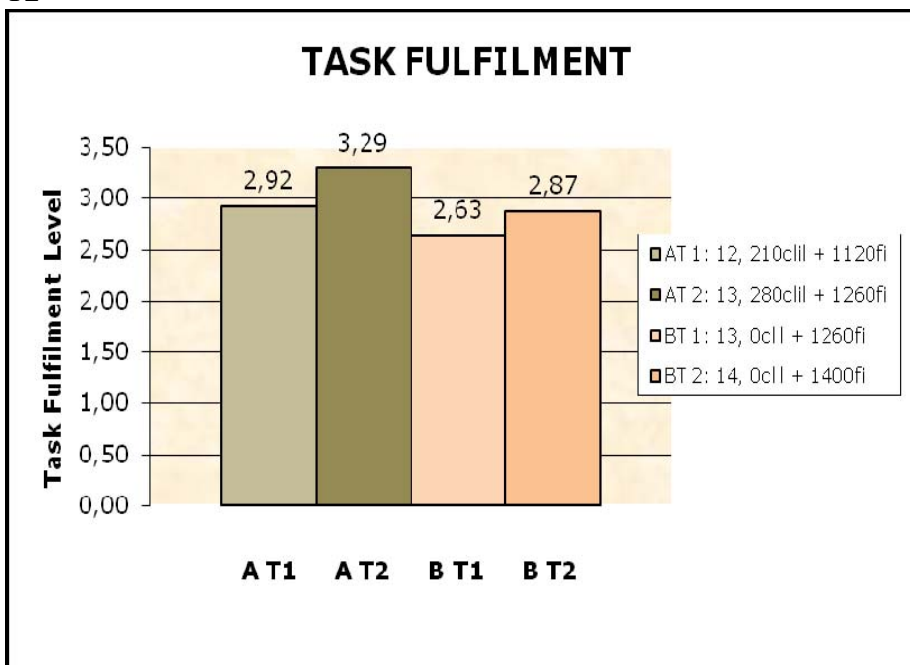
Source: personal

Figure 10 below shows the progress made by both groups in the fulfilment of the task from T1 to T2. Although it being higher in the case of in Group A (FI+CLIL), the difference with Group B was not statistically significant.

Figure 10. Progress in one year in Group A and Group B task fulfilment measure

Source: personal

Organisation

As far as the organisation of the compositions analysed is concerned, Group A's progress (T2-T1) was again higher than Group B's progress (T2-T1). However, the ANOVA statistical analysis revealed that the difference **was not significant** (F[1,96]=0.20, p=0.6565) as in the previous subsection. When measuring the compositions according to the 6 behavioural levels on a scale of 0 (not enough to evaluate) to 5 (very good), Group A (FI+CLIL) obtained an average performance of 2.84 at T1, that is column AT1, versus a 3.24 at T2, that is column AT2. This is a 0.4 improvement. In contrast, Group B's (FI) average results were of 2.49 at T1, that is column BT1, versus 2.76 at T2, that is column BT2. This is a 0.27 improvement. The absolte figures and their graphical representation are shown in figure 11:

Figure 11. Average performance in the organisation measure at T1 and T2

Source: personal

In figure 12 these results are displayed: Group A's (FI+CLIL) were higher than Group B's (FI). The progress from T1 to T2 in Group A (FI+CLIL) was also higher than in Group B (FI) although the results were similar and the difference was not statistically significant. Finally it is interesting to remark that in both this and the previous measure Group A (FI+CLIL) had a higher starting level, as happened with some of the quantitative measures.

Figure 12. Progress in one year in Group A and Group B organisation measure
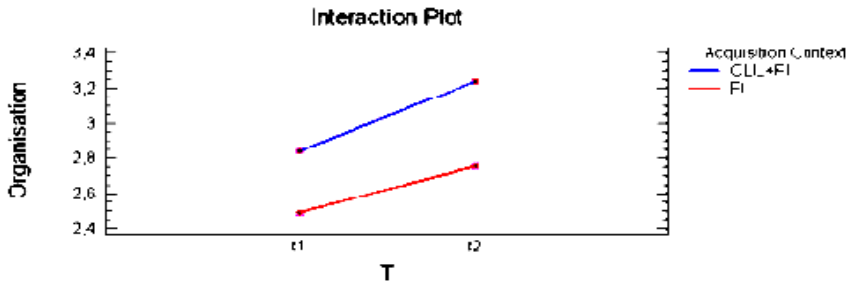
Interaction Plot

Source: personal

Grammar

In this area (accuracy), the analysis of variance of the participants' performance revealed that again Group A's results were higher than Group B's, but that however **no significant differences** between Group A's progress (T2-T1) and Group B's progress (T2-T1) were to be found ($F[1,96]=0.98$, $p=0.3240$).

As can be appreciated in figure 13, participants in Group A (FI+CLIL) obtained an average performance of 2.4 at T1, that is column AT1, versus an average of 3.06 at T2, that is column AT2, (from the 6 behavioural levels on a scale of 0 -not enough to evaluate- to 5 -very good- used in the qualitative measures). Hence the improvement amounted to 0.66. In contrast, progress in Group B (FI) was not so high since they obtained an average result of 2.34 at T1, that is column BT1, versus a 2.7 at T2, that is column BT2. This is a 0.36 improvement.

Figure 13. Average performance in the grammar measure at T1 and T2
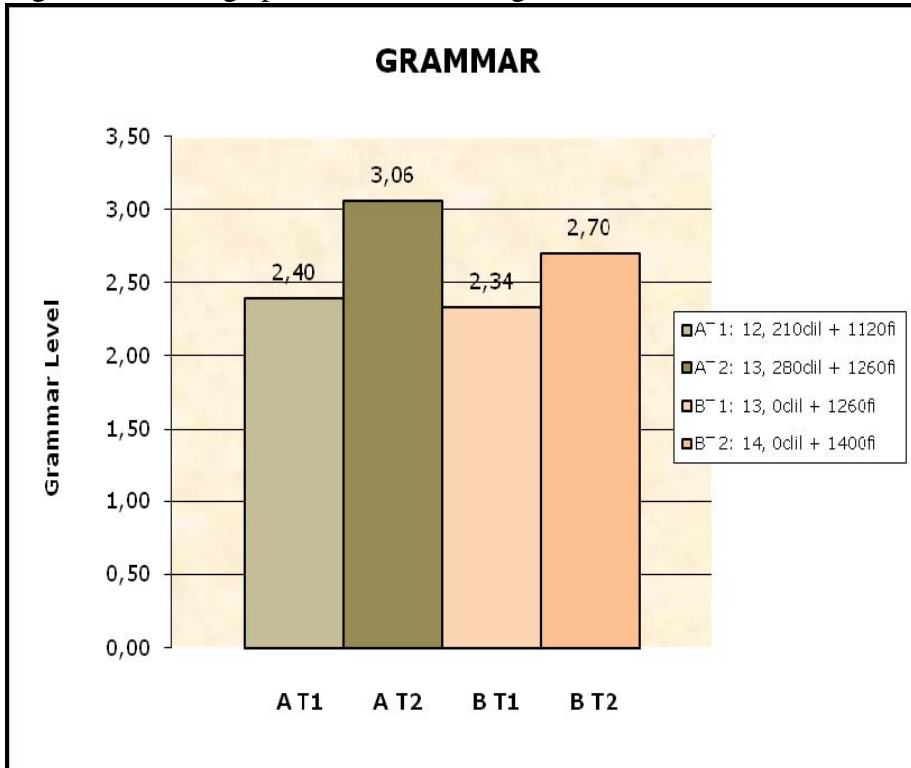
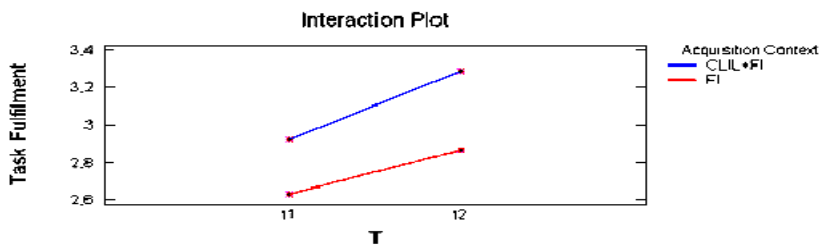Source: personal

Figure 14 below shows higher and faster progress in the case of Group A's (FI+CLIL) results from T1 to T2. However, as stated before, the difference between both groups is not statistically significant.

Figure 14. Progress in one year in Group A and Group B grammar measure

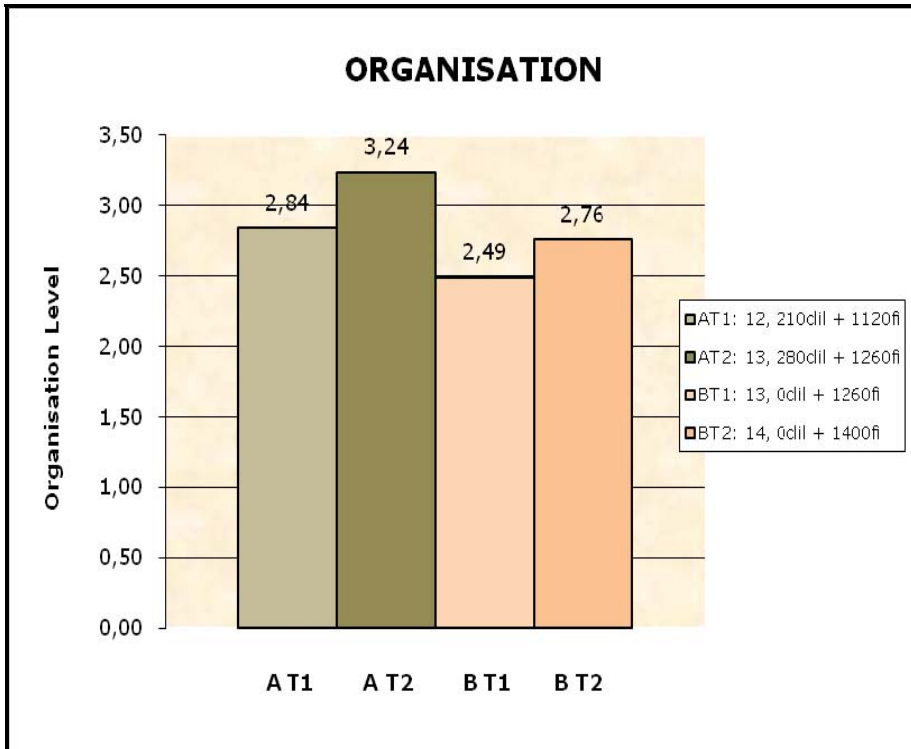Source: personal

Vocabulary

As far as the vocabulary in the compositions analysed when qualitative measures were used, Group A's progress (T2-T1) was again higher than Group B's (T2-T1), but the ANOVA statistical analysis revealed that the difference between the two groups **was not significant** (F[1,96]=2.37, p=0.1256).

Like in the previous subsections, taking into account the 6 behavioural levels on a scale of 0 (not enough to evaluate) to 5 (very good), Group A (FI+CLIL) obtained an average performance of 2.52 at T1, that is column AT1, versus 3.18 at T2, that is column AT2. This is a 0.66 improvement. On the other hand, Group B (FI) obtained an average performance of 2.53 at T1, that is column BT1, versus 2.74 at T2, that is column BT2. This is a 0.21 improvement. This is shown in figure 15:

Figure 15. Average performance in the vocabulary measure at T1 and T2
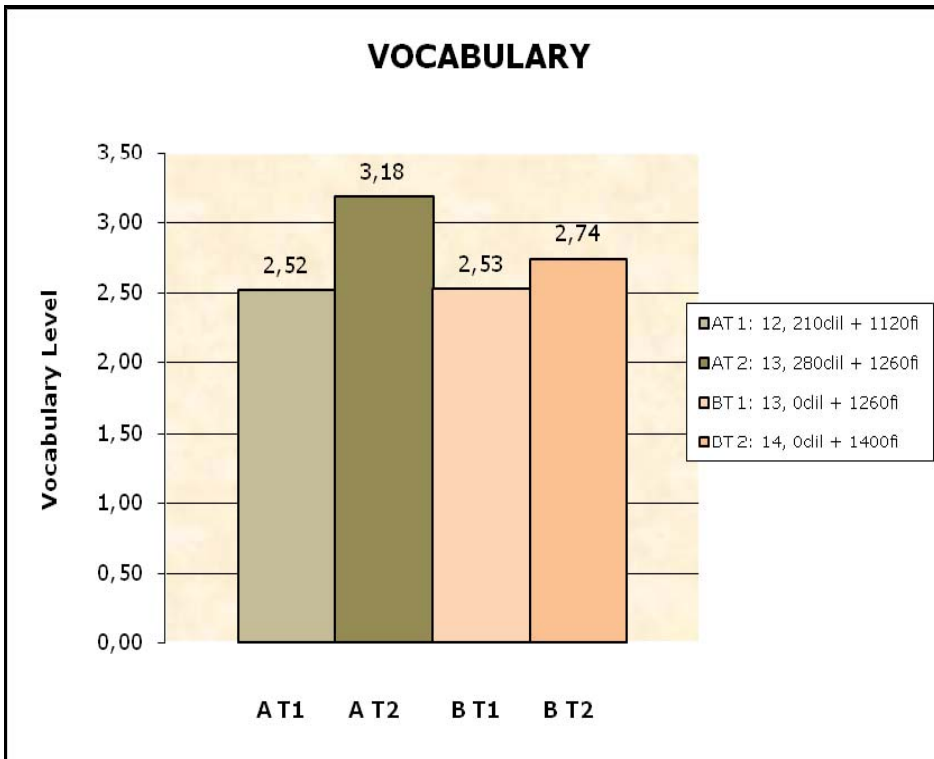
Source: personal

In figure 16 it is very easy to appreciate a higher progress in Group A (FI+CLIL) than in Group B (FI) from T1 to T2. However, the difference in progress, contrary to what it may seem, was not statistically significant.

Figure 16. Progress in one year in Group A and Group B vocabulary measure
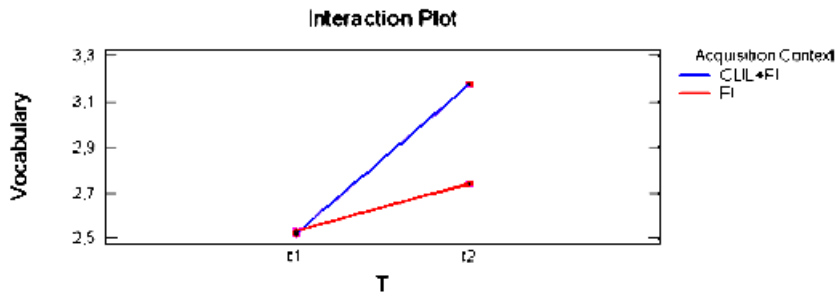
Source: personal

These qualitative results obtained by the subjects in the qualitative measures of their respective progress in writing can be summarised as follows. When analysed with qualitative measures, Group A (FI+CLIL) outperforms Group B (FI) in all areas. However, results do not reach statistical significance in any of them. More specifically, Group A (FI+CLIL) consistently tends to write a better organised, more accurate, lexically richer and more purposeful composition. It is also interesting here to highlight that Group A (FI+CLIL) always has a higher onset level except in the domain of vocabulary, albeit being a year younger. Group B's improvement in all domains of written competence is always inferior to Group A's. We now turn to the results in the domain of reading abilities.

4.1.2 Skill Development: Reading

Regarding reading abilities, the differential effect between Group A's progress (T2-T1) and Group B's progress (T2-T1) were analysed with a cloze test. The results show that although both groups improved Group A's (FI+CLIL) progress was **statistically significant** ($F[1,98]=5.14$, $p=0.0255$).

As figure 17 clearly shows, in a cloze with 20 gaps, Group A (FI+CLIL) obtained an average of 14.39 correct answers at T1, that is column AT1, and 16.08 at T2, that is column AT2. This is a 1.69 improvement. On the other hand, whereas Group B (FI) scored 14.6

correct answers at T1, that is column BT1, their progress was much lower since they only answered an average of 14.82 correct answers at T2, that is column BT2. This is a 0.22 improvement.

Figure 17. Average performance in the reading test (cloze) at T1 and T2
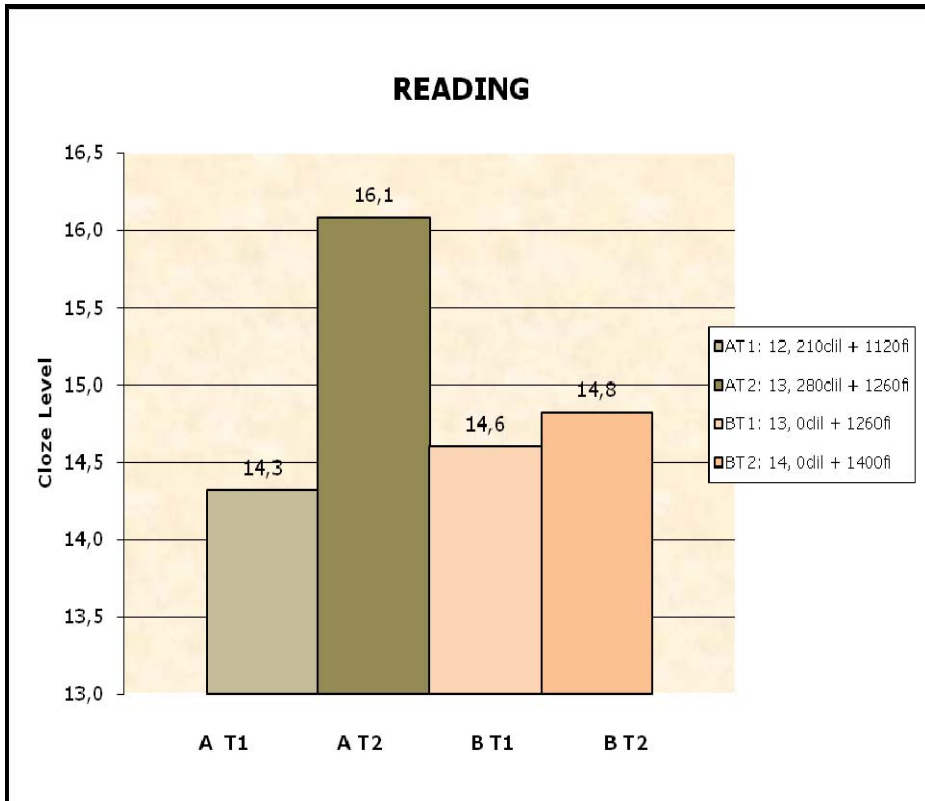
Source: personal

In figure 18 the difference in the progress after one year's treatment between Group A (FI+CLIL) and Group B (FI) is graphically shown.

Figure 18. Progress in one year in Group A and Group B reading test (cloze)
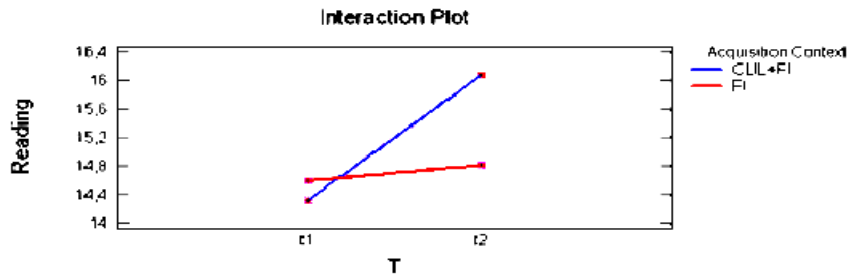
Source: personal

Hence, Group A (FI+CLIL) improves significantly more than Group B (FI) in their capacity to find the right word in a cloze test, a measure which has been associated with reading abilities.

We turn to the results obtained with a dictation task, a measure associated with listening comprehension abilities.

4.1.3 Skill Development: Listening

Regarding listening abilities, the results of the ANOVA statistical analysis run with the data obtained from the dictation between Group A's (FI+CLIL) progress (T2-T1) and Group B's (FI) progress (T2-T1) showed improvement in both groups' development. However, improvement was higher for Group B but it turned out **not to be statistically significant** (F[1,98]=0.01, p=0.924).

From a total number of 116 words in the dictation, at T1 Group A (FI+CLIL) wrote 109.4 correct words, that is column AT1 in figure 19, and the same group at T2 obtained 112.2 correct words, that is column AT2. Hence, the group showed a tendency to improve in 2.8 words. On the other hand, in Group B (FI) progress was similar since they had a score of 109.7 correct written words at T1, that is column BT1, compared to a 112.7 of total correct answers at T2, that is column BT2. This is an improvement of 3 words.

Figure 19. Average performance in the listening test at T1 and T2
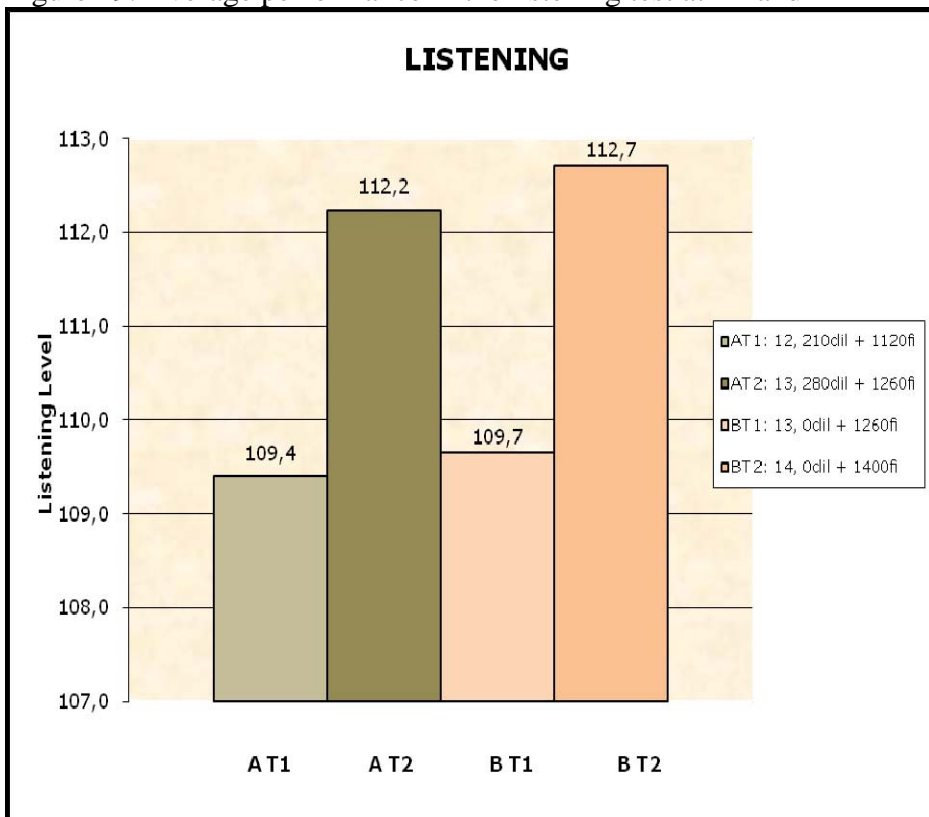
Source: personal

As can be clearly seen in figure 20, the progress in one year time concerning oral comprehension was practically identical in Group A (FI+CLIL) and in Group B (FI).

Figure 20. Progress in one year in Group A and Group B listening test
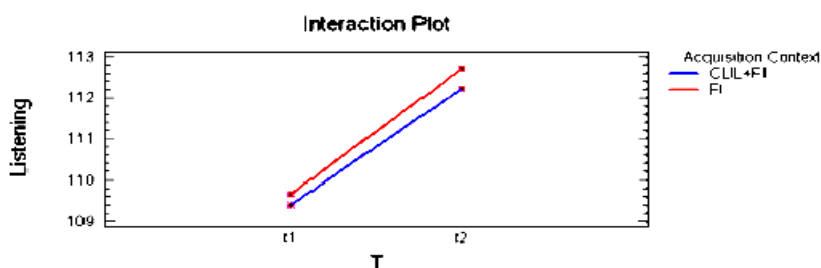
Source: personal

The final set of tests administered to our two groups were grammatical tests sensed to measure the ability to deal with the syntactic system of the language, presented in the following subsection.

4.1.4 Skill Development: Lexico-Grammatical Ability

As for lexico-grammatical ability, similarly to previous subsections, an ANOVA statistical analysis with the significance level set at < 0.05 was performed with the data obtained from the grammar tests. In this area the analysis of variance of the participants' performance revealed **significant differences** between Group A's progress (T2-T1) and Group B's progress (T2-T1) ($F[1,98]=7.39$, $p=0.0078$) in favour of Group A (FI+CLIL).

Indeed, as shown in figure 21, from a total number of 50 questions, Group A (FI+CLIL) at T1 obtained 37.1 correct answers, that is column AT1, and the same group at T2 obtained an average of 39.82 correct answers, that is column AT2. This represents an improvement of 2.72. On the other hand, in Group B (FI) the progress made was not so important since they had a score of 38.46 correct answers at T1, that is column BT1, compared to a 38.76 of total correct answers at T2, that is column BT2. This amounts to an improvement of 0.3.

Figure 21. Average performance in the lexico-grammatical tests at T1 and T2
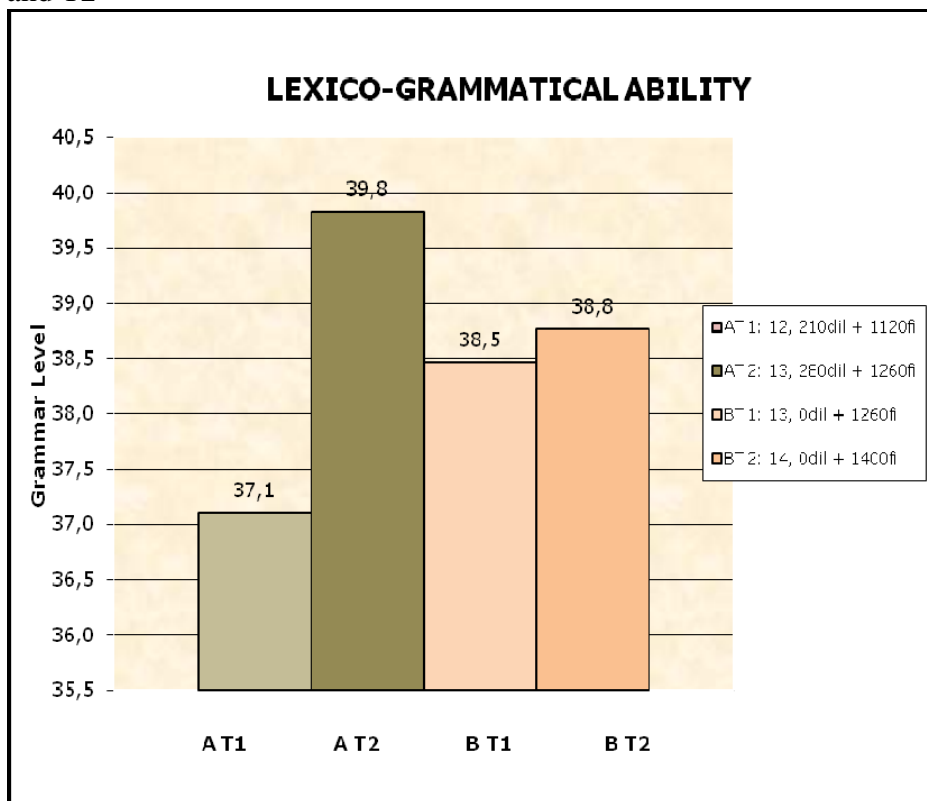
Source: personal

It is interesting to note that the effect of the conventional FI classroom on this grammatical ability test, the main focus in such a context of learning, and certainly in this school, does not seem to have an impact. On the contrary, it is when an additional CLIL ingredient is added with no explicit grammar teaching to the staple diet of FI, that grammar seems to improve.

As figure 22 clearly shows, the progress in Grammar in one year time was higher in Group A (FI+CLIL) than in Group B (FI). Even more interesting is the fact that the results in the domain of accuracy in writing both quantitative (see 4.1.1.1 above) and qualitative (see 4.1.1.2 above) also show Group A's greater progress, the former to a statistically significant degree.

Figure 22. Progress in one year in Group A and Group B lexico-grammatical tests
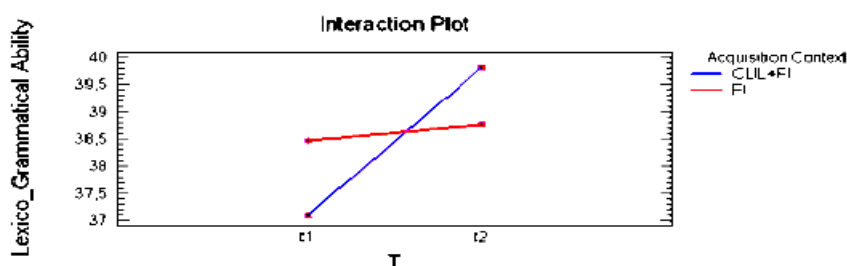
Source: personal

We can summarise this second set of tests as follows: both groups improve their reading, listening and lexico-grammatical abilities. However, Group A (FI+CLIL) makes significantly greater improvement in reading and lexico-grammatical abilities than Group B (FI). That is the subjects read, understand and perform formally better after their respective one-year treatments. Another result to be highlighted is the fact that in the three domains Group A (FI+CLIL) started with a lower level of competence. This is similar to Group A's performance in written fluency and accuracy, but stands in contrast with their onset level in written lexis (when quantitatively measured) and syntactic complexity.

4.2 Age Effects on Contexts of Instruction: Research Question 1b

The second section within the results chapter deals with another key question in our study: the impact of age differences in a FI and a FI + CLIL context in relation to each of the different skills measured. We seek to either confirm or reject the hypothesis (H1b) that, on the one hand, at different ages and with a similar number of hours, younger learners receiving FI+CLIL will benefit more than elder learners only receiving FI because 'the earlier the better' for CLIL instruction. On the other hand, at the same age and with a different number of hours,

learners with a higher total number of hours receiving FI+CLIL will benefit more than learners with fewer hours only receiving FI.

In order to address the age hypothesis another between-groups twofold comparison was established between Group A at T1 and group B at T2. The first set of comparisons was established between Group A (FI+CLIL) and Group B (FI) when they shared a similar number of hours of exposure to the target language, either through FI only or CLIL + FI, at different ages. In this sense, Group A would have 1330 h at 12 years old and Group B would have 1400 h at 14 years old. The second set of comparisons was established when they shared the same age but a different number of hours: Group A (FI+CLIL) 1540 h versus Group B (FI) 1260 h both at 13 years old.

Table 13b[15]. Design

	T1 (2005)	T2 (2006)
A: FI + **CLIL**	Grade7 / 1st ESO (12 yrs.) FI: 1120 h + **CLIL: 210h** = 1330	Grade8 / 2nd ESO (13 yrs.) FI: 1260 h + **CLIL: 280h** = 1540 (+ 210h)
B: FI	Grade8 / 2nd ESO (13 yrs.) FI: 1260 h CLIL: 0h	Grade9 / 3rd ESO (14 yrs.) FI: 1400 h CLIL: 0h (+ 140h)

Source: personal

In the following subsections the results related to the differential effects of age on the learners' linguistic abilities are displayed.

4.2.1 Differential Effects of Age: Writing

This subsection is divided into writing tests results related to quantitative measures (4.2.1.1) and writing tests results related to qualitative measures (4.2.1.2).

4.2.1.1 Age effects on writing: quantitative measures

Concerning quantitative measures, there follows four further subsections according to the measures chosen to calculate syntactic complexity, lexical complexity, accuracy, and fluency.

---

[15] For the sake of reader friendliness table 13 is reproduced here again.

Syntactic Complexity

On the one hand, in order to address the first question regarding age, an ANOVA statistical analysis with the significance level set at < 0.05 was performed to calculate syntactic complexity. The results of this analysis run with the data obtained from the syntactic complexity measure in the compositions between Group A at T1 (12 years old and 1330 h of CLIL+FI) and Group B at T2 (14 years old and 1400 h of FI) turned out **to be statistically significant** (p=0.028) in favour of Group A. This group was found to use subordination at the expense of coordination significantly more than the other group.

On the other hand, in order to address the second question, the results of the analysis run with the data obtained from the syntactic complexity measure in the compositions between Group A at T2 (13 years old and 1540 h of CLIL+FI) and Group B at T1 (13 years old and 1260 h of FI) were these: In spite of being favourable to Group A, they turned out **not to be statistically significant** (p=0.089).

Figure 23 below shows how Group A at T1 (12 years old and 1330 h of CLIL+FI) obtained a coordination index of 0.40, that is column AT1. In contrast, Group B at T2 (14 years old and 1400 h of FI) obtained an average of 0.49 concerning the coordination index, that is column BT2. As for the second comparison, when sharing the same age, in figure 23 it is shown how Group A at T2 (13 years old and 1540 h of CLIL+FI) obtained a coordination index of 0.39, that is column AT2, whereas Group B at T1 (13 years old and 1260 h of FI) obtained a 0.47, that is column BT1.

Figure 23. Average performance in the syntactic complexity measure (coordination index) at T1 and T2

WRITING: SYNTACTIC COMPLEXITY

Legend:
- AT1:12, 210clil+1120fi
- AT2:13, 280clil+1260fi
- BT1:13, 0clil+1260fi
- BT2:14, 0 clil+1400fi
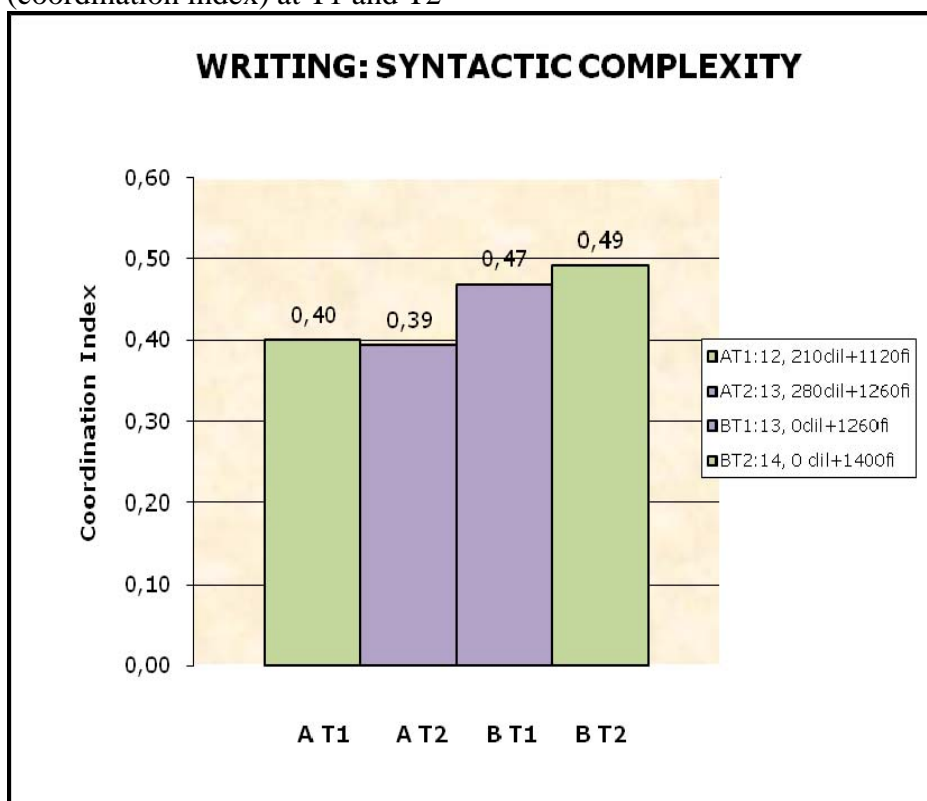
Source: personal

Lexical Complexity

As the ANOVA revealed, the difference between Group A at T1 (12 years old and 1330 h of CLIL+FI) and Group B at T2 (14 years old and 1400 h of FI) concerning the lexical complexity measure was **not statistically significant** (p=0.219). However, it was favourable to Group A again.

On the other hand, the results of the analysis between Group A at T2 (13 years old and 1540 h of CLIL+FI) and Group B at T1 (13 years old and 1260 h of FI), being favourable again to Group A, also turned out **not to be statistically significant** (p=0.08).

The following figure (24) displays how Group A at T1 (12 years old and 1330 h of CLIL+FI) obtained a Guiraud's index of 6.50, that is column AT1, and Group B at T2 (14 years old and 1400 h of FI)

obtained a Guiraud's index of 6.73, that is column BT2. As for the second comparison, when sharing the same age, in the same figure is shown how Group A at T2 (13 years old and 1540 h of CLIL+FI) obtained a Guiraud's index of 6.71, that is column AT2, whereas Group B at T1 (13 years old and 1260 h of FI) obtained a 6.3, that is column BT1.

Figure 24. Average performance in the lexical complexity measure (Guiraud's index) at T1 and T2

Source: personal

Accuracy

The difference between the gains on written accuracy calculated with the quantitative measure errors per word in Group A at T1 (12 years old and 1330 h of CLIL+FI) and Group B at T2 (14 years old and 1400 h of FI) **was statistically significant** (p=0.012) in favour of Group B. As for the second comparison, the results of the analysis between Group A at T2 (13 years old and 1540 h of CLIL+FI) and Group B at T1 (13 years old and 1260 h of FI), although being favourable to Group A this time, turned out **not to be statistically significant** (p=0.19).

As figure 25 below shows, Group A at T1 (12 years old and 1330 h of CLIL+FI) obtained an average performance of 0.12 errors per word, that is column AT1, and Group B at T2 (14 years old and 1400 h of FI) only made an average of 0.086 errors per word, that is column BT2. On the other hand, the results were very different when sharing the same age. In this case Group A at T2 (13 years old and 1540 h of CLIL+FI) obtained 0.078 errors per word, that is column AT2, whereas Group B at T1 (13 years old and 1260 h of FI) made 0.092 errors per word, that is column BT1.

Figure 25. Average performance in the accuracy measure (errors per word) at T1 and T2
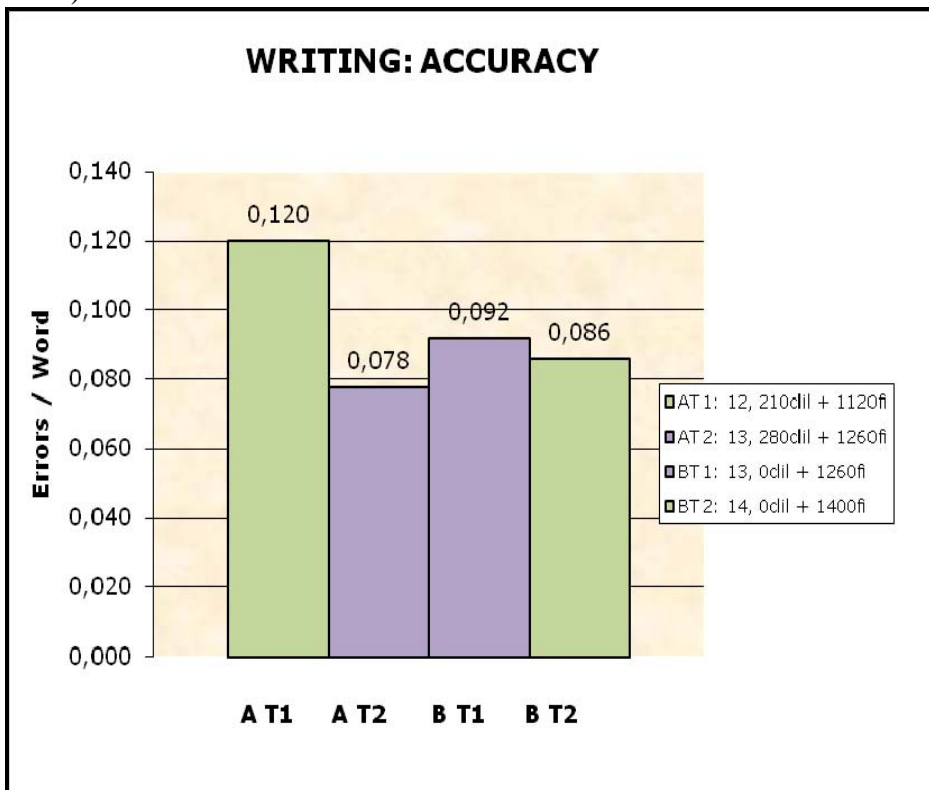
Source: personal

Fluency

The results of the ANOVA statistical analysis run with the data obtained from this fluency measure in the compositions of Group A at T1 (12 years old and 1330 h of CLIL+FI) and Group B at T2 (14 years old and 1400 h of FI) showed greater progress for Group A, although it turned out **not to be statistically significant** (p=0.855).

On the other hand, the results of the same analysis concerning the difference between Group A at T2 (13 years old and 1540 h of CLIL+FI) and Group B at T1 (13 years old and 1260 h of FI) were favourable to Group B but they were **not statistically significant** (p=0.638).

Figure 26 below shows how Group A at T1 (12 years old and 1330 h of CLIL+FI) produced an average of 146.2 words, that is column AT1. In contrast, Group B at T2 (14 years old and 1400 h of FI) obtained an average of 144.7 concerning the total number of words, that is column BT2. As for the second comparison, when sharing the same age, figure 26 shows how Group A at T2 (13 years old and 1540 h of CLIL+FI) produced 145.1 words, that is column AT2, whereas Group B at T1 (13 years old and 1260 h of FI) produced an average of 149.1 total number of words, that is column BT1.

Figure 26. Average performance in the fluency measure (total number of words) at T1 and T2
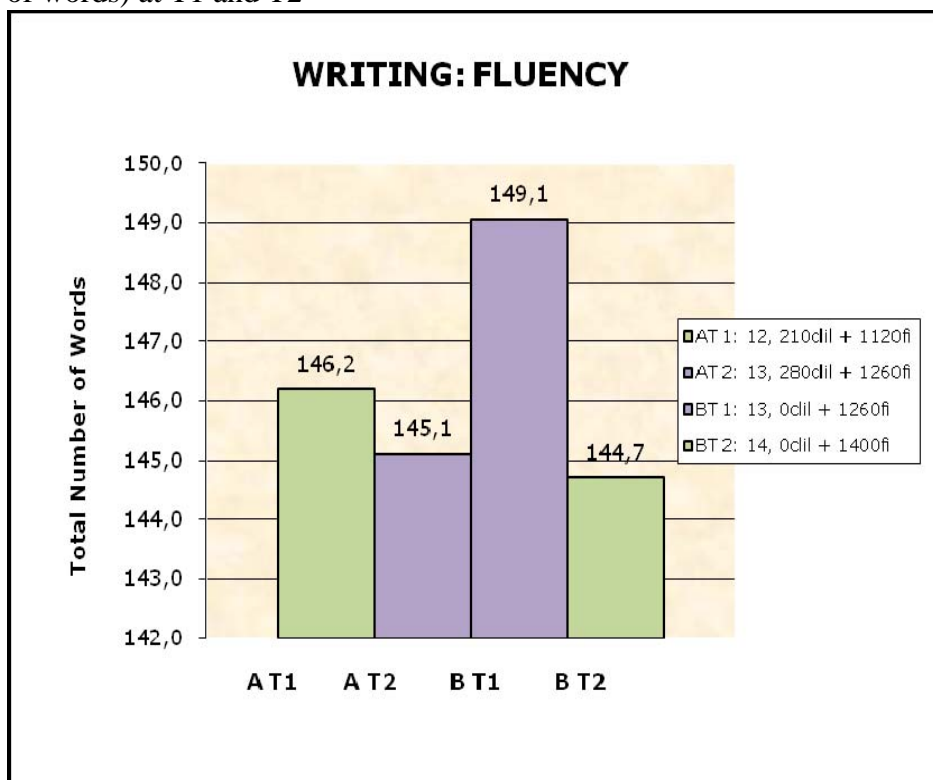
Source: personal

To summarise the results which show the first set of comparisons between both groups at different ages but a similar number of hours, significantly better results yielded for the younger group albeit having had fewer hours of exposure to English in the domain of syntactic complexity. Thus, they used more subordination. On the other hand, significantly better results yielded for the older group in the area of accuracy: they made fewer mistakes.

Concerning the summary about the second set of comparisons, that is groups at the same age but different number of hours, a clear tendency in favour of the group with more hours (extra CLIL hours) in the domain of syntactic complexity, lexical complexity and accuracy was noted. Thus, although not significantly better, at the same age the group with extra CLIL hours used more subordination, made fewer mistakes

and was lexically richer. However, it was less fluent than the group with the same age but fewer hours.

4.2.1.2 Age effects on writing: qualitative measures

Following the same organisation as with the results addressing the issue of the impact of age on groups A and B, concerning the qualitative measures, there follow four further subsections related to task fulfilment, organisation, grammar, and vocabulary in the compositions written by the participants.

Task Fulfilment

Although Group A's results at T1 (12 years old and 1330 h of CLIL+FI) as regards the fulfilment of the task in the compositions analysed was higher than Group B's results at T2 (14 years old and 1400 h of FI), the difference **was not statistically signi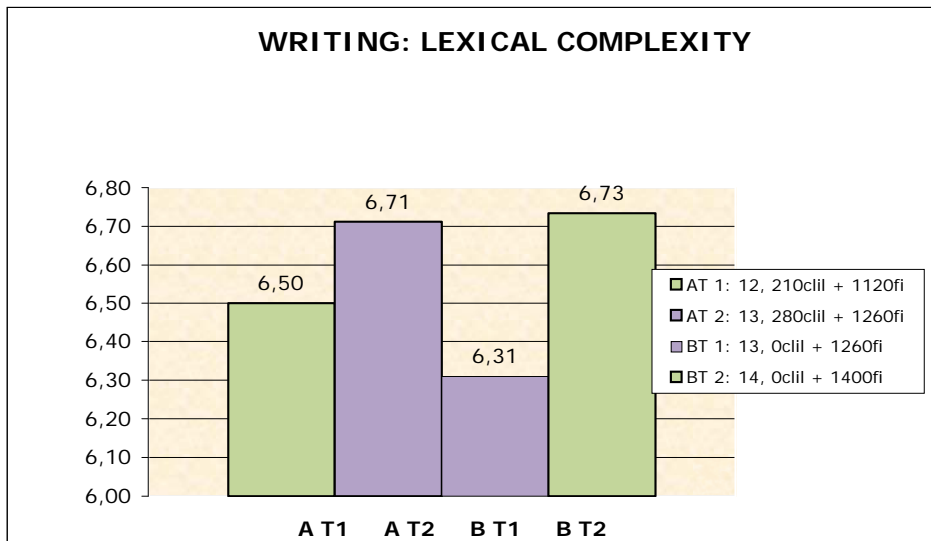ficant** (p=0.819). As for the second part of the hypothesis, the results showed how the difference between Group A at T2 (13 years old and 1540 h of CLIL+FI) and Group B at T1 (13 years old and 1260 h of FI) **was significant** (p=0.001) in favour of Group A.

Taking into account the 6 behavioural levels on a scale of 0 (not enough to evaluate) to 5 (very good), as figure 27 shows, Group A at T1 (12 years old and 1330 h of CLIL+FI) obtained an average performance of 2.92, that is column AT1, and Group B at T2 (14 years old and 1400 h of FI) obtained an average performance of 2.87, that is column BT2. On the other hand, when sharing the same age, Group A at T2 (13 years old and 1540 h of CLIL+FI) obtained a result of 3.29, that is column AT2, whereas Group B result at T1 (13 years old and 1260 h of FI) was 2.63, that is column BT1.

Figure 27. Average performance in the task fulfilment measure at T1 and T2
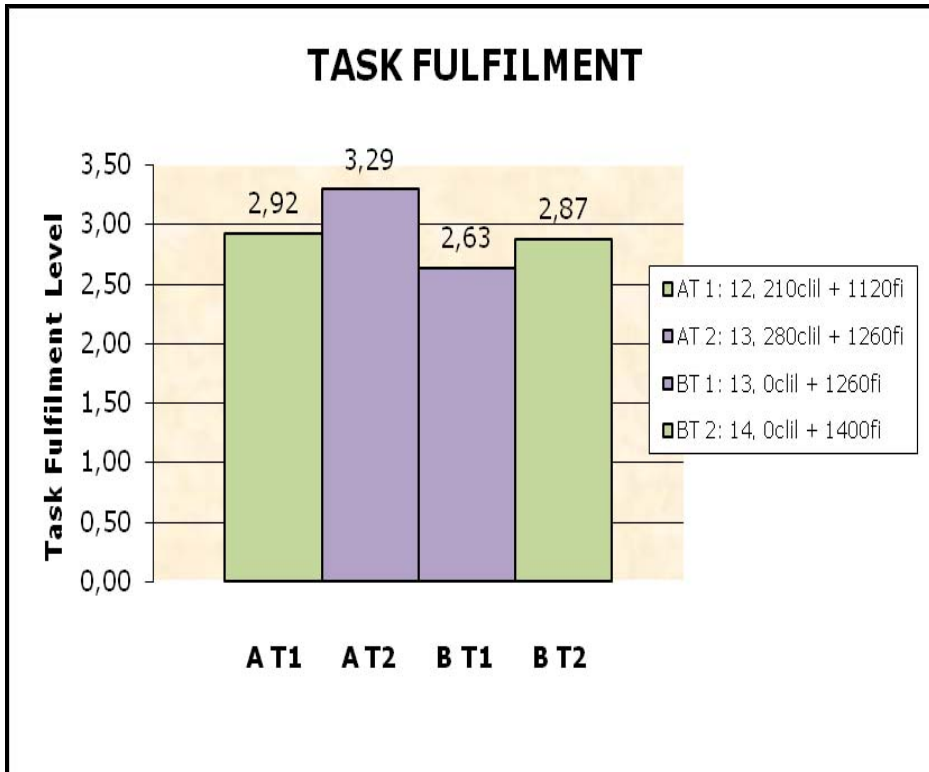
Source: personal

Organisation

As far as the organisation of the compositions analysed is concerned, Group A at T1 result (12 years old and 1330 h of CLIL+FI) was higher than Group B at T2 result (14 years old and 1400 h of FI), but the ANOVA statistical analysis revealed that the difference **was not significant** (p=0.6949). On the other hand, when comparing both groups at the same age, Group A at T2 (13 years old and 1540 h of CLIL+FI) and Group B at T1 (13 years old and 1260 h of FI), the difference **was significant** (p=0.0005) in favour of Group A.

Similar to what has been presented in the previous subsection, taking into account the 6 behavioural levels on a scale of 0 (not enough to evaluate) to 5 (very good), figure 28 shows how Group A's result at T1 (12 years old and 1330 h of CLIL+FI) was 2.84, that is column AT1, whereas at T2 Group B (14 years old and 1400 h of FI) obtained an

average performance of 2.76 as far as organisation is concerned, that is column BT2. In contrast, Group A at T2 (13 years old and 1540 h of CLIL+FI) average result was of 3.24, that is column AT2, and Group B at T1 (13 years old and 1260 h of FI) obtained an average performance of 2.49, that is column BT1.

Figure 28. Average performance in the organisation measure at T1 and T2
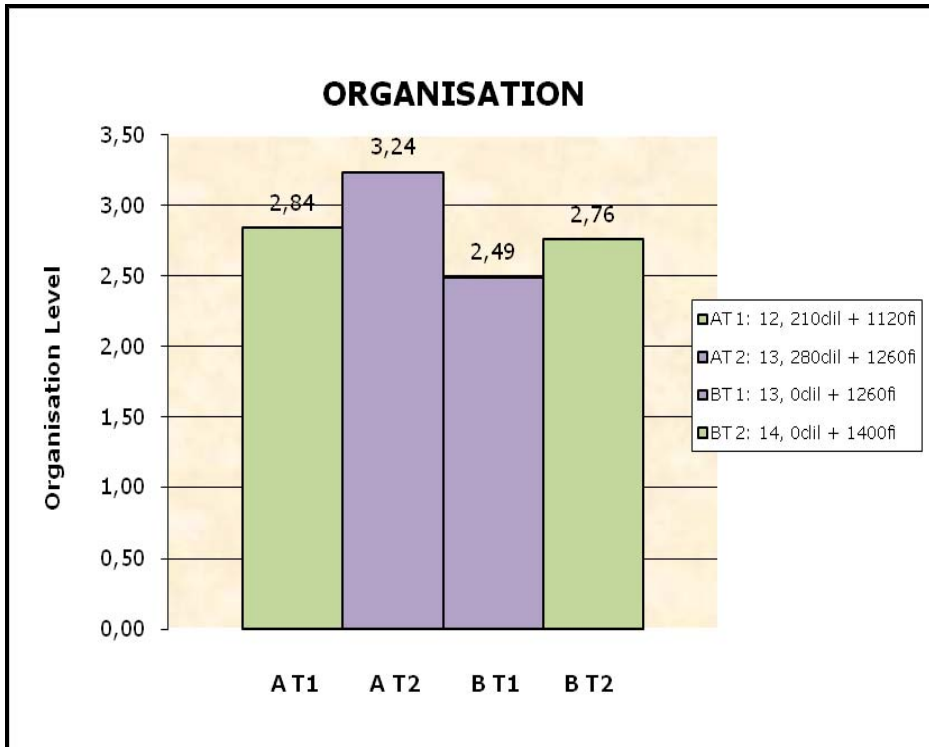
Source: personal

Grammar

In this area (accuracy), the analysis of variance of the participants' performance revealed that again **no significant** differences between Group A at T1 (12 years old and 1330 h of CLIL+FI) and Group B at T2 (14 years old and 1400 h of FI) were found (p=0.17), however, Group A performed slightly better that Group B. In contrast, when comparing the groups with the same age, the difference **was significant** (p=0.001) in favour of Group A.

As can be appreciated in figure 29, participants in Group A at T1 (12 years old and 1330 h of CLIL+FI) obtained an average performance of 2.4, that is column AT1 (from the 6 behavioural levels on a scale of 0 - not enough to evaluate- to 5 -very good- used in the qualitative measures). On the other hand, in Group B at T2 (14 years old and 1400 h of FI) the result was 2.7, that is column BT2. In the second comparison the difference was more important since Group A at T2 (13 years old and 1540 h of CLIL+FI) average result was of 3.06, that is column AT2, whereas Group B at T1 (13 years old and 1260 h of FI) obtained an average performance of 2.34, that is column BT1.

Figure 29. Average performance in the grammar measure at T1 and T2
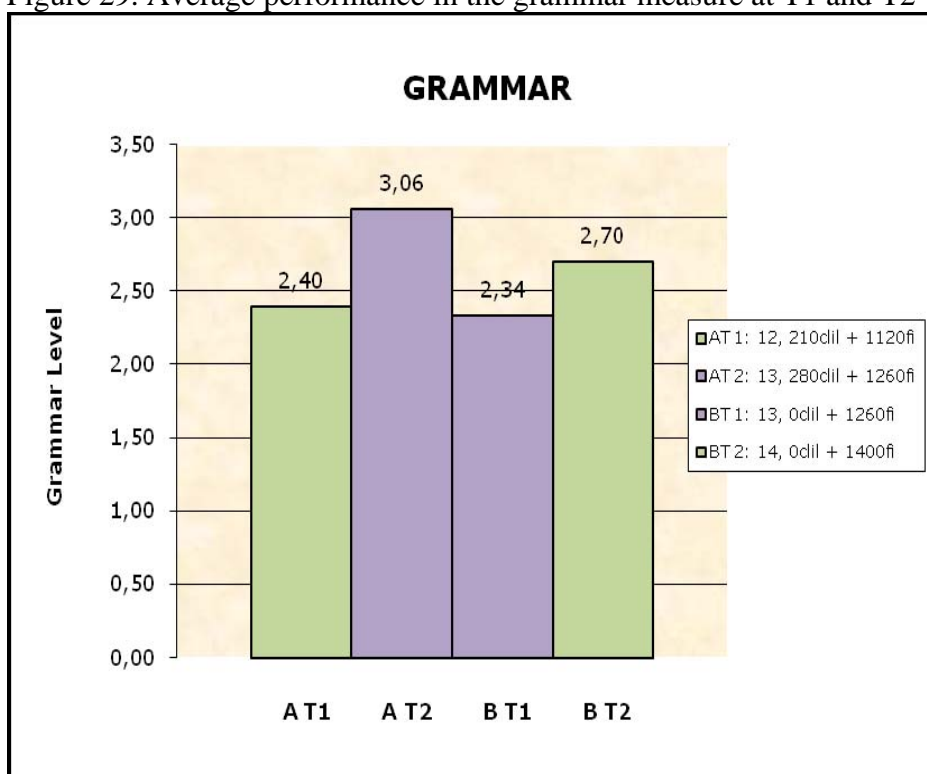
Source: personal

Vocabulary

As far as the vocabulary of the compositions analysed using qualitative measures is concerned, the ANOVA statistical analysis revealed that the difference between Group A at T1 (12 years old and 1330 h of CLIL+FI) and Group B at T2 (14 years old and 1400 h of FI) was **not significant** (p=0.294) but Group B performed better than Group A. However, the difference between Group A at T2 (13 years old and 1540

h of CLIL+FI) and Group B at T1 (13 years old and 1260 h of FI) was more important and the ANOVA statistical analysis revealed that it **was significant** (p=0.002) in favour of Group A.

Just as in the previous subsections, taking into account the 6 behavioural levels on a scale of 0 (not enough to evaluate) to 5 (very good), figure 30 clearly shows how Group A at T1 (12 years old and 1330 h of CLIL+FI) obtained an average performance of 2.52, that is column AT1, and Group B at T2 (14 years old and 1400 h of FI) obtained a result of 2.74, that is column BT2. On the other hand, in the second comparison the difference was again more important since Group A at T2 (13 years old and 1540 h of CLIL+FI) average result was of 3.18, that is column AT2, whereas Group B at T1 (13 years old and 1260 h of FI) obtained an average performance of 2.53, that is column BT1.

Figure 30. Average performance in the vocabulary measure at T1 and T2
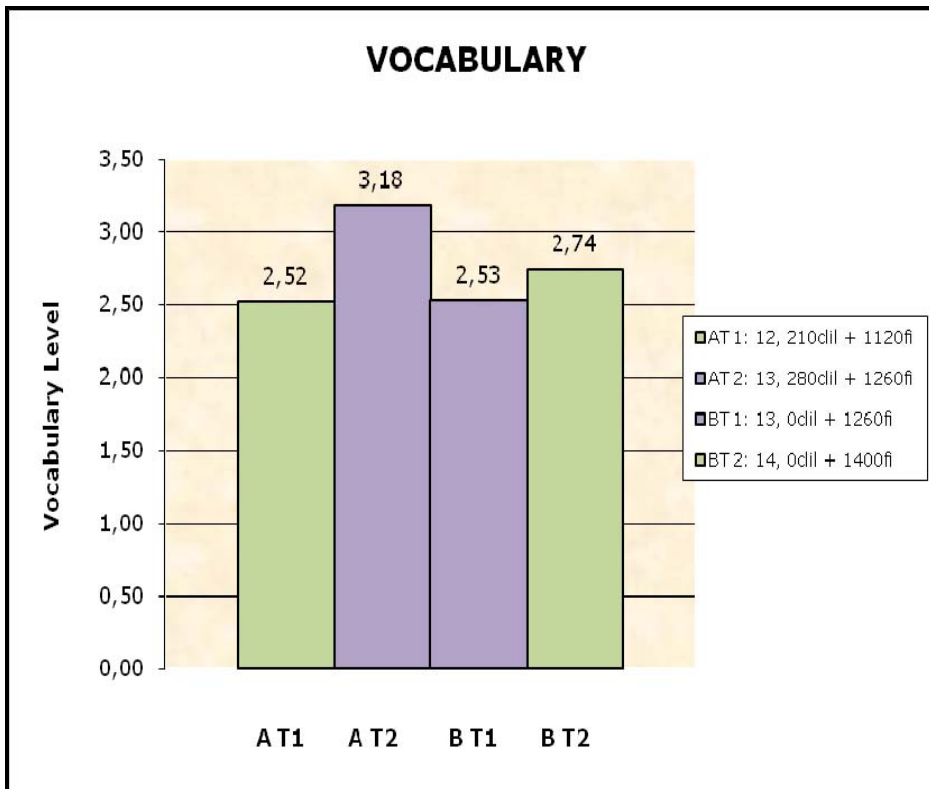
Source: personal

In sum, when analysed by means of qualitative measures, results on the written ability of our two groups of participants turned out not to be statistically significant for the first issue related to age. That is, when measured at different ages but a similar number of hours, their results were not significantly different even if there was a tendency for the younger participants in Group A to surpass the older ones in Group B in task fulfilment and organisation, but not in the areas of grammar and vocabulary. In contrast, when measured at the age of 13, the results of participants with higher number of hours, CLIL+FI, were statistically better for the four areas. That is, at the same age, 13 years, but with extra CLIL hours, the participants' compositions were more purposeful, organised, accurate, and lexically richer.

4.2.2 Differential Effects of Age: Reading

The differential effect between Group A at T1 (12 years old and 1330 h of CLIL+FI) and Group B at T2 (14 years old and 1400 h of FI) regarding results in reading skills analysed with a cloze test was **not statistically significant** (p=0.483) and Group B was slightly better than Group A. On the other hand, when the age of both groups was the same, that is Group A at T2 (13 years old and 1540 h of CLIL+FI) and Group B at T1 (13 years old and 1260 h of FI), the ANOVA statistical analysis revealed that the difference **was significant** (p=0.02) in favour of the group with extra CLIL hours.

As figure 31 clearly displays, in a cloze test with 20 gaps, Group A at T1 (12 years old and 1330 h of CLIL+FI) obtained an average of 14.32 correct answers, that is column AT1, whereas Group B at T2 (14 years old and 1400 h of FI) obtained an average of 14.82 correct answers, that is column BT2. In the second comparison, whereas Group A at T2 (13 years old and 1540 h of CLIL+FI) scored 16.08 correct answers, that is column AT2, Group B at T1 (13 years old and 1260 h of FI), answered an average of 14.6 correct answers, that is column BT1.

Figure 31. Average performance in the reading test (cloze) at T1 and T2
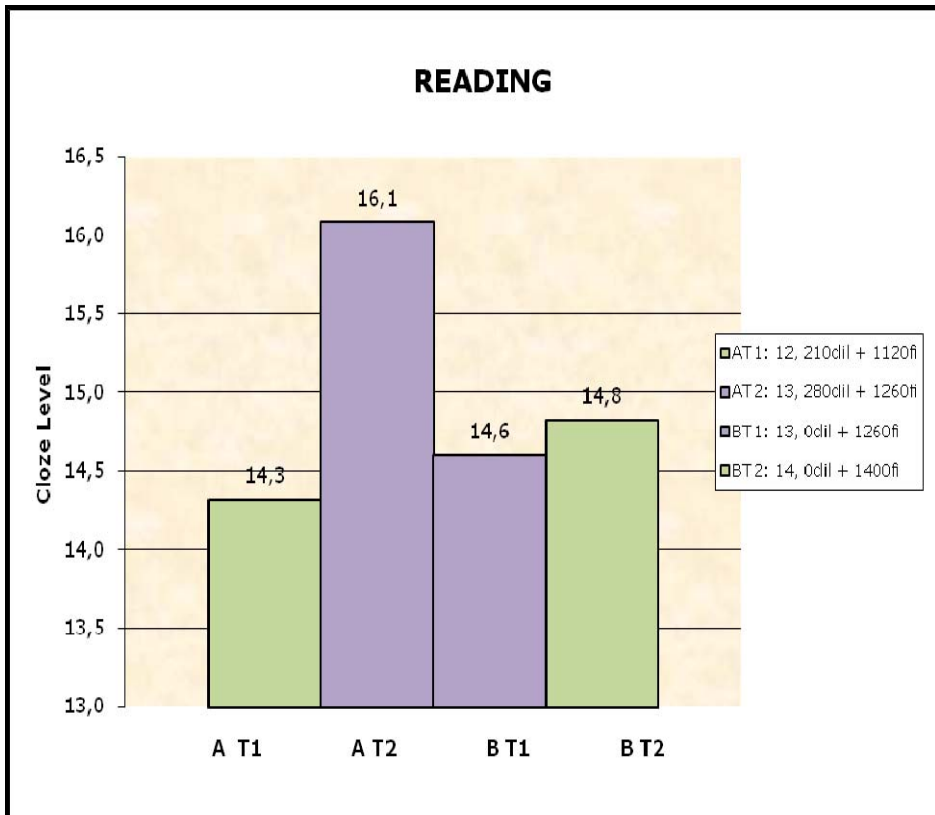
Source: personal

4.2.3 Differential Effects of Age: Listening

The results of the ANOVA statistical analysis run with the data obtained from the dictation concerning the difference between Group A at T1 (12 years old and 1330 h of CLIL+FI) and Group B at T2 (14 years old and 1400 h of FI) turned out to be **statistically significant** (p=0.003) in favour of Group B. As for the second comparison, the difference between the results concerning Group A at T2 (13 years old and 1540 h of CLIL+FI) and Group B at T1 (13 years old and 1260 h of FI) was **not statistically significant** (p=0.28), however better A's results were.

As shown in figure 32, from a total number of 116 words in the dictation, Group A at T1 (12 years old and 1330 h of CLIL+FI) wrote 109.4 correct words, that is column AT1, and Group B at T2 (14 years old and 1400 h of FI) obtained 112.7 correct words, that is column BT2. On the other hand, Group A at T2 (13 years old and 1540 h of CLIL+FI) had a score of 112.2 correct written words, that is column AT2, compared to a 109.66 of total correct answers by Group B at T1, that is column BT1.

Figure 32. Average performance in the listening test at T1 and T2
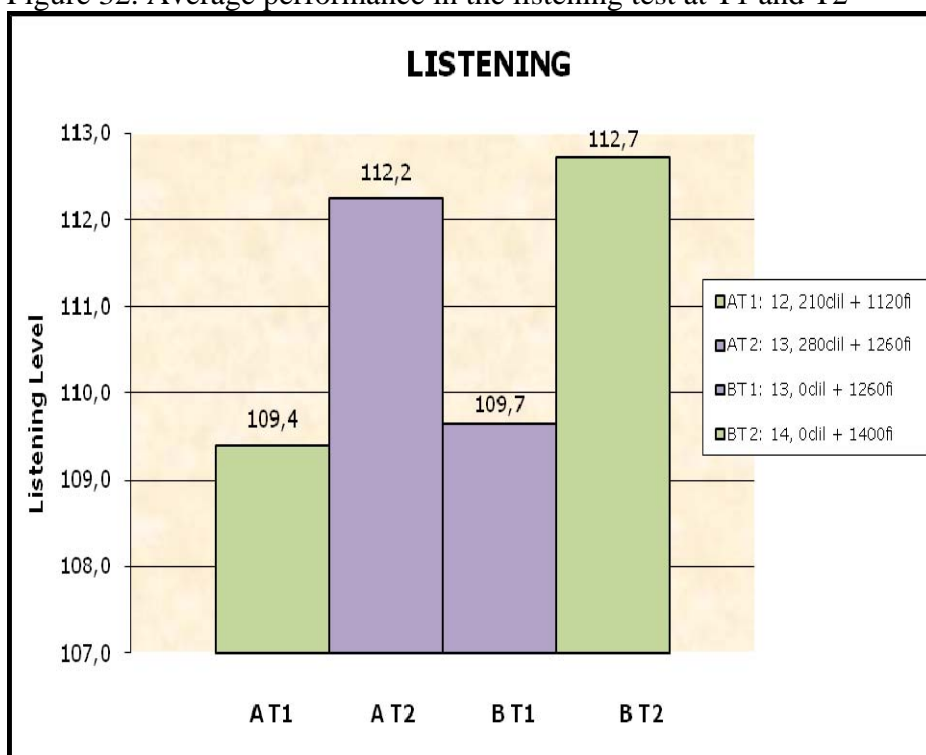
Source: personal

4.2.4 Differential Effects of Age: Lexico-Grammatical Ability

Like in the previous subsections, an ANOVA statistical analysis with the significance level set at < 0.05 was performed with the data obtained from the grammar tests. In this area (lexico-grammatical ability), the analysis of variance of the participants' performance revealed that **no significant differences** between Group A at T1 (12 years old and 1330 h of CLIL+FI) and Group B at T2 (14 years old and

1400 h of FI) were found (p=0.109) albeit group B was better than Group A. When comparing groups with the same age, that is Group A at T2 (13 years old and 1540 h of CLIL+FI) and Group B at T1 (13 years old and 1260 h of FI), **no significant differences** were found either (p=0.11) but Group A outperformed Group B.

Figure 33 below shows that from a total number of 50 questions, Group A at T1 (12 years old and 1330 h of CLIL+FI) obtained 37.1 correct answers, that is column AT1, and Group B at T2 (14 years old and 1400 h of FI) obtained an average of 38.76 correct answers, that is column BT2. On the other hand, Group A at T2 (13 years old and 1540 h of CLIL+FI) had a score of 39.82 correct answers, that is column AT2, compared to a 38.46 of total correct answers in Group B at T1, that is column BT1 (13 years old and 1260 h of FI).

Figure 33. Average performance in the lexico-grammatical tests at T1 and T2

Source: personal

To summarise the results obtained for reading, listening, and lexico-grammatical abilities, with different ages but a similar number of hours, older participants were significantly better in listening skills while no significant difference was shown in the domains of reading and lexico-grammatical ability. As for the second set of comparisons, with the same age, 13, but different number of hours, participants with extra CLIL hours (Group A) were significantly better in reading skills, while no significant difference was shown in the domains of listening and lexico-grammatical ability. However, the group with extra CLIL hours (Group A) showed a higher trend of progress.

It is worth noting that in the first set of comparisons, that is when comparing younger learners (12) with CLIL and FI to older learners (14) with FI only but a similar number of hours, older learners without CLIL instruction were better in 7 of the 11 domains analysed in spite of the lack of CLIL instruction (although they were significantly better only in accuracy in the written tests and in listening skills). This finding will be taken up in our discussion of results.

4.3 Gender Effects on Contexts of Instruction: Research Question 1c

The third section within the results chapter deals with the results concerning the issue of the impact of gender differences in a FI and a FI + CLIL context in relation to each of the different skills measured. Our aim is to either confirm or disconfirm the hypothesis that FI + CLIL female participants will not significantly outperform their male peers in the linguistic progress achieved after a one year CLIL treatment as they reportedly do in FI.

In order to address such a hypothesis a within-group comparison was established between the subsample of female versus male participants in each group.

In the following subsections the results related to the differential effects of gender on the learners' abilities in writing (qualitative measures), reading and listening are displayed leaving aside grammatical abilities measured through discrete point tests. These areas have been chosen among the whole battery of tests for their integrative nature. These integrated tests are the writing test concerning qualitative measures (4.3.1), the cloze test (4.3.2), and the listening test (4.3.3).

4.3.1 Differential Effects of Gender: Writing

In this subsection writing tests results related to qualitative measures (4.3.1.1) are shown.

4.3.1.1 Gender on writing: qualitative measures

Following the same organisation as with the results related to general abilities and age differences when measuring compositions with qualitative measures, four subsections display the results related to task fulfilment, organisation, grammar, and vocabulary. The participants are grouped according to gender, as female and male participant subgroups within the two groups established in our study: Group A (FI+CLIL) and Group B (FI).

Task Fulfilment

The difference between the male and female subgroups in Group A at T1 (12 years old and 1330 h of CLIL+FI) in the fulfilment of the task for the compositions analysed **was significant** (p=0.008) in favour of the female subgroup. One year later, Group A at T2 (13 years old and 1540 h of CLIL+FI), the difference between boys and girls **remained statistically significant** (p=0.0509) in favour of the female subgroup. On the other hand, as for the group without CLIL instruction at T1, the results showed how the difference between boys and girls **was not statistically significant** (p=0.3039). At T2, the ANOVA statistical analysis revealed that the difference **remained non-significant** (p=0.8376).

Taking into account the 6 behavioural levels on a scale of 0 (not enough to evaluate) to 5 (very good), as figure 34 shows, girls in Group A at T1 (12 years old and 1330 h of CLIL+FI) obtained an average performance of 3.3, and boys in the same group and time obtained a result of 2.5. Concerning girls in Group A at T2 (13 years old and 1540 h of CLIL+FI), they obtained a result of 3.6 whereas the boys' result was 3.

On the other hand, as for the results of groups without CLIL instruction, the boys' result in Group B at T1 (13 years old and 1260 h of FI) was 2.5, whereas the girls' average performance was 2.8. One

year later, boys in Group B at T2 (14 years old and 1400 h of FI) obtained an average of 2.9, whereas the girls' score was 2.8

Figure 34. Average performance in the writing test at T1 and T2 in relation to gender
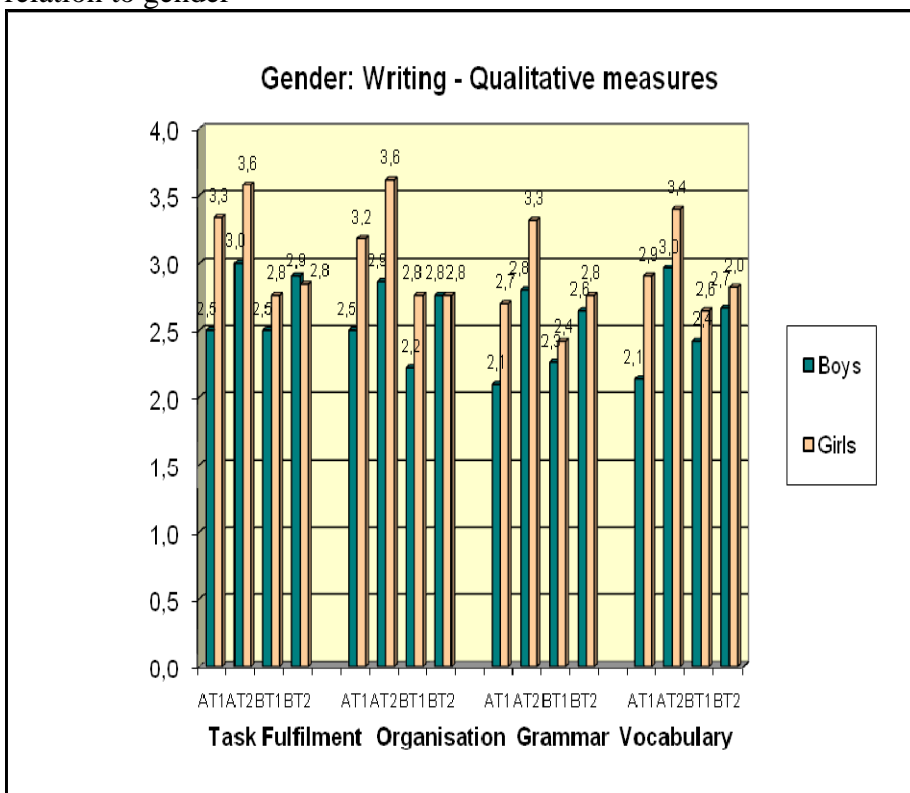
Source: personal

AT1: 12 yrs., 210 CLIL + 1120 FI
AT2: 13 yrs., 280 CLIL+1260 FI
BT1: 13 yrs., 0 CLIL+ 1260 FI
BT2: 14 yrs., 0 CLIL + 1400FI

Organisation

As far as the organisation of the compositions analysed is concerned, the girls' results in Group A at T1 (12 years old and 1330 h of CLIL+FI) were higher than boys' results in the same group and time, and the ANOVA statistical analysis revealed that the difference **was significant** (p=0.0207). The **same occured** with this group at T2 (p=0.0121). On the other hand, when comparing gender differences of

groups without CLIL instruction, girls in Group B at T1 (13 years old and 1260 h of FI) obtained higher results than boys and the difference **was also significant** (p=0.0560). Finally, girls and boys in Group B at T2 obtained exactly the **same result** (p=1).

Similarly to the previous subsection, taking into account the 6 behavioural levels on a scale of 0 (not enough to evaluate) to 5 (very good), figure 34 above shows how girls' results in Group A at T1 (12 years old and 1330 h of CLIL+FI) were 3.2. As for boys, the results in the same group at the same time were 2.5. Girls in Group A at T2 (13 years old and 1540 h of CLIL+FI) obtained a result of 3.6 whereas boys 2.9.

Without a CLIL instruction, results were the following ones: Boys in Group B at T1 (13 years old and 1260 h of FI) obtained an average of 2.2 whereas girls in the same group obtained an average of 2.8. One year later, boys in Group B at T2 (14 years old and 1400 h of FI) scored again 2.8 and girls also obtained an average of 2.8.

Grammar

In this area (accuracy), the analysis of variance of the participants' performance revealed that **no significant differences** between girls and boys in Group A at T1 (12 years old and 1330 h of CLIL+FI) were found (p=0.064). At T2 in the same group gender differences were even smaller and therefore they were **not significant** either (p=0.106). On the other hand, when comparing boys and girls without CLIL instruction at T1, Group B at T1 (13 years old and 1260 h of FI), the difference was again **not significant** (p=0.567). Finally, one year later, Group B at T2 (14 years old and 1400 h of FI), **no significant** gender differences were found (p=0.68).

As can be appreciated in figure 34 above, girls in Group A at T1 (12 years old and 1330 h of CLIL+FI) obtained an average performance of 2.7 (from the 6 behavioural levels on a scale of 0 -not enough to evaluate- to 5 -very good- used in the qualitative measures). In the same group and also at T1 boys obtained a result of 2.1 Concerning girls in Group A at T2 (13 years old and 1540 h of CLIL+FI), they obtained a result of 3.3 whereas boys result was 2.8.

On the other hand, if we now turn to the results of groups without a CLIL instruction, the boys' result in Group B at T1 (13 years old and

1260 h of FI) was 2.3, whereas the girls' average performance was 2.4. One year later the boys in Group B at T2 (14 years old and 1400 h of FI) obtained an average of 2.6, whereas the girls score was 2.8.

Vocabulary

As far as the vocabulary of the compositions analysed is concerned when using qualitative measures, the ANOVA statistical analysis revealed that the difference between boys and girls in Group A at T1 (12 years old and 1330 h of CLIL+FI) **was significant** (p=0.0175). However, in Group A at T2 (13 years old and 1540 h of CLIL+FI) the ANOVA statistical analysis revealed that gender differences **were not significant** (p=0.1841). On the other hand, when comparing boys and girls without CLIL instruction at T1, the difference was again **not significant** (p=0.374). Finally, in Group B at T2 (14 years old and 1400 h of FI) **no significant** gender differences were either found (p=0.547).

Similarly to the previous subsections, taking into account the 6 behavioural levels on a scale of 0 (not enough to evaluate) to 5 (very good), in figure 34 above it is clearly shown how Group A at T1 (12 years old and 1330 h of CLIL+FI) obtained an average performance of 2.9 and boys obtained 2.1. One year later, girls result in the same group was 3.4 whereas boys result was 3.

The results of the groups without CLIL instruction were 2.4 for boys in Group B at T1 (13 years old and 1260 h of FI) and 2.6 for girls in the same group and time. On the other hand, the boys in Group B at T2 (14 years old and 1400 h of FI) obtained a result of 2.7 whereas the girls obtained an average performance of 2.8.

4.3.2 Differential Effects of Gender: Reading

Statistical results were not performed in this particular test because, contrary to expectations, female participants did not outperform their male peers in the linguistic progress achieved after a one year FI treatment and they did so after a one year CLIL treatment. Since the results where the opposite to the ones stated in hypothesis 1c, it was illogical to carry out an ANOVA analysis.

As figure 35 clearly shows, in a cloze with 20 gaps, boys in Group A at T1 (12 years old and 1330 h of CLIL+FI) obtained an average of 14.3 correct answers whereas girls obtained an average of 14.4 correct

answers. Concerning boys in Group A at T2 (13 years old and 1540 h of CLIL+FI), they performed 15.5 correct answers, whereas girls obtained a result of 16.6 correct answers. Therefore, girls were superior in both cases.

If we now turn to the results of the groups without CLIL instruction, boys obtained slightly higher results. Boys in Group B at T1 (13 years old and 1260 h of FI) answered an average of 15 correct answers whereas girls in the same group obtained an average of 14.6 correct answers. One year later, boys in Group B at T2 (14 years old and 1400 h of FI) scored 15.8 correct answers, whereas girls answered an average of 14.4 correct answers.

Figure 35. Average performance in the cloze test at T1 and T2 in relation to gender
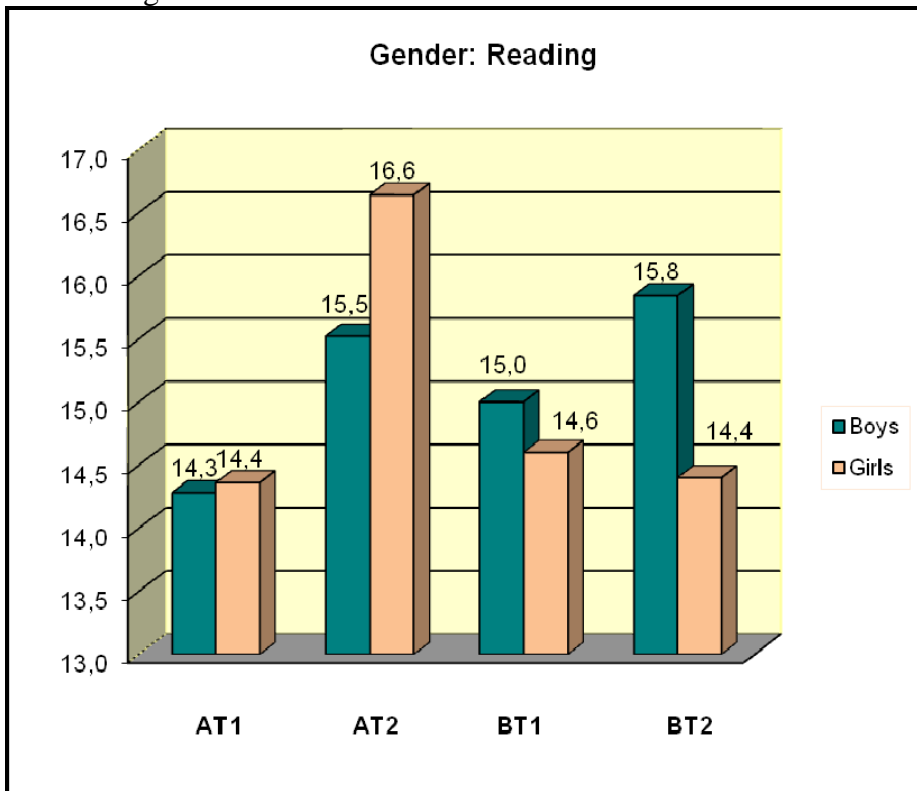
Source: personal

| AT1: 12 yrs., 210 CLIL + 1120 FI |
| AT2: 13 yrs., 280 CLIL+1260 FI |
| BT1: 13 yrs., 0 CLIL+ 1260 FI |
| BT2: 14 yrs., 0 CLIL + 1400FI |

4.3.3 Differential Effects of Gender: Listening

In the results of the ANOVA statistical analysis run with the data obtained from the dictation concerning the difference between girls and boys with CLIL instruction (Group A), girls were superior to boys both at T1 and T2. However, **in no case results were statistically significant** as p=0.695 at T1 and p=0.699 at T2. If we now turn to Group B, at T1 (13 years old and 1260 h of FI) girls were superior to boys but the difference turned out **not to be statistically significant** either (p=0.49). Concerning Group B at T2, boys were superior to girls but the difference **was again not statistically significant** (p=0.11).

As figure 36 shows, from a total number of 116 words in the dictation, boys in Group A at T1 (12 years old and 1330 h of CLIL+FI) wrote 109 correct words whereas girls obtained an average of 109.8 correct answers. Concerning boys in Group A at T2 (13 years old and 1540 h of CLIL+FI), they had a score of 112 correct written words, compared to a 112.5 of total correct answers on behalf of girls. Therefore, girls obtained again higher results in both cases.

On the other hand, if we now observe the results of groups without CLIL instruction, girls obtained higher results at T1 but boys were superior at T2. Boys in Group B at T1 (13 years old and 1260 h of FI) wrote an average of 108.1 correct words whereas girls in the same group obtained an average of 111.2 correct answers. One year later, boys in Group B at T2 (14 years old and 1400 h of FI) scored 113.4 correct answers, whereas girls answered an average of 112 correct answers.

Figure 36. Average performance in the listening test at T1 and T2 in relation to gender
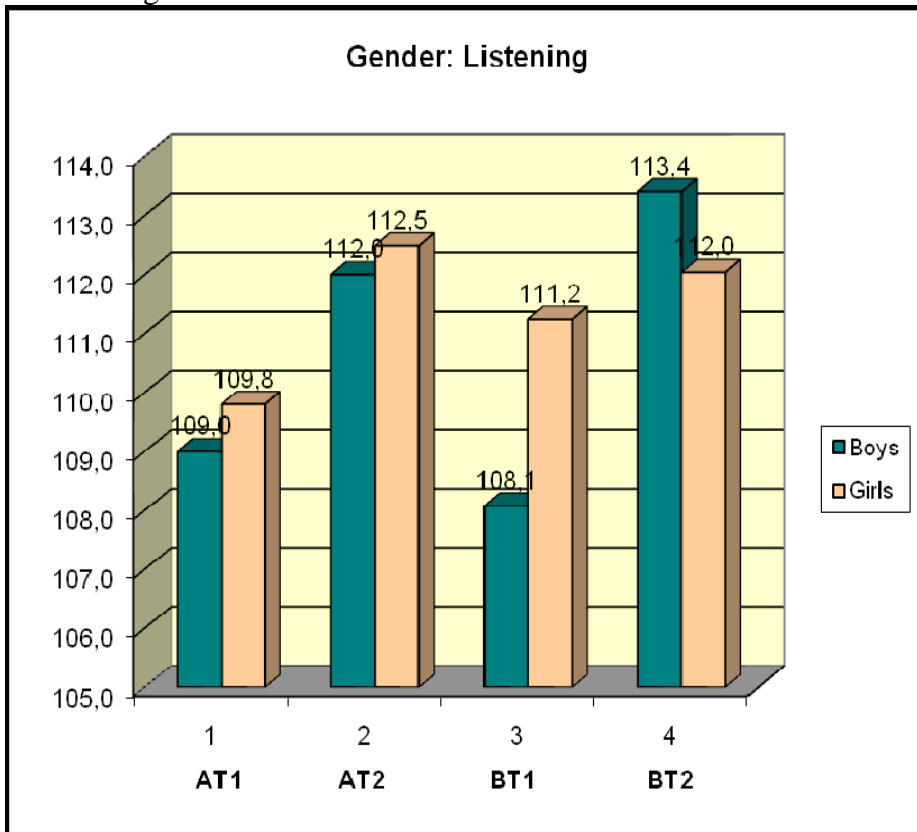
Gender: Listening

Source: personal

| AT1: 12 yrs., 210 CLIL + 1120 FI |
| AT2: 13 yrs., 280 CLIL+1260 FI |
| BT1: 13 yrs., 0 CLIL+ 1260 FI |
| BT2: 14 yrs., 0 CLIL + 1400FI |

To sum up the results regarding gender effects on EFL learners following FI+CLIL and FI only programmes, girls in Group A performed significantly better in the written test as regards task fulfilment and organisation both before and after the CLIL treatment, and in the area of vocabulary only at T1. No significant differences were found for grammar. Girls in Group B (FI only) showed significantly better performance as far as organization of the written test at T1. This was the only area were the female subgroup in Group B was better. Concerning reading and listening, the former was not

reported because, contrary to expectations, girls obtained lower results than boys in the FI group and better results in the CLIL group. The latter results, that is listening, revealed the superiority of girls in the CLIL group. However, girls' progress in the FI group was lower than boys' progress.

4.4 Summary

In this subsection a summary of the present results' chapter is offered. Results have been reported and organized in different sections, each dealing with one of the three research subquestions established in relation to the main question addressed in our study, namely:

How does context of learning affect the linguistic development of young bilingual secondary education EFL leaners when contrasting a group experiencing FI only and a group experiencing FI in combination with CLIL?

Firstly, section 4.1 dealt with the first subquestion, the development of general language skills, and included the results concerning it as referred to by Research Question 1a and the hypothesis derived from it:

RQ1a) When contrasting the differential effects of two different programmes, a FI only and a FI+CLIL, that is with an additional CLIL component, which programme results in linguistic benefits if any and which skills benefit the most if any?

H1a) When contrasting the differential effects of two programmes on learners' linguistic progress, a FI programme and a FI+CLIL with the additional hours, the group in the FI+CLIL will improve significantly more than the other especially in receptive skills.

Subsection 4.1.1 concerned results related to production in writing skills (composition), which in turn was divided into writing tests results related to quantitative measures (4.1.1.1) and writing tests results related to qualitative measures (4.1.1.2). After this, subsection 4.1.2 refered to results related to comprehension in reading skills (cloze). Then, subsection 4.1.3 dealt with results related to comprehension in oral skills (listening test: dictation). Finally, subsection 4.1.4 concerned results related with lexico-grammatical ability.

In the first place, quantitative results related to production in writing skills allowed us to identify the following trends. Firstly, **Group A (FI+CLIL)** outperformed Group B (FI) **only significantly in the domain of accuracy**. However they showed a tendency towards surpassing them as far as syntactic complexity went. Another relevant result was the fact that in the domain of lexical and syntactic complexity Group A (FI+CLIL) had a higher onset level, whereas in the domain of accuracy and fluency, on the contrary, they had a lower onset level.

When looking at each context in turn, on the one hand, Group A (FI+CLIL) after the CLIL treatment wrote shorter texts, which nonetheless were more accurate, lexically richer and syntactically complex. On the other hand, after the treatment, Group B (FI) also wrote shorter texts, more accurate and lexically richer, however less syntactically complex.

Qualitative results related to production in writing skills could be summarised as follows. When analysed with qualitative measures, Group A (FI+CLIL) outperformed Group B (FI) in all areas. However, results did not reach statistical significance. More specifically, Group A (FI+CLIL) consistently tended to write a better organised, more accurate, lexically richer and more purposeful composition. It is also interesting here to highlight that Group A (FI+CLIL) always had a higher onset level except for the domain of vocabulary although being a year younger.

Finally, when looking at results related to comprehension in reading skills (cloze), comprehension in oral skills (listening test: dictation), and results related to lexico-grammatical ability, we could summarise this second set of tests as follows: both groups improved their reading, listening and lexico-grammatical abilities. However, **Group A (FI+CLIL) was significantly faster in reading and lexico-grammatical abilities** than Group B (FI). Another result worth highlighting was the fact that in the three domains Group A (FI+CLIL) started with a lower onset level.

Secondly, section 4.2 dealt with the second research subquestion related to whether changes occured irrespective of age differences, and included the results concerning this issue as referred to by Research Question 1b and the hypothesis derived from it:

RQ1b) Does age have an impact on the potential benefits learners may achieve in either of the two contexts compared: FI and FI+CLIL?

H1b) At different ages and with a similar number of hours, younger learners receiving FI+CLIL will benefit more than elder learners only receiving FI because 'the earlier the better' for CLIL instruction. On the other hand, at the same age and with a different number of hours, learners with a higher total number of hours receiving FI+CLIL will benefit more than learners with fewer hours only receiving FI.

Subsection 4.2.1 refered to results related to production in writing skills (composition), which in turn was divided into writing tests results related to quantitative measures (4.2.1.1) and writing tests results related to qualitative measures (4.2.1.2). After this, subsection 4.2.2 refered to results related to comprehension in reading skills (cloze). Then, subsection 4.2.3 dealt with results related to comprehension in oral skills (listening test: dictation). Finally, subsection 4.2.4 concerned results related to lexico-grammatical ability.

To summarise the first set of comparisons between both groups at different ages but a similar number of hours by means of quantitative measures, **significantly better results yielded for the younger group** (FI+CLIL) albeit having had fewer hours of exposure to English in the domain of **syntactic complexity**. Thus, they used more subordination. On the other hand**, significantly better results yielded for the older group** (FI) in the area of **accuracy**: they made fewer mistakes.

Concerning the summary of the second set of comparisons, that is groups at the same age but different number of hours, a clear tendency in favour of the group with more hours (extra CLIL hours) in the domain of syntactic complexity, lexical complexity and accuracy was noted. Thus, although not significantly better, at the same age the group with extra CLIL hours used more subordination, made fewer mistakes and was lexically richer. However, it was less fluent than the group with the same age but fewer hours.

On the other hand, when analysed by means of qualitative measures, results on the written ability of our two groups of participants turned out not to be statistically significant for the first issue related to age. That is, when measured at different ages but a similar number of hours, their results were not significantly different even if there was a tendency for the younger participants to surpass the older in task

fulfilment and organisation, but not in the areas of grammar and vocabulary. In contrast, **when measured at the age of 13, the results of participants with higher number of hours, FI+CLIL, were statistically better for the four areas**. That is, at the same age but with extra CLIL hours participants' compositions were more purposeful, organised, accurate, and lexically richer.

Finally, to summarise the results obtained for reading, listening, and lexico-grammatical abilities, with different ages but a similar number of hours, **older participants were significantly better in listening skills** while no significant difference was shown in the domains of reading and lexico-grammatical ability. As for the second set of comparisons, with the same age, 13, but different number of hours, **participants with extra CLIL hours (Group A) were significantly better in reading skills**, while no significant difference was shown in the domains of listening and lexico-grammatical ability. However, the group with extra CLIL hours (Group A) showed a higher trend.

It is interesting that, as noted above, in the first set of comparisons, that is when comparing younger learners (12) with CLIL and FI to older learners (14) with FI only but a similar number of hours, older learners without CLIL instruction were better in 7 of the 11 domains analysed in spite of the lack of CLIL instruction (although they were significantly better only in accuracy in the written tests and in listening skills).

In the third place, section 4.3 dealt with the third research subquestion, the issue of the impact of gender differences in a FI and a FI + CLIL context in relation to each of the different skills measured in this study. These implied results concerning Research Question 1c and the hypothesis derived from it:

RQ1c) Does gender have an impact in the benefits obtained with either a FI context of learning, or a FI+CLIL context?

H1c) FI + CLIL female participants will not significantly outperform their male peers in the linguistic progress achieved after a one year CLIL treatment as they reportedly do in FI.

Concerning the subsections within 4.3, differences in progress in each of the two contexts of learning in relation to this issue were measured for three linguistic abilities: writing (subsection 4.3.1), reading (subsection 4.3.2), and listening (subsection 4.3.3).

To sum up results related to this issue, **girls in Group A (FI + CLIL) performed significantly better in the written test as regards task fulfilment and organisation both before and after the CLIL treatment, and in the area of vocabulary only at T1**. No significant differences were found for grammar in the written tests. **Girls in Group B (FI only) showed significantly better performance as far as organization of the written test at T1**. This was the only area were the female subgroup in Group B was better. Concerning reading and listening, no significant differences were found in any case.

5. DISCUSSION

In this chapter the results earlier presented are used to address the three subhypotheses (hypothesis 1a, 1b, and 1c) put forward in this research study concerning the effects of CLIL programmes. The results are also discussed in the light of previous research on the issues examined.

The organisation of the chapter is as follows: Section 5.1 discusses the results obtained in order to address the first subhypothesis of the study. Hypothesis 1a states that learners following an English programme in which CLIL instruction is added to conventional formal instruction would significantly outperform learners following conventional formal instruction, and that this would be most clearly so as regards their receptive skills, that is reading and listening. The review of studies related to the Canadian immersion programmes (Genesee, 1987; Harley *et al.,* 1990; Swain and Lapkin, 1990; Johnson and Swain, 1997; Wechse, 2002; Lyster, 2007; among others) and to European CLIL programmes (Dalton-Puffer, 2007, among others) revealed CLIL advantages concerning receptive skills, however also a weakness in the CLIL learners' productive skills, spoken and written. For such a reason in the present study it has been hypothesized that the group in the FI+CLIL would show progress after a year of treatment significantly more than the FI only group, especially in their receptive skills.

Section 5.2 deals with the second hypothesis (1b) of the study which is doublefold. Firstly it states that at different ages and with a similar number of hours, younger learners receiving FI+CLIL would make greater progress in their English than elder learners only receiving FI because 'the earlier the better' for CLIL instruction. Secondly, it also hypothesized that at the same age and with a different number of hours, learners with a higher total number of hours, hence receiving FI+CLIL, would benefit more than learners with fewer hours and only receiving FI in English.

The third section, section 5.3, addresses the third hypothesis (1c) regarding the impact of gender differences in both a FI and a FI + CLIL context in relation to each of the different skills measured in this study. Finally, section 5.4 provides a summary of the chapter.

Chapter 5

5.1 Linguistic Progress and Skill Development

Hypothesis 1a in this study stated that when contrasting the differential effects of two different programmes on learners' linguistic progress over an academic year, a FI programme and a FI+CLIL with additional hours, the group in the FI+CLIL would improve significantly more than the group with only FI especially in their receptive skills.

Table 18 below displays the results obtained through the statistical analyses presented in chapter 4 for a better understanding of the ensuing discussion. The left hand-side column lists the different skills gauged, writing, reading, listening and lexico-grammatical abilities. The quantitative and qualitative measures used in the case of writing are broken down as syntactic complexity, lexical complexity, accuracy, and fluency (quantitative measures); and task fulfilment, organisation, grammar, and vocabulary (qualitative measures) respectively. The central column shows the results obtained by Group A, experiencing a FI+CLIL context of learning, and the right-hand side column those by Group B, experiencing a FI only. The upper boxes in these two columns include the number of hours of instruction accumulated by each group at each data collection time, T1 and T2. The boxes to the right of each skill state the results obtained for each measure at both data collection times. It must be remembered that bold results are those that reach significant difference in that the difference in progress between Group A and Group B is significant in favour of Group A as it turns out that there is no significant difference in favour of Group B.

A quick overview of the relevant key findings in the table, as they are listed on the left-hand side column, from top to bottom, are as follows. Group A's results as far as written accuracy show a progress of 0.042 over one academic year. This is significantly higher than Group B's results which only improved 0.006 from T1 to T2. As for reading, Group A obtained a 1.69 figure which is significantly higher than Group B's 0.22 improvement. In the case of listening, Group A's progress reached a value of 2.8 whereas Group B's progress reached 3.1. This, together with lexical complexity in writing, is the only skill where Group B outperfoms Group A, albeit non-significantly (marked up as italics in table 18). Finally, in the area of lexico-grammatical ability Group A's results show a progress of 2.72 whereas Group B's progress reached a figure of 0.3.

Table 18: Skill Results per Context (progress)

	CLIL(Group A) AT1: 12 yrs. 210CLIL+1120FI=1330h AT2:13 yrs. 280CLIL+1260FI=1540h	**FI(Group B)** BT1:13 yrs. 0CLIL+1260FI=1260h BT2: 14 yrs. 0CLIL+1400FI=1400h
Writing		
Syntactic complexity	AT1: 0.40 ;A T2: 0.39 (+0.01)	BT1: 0.47 ; BT2: 0.49 (-0.02)
Lexical complexity	AT1: 6.50 ; AT2: 6.71 (+0.21)	BT1: 6.31 ; BT2: 6.73 *(+0.42)*
Accuracy	**AT1: 0.120 ; AT2: 0.078 (+0.042*)**	BT1: 0.092 ; BT2: 0.086 (+0.006)
Fluency	AT1: 146.2 ; AT2: 145.1 (-1.1)	BT1: 149.1 ; BT2: 144.7 (-4.4)
Task Fulfilment	AT1: 2.92 ; AT2: 3.29 (+0.37)	BT1: 2.63 ; BT2: 2.87 (+0.24)
Organisation	AT1: 2.84 ; AT2: 3.24 (+0.4)	BT1: 2.49 ; BT2: 2.76 (+0.27)
Grammar	AT1: 2.40 ; AT2: 3.06 (+0.66)	BT1: 2.34 ; BT2: 2.70 (+0.36)
Vocabulary	AT1: 2.52 ; AT2: 3.18 (+0.66)	BT1: 2.53 ; BT2: 2.74 (+0.21)
Reading	**AT1: 14.3 ; AT2: 16.1 (+1.69*)**	BT1: 14.6 ; BT2: 14.8 (+0.22)
Listening	AT1: 109.4 ; AT2: 112.2 (+2.8)	BT1: 109.7 ; BT2: 112.7 *(+3.1)*
Grammar (Lexico-grammatical ability)	**AT1: 37.1 ; AT2: 39.8 (+2.72*)**	BT1: 38.5 ; BT2: 38.8 (+0.3)

Source: personal
Note: the higher the value for syntactic complexity and accuracy the lower the competence level
T1: first data collection time   T2: second data collection time
A: Group A   B: Group B

It is worth having a more detailed appraisal of these results for each of the skills. Firstly, the **quantitative results obtained for production in writing as shown in a composition task indicate that** the group with

the CLIL treatment, that is **Group A (FI+CLIL), significantly outperformed** the group following formal instruction classes without a CLIL treatment, that is **Group B (FI), in the domain of accuracy**. In addition they also tended to outperform them as far as syntactic complexity went. These results were not the same in the case of the other quantitative measures, that is lexical complexity and fluency. This significant leap forward in accuracy is most relevant in the case of the CLIL students as CLIL courses are said to focus on meaning rather than on form. Hence only the extra amount of practice or transfer of skills can explain them, as is discussed further below. Secondly, when the learners' written production was analysed with qualitative measures, Group A (FI+CLIL) outperformed Group B (FI). However results did not reach statistical significance.

When we try and picture out these results in a more global manner we see that **Group A (FI+CLIL)** consistently tended to **write a better organised, more accurate, lexically richer and more purposeful composition.**

We now turn to results related to **reading comprehension**, as tested by means of a **cloze test.** The results summarised in table 18 show that the group with CLIL instruction, that is **Group A (FI+CLIL), performed significantly better than** the group with formal instruction only, that is **Group B (FI)** in the test.

When looking at **listening comprehension**, as tested by means of a **dictation** given by the teacher, our results reveal that this is a domain in which **Group A did not outperform Group B**. Indeed, both groups showed improvement at T2, however, and contrary to what we had hypothesized, Group B (FI) even showed a tendency towards higher results than the group following a CLIL treatment, that is Group A (FI+CLIL).

Finally, when turning to the last linguistic domain scrutinized, **grammar**, as tested through a **fill in the gaps task,** and an **error correction task**, results related to **lexico-grammatical ability** again indicated that **Group A (FI+CLIL) performed significantly better than Group B (FI)**.

In the light of these findings summarised we can address the first hypothesis (1a). Our results show improvement in written production as Group A significantly outperforms Group B in accuracy, and tends to also outperform it in syntactic complexity

and in the whole set of qualitative measures (task fulfilment, organisation, grammar and vocabulary). This is also true in the case of reading comprehension and lexico-grammatical competence, where Group A is significantly better than Group B. It is only in the domain of listening comprehension that Group B tends to perform better than Group A.

These findings allow us to state that our first hypothesis (1a) concerning the greater progress in receptive skills for the CLIL group is only partially confirmed. Indeed, whereas reading improves significantly, listening does not. Furthermore, our findings show a significant improvement in productive skills, something which we had not hypothesised, as writing and particularly accuracy, significatively progress and so do lexico-grammatical abilities. This is in contrast with findings published in previous studies, as will be discussed further below.

Furthermore, although it is true that significant benefits do not accrue in all skills and measurements, it is also true that **tendencies in the differential progress between both groups can allow us to establish the benefits of the school's CLIL programme**. On the basis of the existing results it can be argued that an academic year might not have been sufficient for learners to register more substantial benefits and that a longer course of study might eventually show that tendencies would become significant differences.

Hence we would posit that **our results confirm the effectiveness of a CLIL programme. This is something which previous research conducted in educational contexts similar to ours had also confirmed** (in Catalonia, as in the rest of Spain, the foreign target language, English, is hardly ever used and heard outside the school setting).

Several general considerations concerning such general progress made by the CLIL group should be made here:

Firstly, when we review the research conducted in such settings, and more specifically in other bilingual contexts, such as the Basque Country, studies which state detailed results for each skill seem to report similar findings to ours (Lasagabaster, 2008; Muñoz and Navés, 2007; Ruiz de Zarobe, 2008; Victori and Vallbona, 2008; Villareal and Gacía Mayo, 2009, Navés and Victori, 2010). This is in contrast to

studies from Europe, as CLIL students here tend to show an improvement not only in receptive but also in productive skills. It is interesting to highlight that in Lasagabaster's study (2008), like in ours, CLIL younger groups also scored lower than non-CLIL one year older groups in the listening tests (in the present study CLIL learners scored lower than FI learners for the listening ability not only when they were younger but also when both groups shared the same age).

Second, it has to be noted that the CLIL group has the "extra" hours (and hence more hours), as it is often the case in most research studies since it is difficult to have valid control groups of formal instruction with the same number of hours as the CLIL group. However, **it is of utmost importance and worth noting that in spite of Group A having a higher number of hours (70) than Group B, when measured at T1 they do not always outperform Group B at that point**. For example, whereas it is true that as far as written competence went, in the domain of lexical and syntactic complexity Group A (FI+CLIL) started at a higher onset level, in contrast, in the domain of accuracy, vocabulary and fluency, on the contrary, they had a lower onset level, just as in the domains of reading, listening comprehension and lexico-grammatical ability, in which Group A had lower values at T1 than Group B.

Hence, it could be argued that in such domains in which Group A is lower at T1 quantity of hours is not what matters, but other factors such as quality, readiness to learn, motivation, etc. A possible explanation would be the maturational constraints of Group A who is a year younger than Group B, an issue further discussed when we tackle our second hypothesis. What is interesting here is that even in some of those domains in which Group A had lower onset levels, such as reading and lexico-grammatical ability, they still outperformed Group B at T2, after 70 extra CLIL hours + 140 FI hours.

We now turn to a second set of considerations concerning the specific language skills analysed. We will first address the issue of accuracy, that is grammatical ability. The significant improvement found in the area of accuracy in the writing skill and in lexico-grammatical abilities is a rather surprising finding. Opposite results were obtained by the empirical studies carried out in Canada and Europe. In Canada, this led to a concern for fostering accuracy, as proposed by Harley *et al.* (1990), and more recently Lyster (2007) in Canada, or Pérez-Vidal (2007c) in Europe (explained in detail in section 2.3.2, *CLIL Research*). More

specifically, these authors have proposed balancing the experiential and analytical approaches, that is introducing approaches that focus on form in order to compensate for the low level in immersion/CLIL students' accuracy. Therefore, the fact that accuracy in the writing skill and lexico-grammatical abilities in general showed significant improvement in the case of our participants who followed a CLIL methodology might be explained by transfer of knowledge and skills from a FI context to a CLIL context, since they are "often" and "very often" practised in the FI context. This idea is further developed below.

Secondly, the decrease in listening abilities as regards CLIL participants may be explained by the test they were administered: a dictation. As explained in section 3.3.5 (Analyses/Measures), it has to be born in mind that the evaluation of the dictations caused some scoring problems: to be scored as correct, a word had to provide strong evidence of the learner's having heard and recognised the word, even if he/she could not spell it correctly. That is, spelling was not required for a response to be scored as correct since it was meaning comprehension rather than form (spelling) that was meant to be tested. It was not enough for our participants to attempt a representation of the sounds that they heard, without making sense of those sounds. The difficulty was to really know whether, although having a correct spelling (or not), the learner had recognised the word giving to those particular sounds the sense that the specific context really meant. For example, "(...) the british explorer (...)", was scored as correct because the meaning of *British* was understood. However, in "(...) he was to late (...)" we did not really know wether the mistake was due to the fact that the learner did not know the meaning of *too* and thus its correct comprehension was impossible, or simply because he/she did not know how to spell it. In this particular case it was considered a mistake.

We would now like to turn to an interpretation of our results in the light of the theories and models presented in chapter 2, more specifically the views on the role of practice and skill transfer models.

As regards the issue of transfer of knowledge and skills, one of the main differences between one learning context (CLIL) and the other (FI) is the type and amount of practice that learners engage in while in the classroom. We must remember that our setting is one where little

practice can be expected outside the classroom walls[16]. The study of practice in SLA literature has been recently retackled, especially with DeKeyser's 2007(a) monographic book on practice, claiming that not only the amount of practice but also the type is crucial to language learning. As explained in section 2.1.1.4 (*The Cognitive Period: Skill Acquisition Theories*), previous studies on practice had assumed a dual division between input practice and output practice. Muranoi (2007) defines output practice as "any activity designed to provide L2 learners with opportunities to produce output". Comprehension-based approaches to input practice argue that "learners will be able to develop their L2 if they are allowed to listen (i.e., understand) to the L2" (Salaberry, 1997). So, in these authors' view, input practice would include any type of activity directed at L2 learners' comprehension of the input provided. Two confronted positions have developed over the years on this issue. VanPattern and colleagues, defending the position within the input processing studies that comprehension practice alone is enough to bring about significant development, not only in comprehension but also in production (VanPatten and Cadierno, 1993a). On the contrary, the skill-specificity theory approach, represented by DeKeyser and Sokalski (1996) and DeKeyser (1997), which replicated VanPatten and Cadierno's (1993a) study, reached the conclusion that "input practice is better for comprehension and output practice for production" (DeKeyser and Sokalski, 1996: 635).

Thus, we can expect that in learning contexts where sufficient input practice is provided, comprehension skills (both reading and listening) will improve after a certain period of time. What seems not so straightforward is whether production skills (speaking and writing) will also improve in learning contexts where only comprehension practice is provided (with very limited or almost inexistent production practice).

In our research study, as explained in section 3.3.3 (Treatment), each of both contexts involve particular patterns of language skills practiced. As can be seen in table 19 below, the boxes to the right of each skill state the amount of practice learners experience in the corresponding skill according to each context of learning (CLIL or FI). On the one hand, in our FI context the writing and reading skills are mostly practiced, at least once a week. Listening is practiced in a limited way, through teacher talk, and lexico-grammatical abilities are very often

---

[16] However, most of the learners may be following extra-curricular classes as was mentioned in section 3.3.1 (*Context and Participants*).

practiced, at least once a week and very often in every single session. Oral production practice is limited. On the other hand, in the CLIL context, whereas reading is highly practiced in every class session with a considerable amount of authentic texts unusual in FI, practice in listening and writing abilities is limited to teacher talk and very short exercises. Furthermore, lexico-grammatical abilities are hardly ever practiced[17].

Table 19: Skill Practice

SKILL	CLIL Context	FI Context
Writing	**Seldom practiced** (short exercises)	**Often practiced** (at least once a week)
Reading	**Highly practiced** (every class session)	**Often practiced** (at least once a week)
Listening	**Seldom practiced** (teacher talk)	**Seldom practiced** (teacher talk)
Grammar (Lexico-grammatical ability)	**Very seldom practiced** (once a month)	**Very often practiced** (at least once a week)

Source: personal

In addition to the impact of practice within contexts we should take into account the transferability of practice occurring in a particular context onto another. As Group A in our study experiences a CLIL context together with a FI setting, their ability to transfer linguistic skills and competences learnt in the FI classes to the communication situations encountered in CLIL sessions will be at play. This might explain why, although writing skills and lexico-grammatical abilities are hardly practiced in the CLIL sessions, Group A participants obtain high results in these domains of competence. It could be argued that the amounts of writing and grammar practice typical of FI are used in the CLIL context and what students proceduralize in a FI context is automatized while in the CLIL setting. That is, the accumulated experience of FI is what may

---

[17] It must be noted that oral production, although not contemplated in our study, is also important. In CLIL contexts it is practiced to a higher degree than in FI classes as learners always use the target language to interact with the teacher. Peer work, however, seldom involves it. In contrast, one might say that in FI peer work is controlled for the language being used when carrying out tasks, but teacher to learners conversations may show high amounts of code-switching from English to Catalan/Spanish.

play a major role in the relative benefits of a CLIL context enjoyed by the learner. We now turn to the second hypothesis.

5.2 Linguistic Progress and Age

The second hypothesis (1b) enquired whether age had an impact on the potential benefits learners may achieve in either of the two contexts compared: FI and FI+CLIL. It was hypothesized that: 1) at different ages and with a similar number of hours, younger learners receiving FI+CLIL would benefit more than elder learners only receiving FI because 'the earlier the better' for CLIL instruction; 2) at the same age and with a different number of hours, learners with a higher total number of hours receiving FI+CLIL would benefit more than learners with fewer hours only receiving FI.

Table 20 below displays the results obtained through the statistical analyses presented in chapter 4 for a better understanding of the ensuing discussion. As in the previous section, the left hand-side column lists the different skills gauged, writing, reading, listening and lexico-grammatical abilities. The quantitative and qualitative measures used in the case of writing are broken down as syntactic complexity, lexical complexity, accuracy, and fluency (quantitative measures); and task fulfilment, organisation, grammar, and vocabulary (qualitative measures). The central column shows the results obtained when comparing groups with different ages but a similar number of hours. Therefore it shows the results by Group A (FI+CLIL) at T1 (12 years old and 1330h) and the results by Group B (FI) at T2 (14 years old and 1400h). The right-hand side column shows the results obtained when comparing groups at the same age but a different number of hours. That is the results obtained by Group A (FI+CLIL) at T2 (13 years old and 1540h), and the results by Group B (FI) at T1 (13 years old and 1260h). The upper boxes in these two columns include the number of hours of instruction accumulated by each group at each data collection time, T1 and T2. The boxes to the right of each skill state the results obtained for each measure and each group at each data collection times. It must be remembered that bold results are those that reach significant difference when comparing groups (p value < 0.05).

A quick overview of the key findings in table 20 is as follows, as they are listed on the left-hand side column, from top to bottom. Concerning the first set of comparisons**, at different ages but a similar number of hours,** in the **writing** area, **younger learners with CLIL** instruction

obtained **significantly higher results** than older learners with FI only in the domain of **syntactic complexity**. In the domain of **written accuracy and listening**, **older learners with FI only** were **significantly better than younger learners with CLIL**. Concerning the second set of comparisons, **at the same age and different number of hours**, the group with **more hours (FI+CLIL)** obtained **higher results than** the group with **fewer hours (FI) in all areas but written fluency. Results were significant in all written holistic measures and in the reading skill.**

Table 20: Skill Results per Context and Age (p value)

	Diff age, similar n hours AT1: 12 yrs. 210CLIL+1120FI=1330h BT2: 14 yrs. 0CLIL+1400FI=1400h	Same age, diff n hours AT2: 13 yrs. 280CLIL+1260FI=1540h BT1: 13 yrs. 0CLIL+1260FI=1260h
Writing Syntactic complexity	**AT1: 0.40 ; BT2: 0.49 (p=0.028*)**	AT2: 0.39 ; BT1: 0.47 (p=0.089)
Lexical complexity	AT1: 6.50 ; BT2: 6.73 (p=0.219)	AT2: 6.71 ; BT1: 6.31 (p=0.08)
Accuracy	**AT1: 0.120 ; BT2: 0.086 (p=0.012*)**	AT2: 0.078 ; BT1: 0.092 (p=0.19)
Fluency	AT1: 146.2 ; BT2: 144.7 (p=0.855)	AT2: 145.1 ; BT1: 149.1 (p=0.638)
Task Fulfilment	AT1: 2.92 ; BT2: 2.87 (p=0.819)	**AT2: 3.29 ; BT1: 2.63 (p=0.001*)**
Organisation	AT1: 2.84 ; BT2: 2.76 (p=0.694)	**AT2: 3.24 ; BT1: 2.49 (p=0.0005*)**
Grammar	AT1: 2.40 ; BT2: 2.70 (p=0.17)	**AT2: 3.06 ; BT1: 2.34 (p=0.001*)**
Vocabulary	AT1: 2.52 ; BT2: 2.74 (p=0.294)	**AT2: 3.18 ; BT1: 2.53 (p=0.002*)**
Reading	AT1: 14.3 ;BT2: 14.82 (p=0.48)	**AT2: 16.08 ; BT1: 14.6 (p=0.02*)**
Listening	**AT1: 109.4 ;BT2: 112.7 (p=0.003*)**	AT2: 112.2 ; BT1: 109.7 (p=0.28)

Grammar (Lexico-grammatical abil.)	AT1: 37.1 ; BT2: 38.8 (p=0.109)	AT2: 39.8 ; BT1: 38.4 (p=0.11)

Source: personal
Note: the higher the value for syntactic complexity and accuracy the lower the competence level
T1: first data collection time   T2: second data collection time
A: Group A (CLIL+FI)   B: Group B (FI)

It is worth having a more detailed account of the previous results. We first look at the first set of comparisons between both groups at **different ages but a similar number of hours.** As far as **writing** is concerned we have two contrasting results. On the one hand, when measured **quantitatively,** the **younger group (FI+CLIL)** shows **significantly better results** in the domain of **syntactic complexity** as they used more subordination. On the other hand**, significantly better results were yielded by the older group (FI)** in the area of **accuracy: they made fewer mistakes.** This might be explained by maturational differences, as is discussed below. When analysed by means of **qualitative measures**, results on the written ability of our two groups of participants turned out not to be statistically significant. That is, when measured at different ages but a similar number of hours, their results were not significantly different even if there was a **tendency for the younger participants (FI+CLIL) to surpass the older ones (FI) in task fulfilment and organisation**, but not in the areas of grammar and vocabulary.

To summarise the results obtained for **reading** skills, **no significant differences** were shown in the domains of reading. As far as **listening** abilities, **older participants (FI) were significantly better in listening skills than younger ones (FI+CLIL)**. Hence it can be argued here again that age matters for progress in listening because the more mature learners are, the greater progress they achieve.

Finally, concerning **lexico-grammatical abilities**, there was no significant difference between both groups.

Concerning the second set of comparisons, groups having the **same age but a different number of hours**, we also obtained two contrasting findings in the area of writing. On the one hand, when measured **quantitatively,** we saw a clear tendency in favour of the group with

more hours (extra CLIL hours) in the domain of syntactic complexity, lexical complexity and accuracy. Thus, **at the same age Group A (FI+CLIL) used more subordination, made fewer mistakes and was lexically richer**. On the other hand, the group with fewer hours, **Group B (FI), obtained better results in fluency than Group A (FI+CLIL)**.

When analysed by means of qualitative measures, the results of participants with our higher number of hours, CLIL+FI, were statistically better for the four areas. That is, **at the same age but with extra CLIL hours ours participants' compositions were more purposeful, organised, accurate, and lexically richer.**

Concerning **reading** skills, **participants with extra CLIL hours (Group A) were significantly better.** As far as **listening** abilities, no significant difference was shown in this area.

Finally, concerning **lexico-grammatical abilities**, no significant difference was found between Group A (FI+CLIL) and B (FI) at the same age, 13, but different number of hours.

A general overview of these results offers two sets of findings. With the first comparison, that is **when comparing younger learners (12) with CLIL and FI to older learners (14) with only FI but a similar number of hours, an older age seems to make a difference** as **older learners without CLIL instruction were better in 7 of the 11 domains analysed in spite of the lack of CLIL instruction**: written lexical complexity, written accuracy, written grammar, written vocabulary, reading, listening, and lexico-grammatical ability. These differences were significatively higher only in written accuracy and in listening skills. Hence, it can be stated that **an older age counterbalances the positive impact of a CLIL programme in our study.** With the second comparison, that is **when comparing learners at the same age (13) but different number of hours**, 1540h (FI+CLIL) versus 1260 (FI), learners with a higher total number of hours receiving FI+CLIL obtain higher results than learners with fewer hours only receiving FI in all areas but fluency. **Therefore, the total number of hours and the treatment seem to make a difference since learners with more hours and CLIL instruction are better.**

In the light of these findings we can address the second hypothesis (1b). According to the first part of our second hypothesis, younger participants enrolled in the CLIL programme would benefit more than

older students only receiving FI because 'the earlier the better' for CLIL instruction. The results show that **the first part of our hypothesis can not be confirmed because despite the statistically significant difference in favour of the CLIL younger students in the written syntactic complexity area, there are also statistically significant differences in favour of the FI older students in written accuracy and in the listening skill. In the same vein, sometimes younger CLIL students show a tendency towards improvement, but the same occurs with older FI participants. We might interpret this on cognitive grounds and justify the result by the fact that FI with its high amount of grammar practice is beneficial to this age range learners.**

Contrary to what happens with the first part, the second part of our hypothesis can be confirmed because at the same age and with a different number of hours, learners with a higher total number of hours receiving FI+CLIL benefit more than learners with fewer hours only receiving FI. Our findings show that FI+CLIL participants scored higher than FI learners in all areas but written fluency, as can be seen in table 20 above. Although showing a tendency towards improvement in some areas, differences turned out to be significant in all holistic measures concerning writing and in reading skills. This finding was quite expected since not only CLIL students had enjoyed 280 hours of extra exposure, but also these were CLIL hours.

Some general considerations concerning such findings should be made here:

The first consideration is that a possible explanation for not finding a clear difference when comparing learners at different ages but with a similar number of hours could be the existing empirical evidence from the field of SLA which states that it is by reaching an upper intermediate level through formal instruction that CLIL approaches to education can be beneficial to learners and help them improve receptive skills, general fluency, vocabulary, and self-regulatory abilities. Therefore, it could be argued that **the younger participants in our research study had not reached this threshold level yet in order to draw on the learning opportunities they have in the FI context for greater improvement through CLIL.**

Another consideration concerns **maturational constraints**. The younger group (FI+CLIL) had the same number of hours as the older group (FI) but their age affects how they benefit from these hours. As explained in detail below, studies completed in formal learning contexts have recurrently shown that **older learners are faster and better learners than younger ones in most aspects of acquisition due to their higher capacity for abstraction and logical thinking**.

Now we turn to discuss these topics in the light of the theories and models concerning age presented in chapter 2 section 2.1.4 (*Individual Differences: Age*)

As explained in section 2.1.4 (Individual Differences), one of the most topical issues in many European educational systems is whether it is better to start foreign language teaching at an early age, or whether it is better at a later stage –without establishing an early first contact with the foreign language. Although research undertaken in naturalistic settings confirms that young starters ultimately achieve higher competence in the L2, studies carried out in school settings with a foreign language situation, not a second language situation, are not so straightforward and, in fact, they usually conclude that older starters show a faster rate of acquisition (Muñoz, 2006). The Age Factor issue is relevant for any context of language acquisition, be it formal instruction, immersion, or any other (study abroad, for example). In this sense, studies completed in the Canadian context also demonstrate that late immersion students perform as well as early immersion students in some language assessments, despite the latter having accumulated two to three times more instruction learning the L2 (Turnbull *et al.*, 1998). Studies completed in formal learning contexts have thus recurrently shown that older learners are faster and better learners than younger ones in most aspects of acquisition, even in the case of pronunciation, the skill that –at least from a theoretical point of view– may benefit most from this early start (García Mayo and García Lecumberri, 2003; Muñoz, 2006). **Our results in a CLIL context are consistent with such evidence in that we also found that** *the older, the better* as has just been summarised. This is in contrast with empirical data collected in the rest of Spain such as studies by Lasagabaster (2008), or Navés and Victori (2010) in that they seemed to find a combined positive effect of early foreign language learning and CLIL. As Lasagabaster points out:

> In a study where CLIL students were a year younger than the control students (FI) but had had the same hours of exposure to English, CLIL students caught up with older FI learners. These results may help to shed some light on the combined effect of early foreign language learning and a CLIL approach and provide answers to some practical questions that are usually posed to language planners. (Lasagabaster, 2008:39).

Navés & Victori (2010) in Catalonia reached the same kind of results: CLIL learners in Grade 8 outperformed non-CLIL students two grades ahead. These authors concluded that such results were due to the qualitative effect of CLIL hours.

All in all, as Lasagabaster (2008) explains, folk beliefs are playing a paramount role in this respect, as it is widely held that *the younger, the better* in foreign language learning. Folk values exist despite the empirical evidence previously mentioned which demonstrates that it is not always the case in formal language learning contexts such as schools located in countries where the target language is seldom heard in the streets. One of the main reasons for the extended belief of *the younger, the better* lies in the idea that children are supposed to be better at acquiring languages implicitly (whereas older students and adults benefit more from explicit teaching), and for such implicit learning to take place, massive amounts of input are needed. As a result of this, it is similarly believed that this implicit learning can only be provided in second language naturalistic contexts or in immersion programmes (DeKeyser, 2000). This belief is shared not only by parents, but also by teachers and language planners, which is why most European governments have decided to lower the starting age of learning a foreign language (Eurydice, 2005; Navés, 2006).

Finally, in another study also conducted in Spain, Egiguren (2006), observed that:

> Basque 8 year-old starters catch up with 4 year-olds in just a year and a half. The teaching of Arts in English among the 8 year-olds was enough to eliminate significant differences among the two groups once they had reached the age of 10 (Egiguren, 2006: 35).

This author points to the possibility that **it was not the amount of hours of exposure, but rather the quality of the exposure** which

might account for the fact that the group which had started EFL earlier (age 4) and had had more hours of exposure was caught up at the age of 11 by another group starting later (age 8) and having had fewer hours of exposure, but part of the hours being CLIL hours. In this author's study it coud be questioned whether older learners obtained better results due to the quality of the exposure (CLIL methodology) or due to a combined effect of CLIL with a later foreign language learning. In that case, Egiguren's findings would be more in line with the present dissertation, since our results confirmed that *the older, the better*. The third hypothesis of the study is discussed in the following section.

5.3 Linguistic Progress and Gender

The third hypothesis (1c) enquired whether gender had an impact on the potential benefits learners may achieve in either of the two contexts compared: FI and FI+CLIL. It was hypothesized that FI+CLIL female participants would not significantly outperform their male peers in the linguistic progress achieved after a one year CLIL treatment as they reportedly would do in FI.

Table 21 below displays the results obtained through the statistical analyses presented in chapter 4 for a better understanding of the ensuing discussion. As in the previous sections, the left hand-side column lists the different skills gauged, writing, reading and listening. As explained more in detail in chapter 4 (*Results*) these have been the skills chosen to discuss the third hypothesis for the integrative nature of their measures. The qualitative measures used in the case of writing are broken down as task fulfilment, organisation, grammar, and vocabulary. The central column shows the results obtained by Group A, experiencing a CLIL context of learning, and the right-hand side column those by Group B, experiencing a FI only. Both groups are divided according to males' and females' results. The upper boxes in these two columns include the number of hours of instruction accumulated by each group at each data collection time, T1 and T2. The boxes to the right of each skill state the results obtained for each measure and each group at both data collection times and for males and females. In this case comparisons are performed between male and female participants at the same data collection time. It must be remembered that bold results are those that reach significant difference when comparing males' and females' groups within each group (p value < 0.05).

A quick overview of the key findings in the table are as follows. On the one hand, **in a FI context female participants were statistically better only in the organisation of the written task** but there was a tendency for female participants to obtain higher scores in 7 of the 12 domains analysed: written task fulfilment at T1, written organisation at T1, written grammar both at T1 and T2, written vocabulary both at T1 and T2, and listening at T1. On the other hand, **in a FI+CLIL context female participants were superior in the 12 domains under control and they obtained significant differences in 5 of the 10 areas statistically analysed** (two of them were not statistically analysed due to reasons explained in section 4.3.3: *Differential effects of gender: reading*), **all of them in the writing task:** written task fulfilment both at T1 and T2, written organisation also at both data collection times, and written vocabulary at T1.

Table 21: Skill Results per Context and Gender (p value)

	CLIL(Group A): Males/Females	FI(Group B): Males/Females
	AT1: 12, 210CLIL+1120FI =1330h	BT1: 13, 0CLIL+1260FI =1260h
	AT2: 13, 280CLIL+1260FI =1540h	BT2: 14, 0CLIL+1400FI =1400h
Writing Task Fulfilment	**AT1m: 2.5 ;AT1f: 3.3 (p=0.0089*)**	BT1m: 2.5 ; BT1f: 2.8 (p=0.303)
	AT2m: 3.0 ;AT2f: 3.6 (p=0.0509*)	BT2m: 2.9 ; BT2f: 2.8 (p=0.8376)
Organisation	**AT1m: 2.5 ;AT1f: 3.2 (p=0.0207*)**	**BT1m: 2.2 ; BT1f: 2.8 (p=0.0560*)**
	AT2m: 2.9 ;AT2f: 3.6 (p=0.0121*)	BT2m: 2.8 ; BT2f: 2.8 (p=1)
Grammar	AT1m: 2.1 ;AT1f: 2.7 (p=0.064)	BT1m: 2.3 ; BT1f: 2.4 (p=0.567)
	AT2m: 2.8 ;AT2f: 3.3 (p=0.106)	BT2m: 2.6 ; BT2f: 2.8 (p=0.68)
Vocabulary	**AT1m: 2.1 ;AT1f: 2.9 (p=0.0175*)**	BT1m: 2.4 ; BT1f: 2.6 (p=0.374)
	AT2m: 3.0 ;AT2f: 3.4 (p=0.1841)	BT2m: 2.7 ; BT2f: 2.8 (p=0.547)
Reading	AT1m: 14.3 ;AT1f: 14.4 (no need)	BT1m: 15.0 ; BT1f: 14.6 (no need)
	AT2m: 15.5 ;AT2f: 16.6 (no need)	BT2m: 15.8 ; BT2f: 14.4 (no need)

| **Listening** | AT1m: 109.0 ;AT1f: 109.8 (p=0.695) | BT1m: 108.1 ; BT1f: 111.2 (p=0.49) |
| | AT2m: 112.0 ;AT2f: 112.5 (p=0.699) | BT2m: 113.4 ; BT2f: 112.0 (p=0.11) |

Source: personal
T1: first data collection time   T2: second data collection time
A: Group A (CLIL+FI)   B: Group B (FI)
m: males   f: females

It is worth having a more detailed account of the previous results. As the table shows, **girls in Group A (FI + CLIL) performed significantly better in the written test as regards task fulfilment and organisation both before and after the CLIL treatment, and in the area of vocabulary, but only at T1**. No significant differences were found for grammar in the written tests. Concerning reading and listening, girls in Group A (FI + CLIL) performed better in all cases both before and after the CLIL treatment but no significant differences were found in any case. **Girls in Group B (FI only) showed significantly better performance as far as organization of the written test at T1**. This was the only area were the female subgroup in Group B obtained significantly better results than the male subgroup. Concerning reading and listening, results were mixed and no significant differences were found in any case.

In the light of these findings we can address the third hypothesis (1c). FI+CLIL female participants do not significantly outperform their male peers in the linguistic progress achieved after a one year CLIL treatment as they reportedly did in FI. **Our findings show that the CLIL context is even more beneficial for female participants in terms of an improvement of the foreign language than the FI context**. **Hence, our third hypothesis (1c) can not be confirmed.** As will be explained below, following the literature, it was expected that the female participants' results would have been better than the male participants' results in a FI context, whereas that this finding should have been the opposite in a FI+CLIL context because male students might feel more motivated to learn both the language and the subject matter, thus obtaining higher scores.

Some general considerations concerning such findings should be made here:

It could be argued that CLIL does help balance out gender differences in our study if we compare findings concerning male participants'

results at T2 when following a CLIL programme (that is AT2m) to findings concerning female participants' results at T2 when following a FI programme (that is BT1f). As table 22 below shows (a table with the same layout as the previous one), in this case male participants' results are always higher. A possible explanation for this result is the extra 280 CLIL hours of input male participants receive. However, our findings also show that female participants at T2 following a CLIL programme and when they are 13 (that is AT2f) obtain better results than their male counterparts.

Table 22: Skill Results per Context and Gender (p value)

	CLIL(Group A): Males/Females	FI(Group B): Males/Females
	AT1: 12 yrs. 210CLIL+1120FI=1330h AT2: 13 yrs. 280CLIL+1260FI=1540h	BT1: 13 yrs. 0CLIL+1260FI=1260h BT2: 14 yrs. 0CLIL+1400FI=1400h
Writing Task Fulfilment	AT2m: 3.0 ; AT2f: 3.6	BT1f: 2.8
Organisation	AT2m: 2.9 ; AT2f: 3.6	BT1f: 2.8
Grammar	AT2m: 2.8 ; AT2f: 3.3	BT1f: 2.4
Vocabulary	AT2m: 3.0 ; AT2f: 3.4	BT1f: 2.6
Reading	AT2m: 15.5 ; AT2f: 16.6	BT1f: 14.6
Listening	AT2m: 112.0 ; AT2f: 112.5	BT1f: 111.2

Source: personal
T1: first data collection time   T2: second data collection time
A: Group A (CLIL+FI)   B: Group B (FI)
m: males   f: females

We now turn to discuss these topics in the light of the theories and models concerning gender presented in chapter 2 section 2.1.4 (*Individual Differences: Gender*)

Research studies which have widely covered gender issues during the last three decades suggest that gender plays a significant role in foreign language performance, as there seems to be a female oriented culture

that spreads the idea that learning foreign languages is a feminine terrain (Kobayashi, 2002) and, subsequently, male students feel less confident and obtain worse scores (Onwuegbuzie, Bailey and Daley, 2001; Oxford, 1993). As Lasagabaster (2008) explains, research studies undertaken in different contexts show that women are more inclined to study foreign languages and second languages and they usually outperform their male counterparts (Pavlenko and Piller, 2008; Sunderland, 2000). However, CLIL programmes seem to help blur these gender-based differences. A study of foreign language learners by Schmidt, Boraie and Kassagby (1996) may help to explain this. They concluded that females were better foreign language learners because they were more intrinsically motivated, whereas males expressed more ex- trinsically motivated reasons. This is why CLIL programmes may help balance out gender differences, as male students might feel more motivated to learn both the language and the subject matter, enabling them to obtain higher scores in the subject concerned. Similarly, Marsh (2008) and Coyle (2007) also consider that CLIL can address gender issues in language learning, however more studies are needed to support such a statement.

Therefore, **our findings do not tally with the results obtained by Merisuo-Storm (2007) who stated that among CLIL groups no gender-based differences were observed. They do not tally either with the opinions expressed by Marsh (2008) and Coyle (2007) just summarised. On the contrary, our results show that the CLIL approach does not seem to vanish the differences observed in traditional foreign language teaching contexts when gender is considered: contrary to what we had hypothesised, female participants outperformed their male counterparts not only in a FI context but also in a CLIL context, our third research question being refuted.**

Subsequently, in the present study the CLIL component does not seem strong enough to motivate male participants' fostering them to learn both the language and the subject matter and enabling them to obtain higher scores in the subject concerned. Maybe the female oriented culture explained by Kobayashi's (2002) that spreads the idea that learning foreign languages is a feminine terrain and, subsequently, male students feel less confident and obtain worse scores plays a major role than it was thought in the subject of our particular CLIL context (Science). Should this be the case, our participants might be having the feeling of learning more a foreign language (English) than a content

one (Science).

Hence, further studies are called for in this area. If, as in the previous sections, we turn to research conducted in the rest of Spain, not many studies which tackle this gender issue are found. More specifically, in other bilingual contexts similar to Catalonia, the Basque Country, Lasagabaster (2008) hypothesized that there would be no statistically significant differences between male and female students enrolled in CLIL programmes in secondary education. Contrary to expectations, female students outperformed their male counterparts in all the English tests –but speaking- and in overall English competence. Subsequently, as the author explains, it could not be concluded that CLIL helps to balance English results on gender grounds, as female students outscored their male peers. Hence, our findings coincide with Lasagabaster's (2008). Therefore, as this author states:

> The gender issue has to be further examined and, if future results coincide with ours, it will be necessary to analyze in depth the reasons fot these gender-based differences. Foreign language skills are equally positive for both male and female citizens and the necessary measures need to be implemented in order to smooth out the differences revealed in this study, as their disappearance should be the objective of any education system (Lasagabaster, 2008: 40).

5.4 Summary

In this section a summary of the discussion chapter is offered. This chapter was organized into different sections each one dealing with one of the three subhypotheses (1a, 1b, and 1c) established in relation to the main question.

Firstly, **section 5.1** tackled the issue of language progress and skill development. It discussed the results obtained in order to address the first hypothesis (1a) derived from the first subquestion.

Our results concerning this first hypothesis for written production show that the group with FI+CLIL (Group A) significantly outperformed the group with FI only (Group B) in accuracy, and tended to also outperform it in syntactic complexity and in the whole set of qualitative measures (task fulfilment, organisation, grammar and vocabulary). This is also true in the case of reading

comprehension and lexico-grammatical competence, where Group A was significantly better than Group B. It was only in the domain of listening comprehension that Group B, a year older than Group A, tended to perform better than Group A.

These findings lead us to state that our first hypothesis (1a) concerning the greater progress in receptive skills for the FI+CLIL group is only partially confirmed. Reading but not listening improves significantly. Furthermore, our findings show a significant improvement in productive skills on behalf of the FI+CLIL group, something which we had not hypothesised, as writing and particularly accuracy, significatively progress and so do lexico-grammatical abilities. This is also in contrast with findings published in previous studies.

Although it is true that significant benefits do not accrue in all skills and measurements, it is also true that **tendencies in the differential progress can allow us to establish the benefits of the school's CLIL programme**.

Hence we would posit that **our results confirm the effectiveness of a CLIL programme. This is something which previous research conducted in educational contexts similar to ours has also confirmed.**

It has to be noted that the FI+CLIL group has the "extra" hours (and hence more hours), as it is often the case in most research studies since it is difficult to have valid control groups of formal instruction with the same number of hours as the CLIL group. However, **it is of utmost importance and worth noting that in spite of Group A (FI+CLIL) having a higher number of hours (70) than Group B (FI), they do not always outperform Group B at the moment when they are measured at T1**. Namely, although it is true that Group A (FI+CLIL) started with a higher onset level as far as written competence went, in the domain of lexical and syntactic complexity, in contrast, in the domain of accuracy, vocabulary and fluency, on the contrary, they had a lower onset level. This was just as in the domains of reading and listening comprehension and lexico-grammatical ability, in which Group A had lower values at T1 than Group B.

Concerning the specific language skills analysed, the fact that accuracy in the writing skill and lexico-grammatical abilities in general showed a

significant improvement in the case of our participants who followed a CLIL methodology might be explained by transfer of knowledge and skills from a FI context to a CLIL context, since they are "often" and "very often" respectively practised in the FI context.

As Group A in our study experiences a CLIL context together with a FI setting, their ability to transfer linguistic skills and competences learnt in the FI classes to the communication situations encountered in CLIL sessions will be at play. This might explain why, although writing skills and lexico-grammatical abilities are hardly practiced in the CLIL sessions, these participants obtain high results. It could be argued that the amounts of writing and grammar practice typical of FI are used in the CLIL context and that what students proceduralize in a FI context is automatized while the CLIL setting. That is, the accumulated experience of FI is what may play a major role in the relative benefits of a CLIL context enjoyed by the learner.

Secondly, **section 5.2** dealt with the the the issue of language progress and age: whether changes occured irrespective of age differences. It discussed the results obtained in order to address the second hypothesis (1b) derived from the second subquestion.

The results show that **the first part of our hypothesis can not be confirmed because despite the statistically significant difference in favour of the FI+CLIL younger students in the written syntactic complexity area, there are also statistically significant differences in favour of the FI older students in written accuracy and in the listening skill. In the same vein, sometimes younger FI+CLIL students show a tendency towards improvement, but the same occurs with older FI participants and we do not know whether it is due to the FI context or due to the age factor.**

Contrary to what happens with the first part, the second part of our hypothesis can be confirmed because at the same age and with a different number of hours, learners with a higher total number of hours receiving FI+CLIL benefit more than learners with fewer hours only receiving FI. Our findings show that FI+CLIL participants scored higher than FI learners in all areas but written fluency. Although showing a tendency towards improvement in some areas, differences turned out to be significant in all holistic measures concerning writing and in reading skills. This finding was quite expected since not only CLIL students had enjoyed 280 hours of extra exposure, but also these were CLIL hours.

A possible explanation for the finding concerning the first part of our second hypothesis (1b) could be the existing empirical evidence from the field of SLA which states that it is by reaching an upper intermediate level through formal instruction that CLIL approaches to education can be beneficial to learners and help them improve receptive skills, general fluency, vocabulary, and self-regulatory abilities. Therefore, it could be argued that **the younger participants in our research study had not reached this threshold level yet in order to draw on the learning opportunities they have in the FI context for greater improvement through CLIL.**

Another explanation concerns **maturational constraints**. The younger group (FI+CLIL) had the same number of hours as the older group (FI) but their age affects how they benefit from these hours. As explained in detail below, studies completed in formal learning contexts have recurrently shown that **older learners are faster and better learners than younger ones in most aspects of acquisition due to their higher capacity for abstraction and logical thinking**.

Studies completed in formal learning contexts have thus recurrently shown that older learners are faster and better learners than younger ones in most aspects of acquisition, even in the case of pronunciation, the skill that –at least from a theoretical point of view– may benefit most from this early start (García Mayo and García Lecumberri, 2003; Muñoz, 2006). **Our results in a CLIL context are consistent with this in that we also found that** *the older, the better* as has just been summarised. This is in contrast with empirical data collected in the rest of Spain such as studies by Lasagabaster (2008), or Navés and Victori (2010) in that they seemed to find a combined positive effect of early foreign language learning and CLIL.

In the third place, **section 5.3** dealt with the issue of the impact of gender differences in a FI and a FI + CLIL context in relation to each of the different skills measured. The results concerning the third research subquestion and the hypothesis derived from it were discussed in this section.

Our third hypothesis (1c) can not be confirmed because our findings show that the CLIL context is even more beneficial for female participants in terms of an improvement in their foreign language skills than the FI context. Following the literature, it was

expected that female participants' results would be better than male participants' results in a FI context, whereas this finding should be the opposite in a FI+CLIL context because male students might feel more motivated to learn both the language and the subject matter, thus obtaining higher scores.

However, our findings do not tally with the results obtained by previous research studies (Merisuo-Storm, 2007) who stated that among CLIL groups no gender-based differences were observed. They do not tally either with the opinions expressed by Marsh (2008) and Coyle (2007) explained in section 5.3, our third research question being refuted. On the contrary, our results show that the CLIL approach does not seem to vanish the differences observed in traditional foreign language teaching contexts when gender is considered: contrary to expectations, female participants still outperformed their male counterparts not only in a FI context but also in a CLIL context.

To conclude, as Lasagabaster explains,

> foreign language skills are equally positive for both male and female citizens and necessary measures need to be implemented in order to smooth out the differences revealed in this study, as would be advisable to occur in any educational system (Lasagabaster, 2008: 40).

6. SUMMARY AND CONCLUSIONS

This reseach seeks to tackle one of the main current areas of interest in SLA research these days, the contrast of different learning contexts and the effects these have on learner's linguistic development. More specifically, these learning contexts are the conventional 'by default' context of formal instruction (FI) in contrast with development in CLIL educational contexts.

In other words, the present study focuses on whether or not the acquisition of a language which is almost only heard and practised in the language classroom as the object of instruction, i.e. a FI context, presents significant differences with respect to the acquisition of the same language which is, in addition to the FI context, also heard and practised in the language classroom as the vehicle of instruction, i.e. a CLIL context. And, likewise, it focuses on whether or not the degree of influence of individual factors such as age, or gender reaches levels of significance on the level of competence attained in each different context.

In order to measure the differential linguistic gains achieved by the two groups of participants analysed, one following FI and in parallel CLIL instruction and the other one following a FI only programme, one main research question with three research subquestions derived from it were established. **The main research question** asked how context of learning affected the linguistic development of young bilingual secondary education EFL leaners when contrasting a group experiencing FI only and a group experiencing FI in combination with CLIL. **The first research subquestion** further asked which programme resulted in linguistic benefits if any and which skills benefited the most if any. In order to increase the reliability of the results, quantitative and qualitative measures have been adopted.

Results obtained to answer these first questions **confirmed the effectiveness of the CLIL programme**, something which previous research had already shown. However, significant benefits did not accrue in all skills and measurements. Therefore, **Hypothesis 1a,** which predicted that when contrasting the differential effects on learners' linguistic progress of the two programmes, the group in the FI+CLIL would improve significantly more than the other especially in receptive skills, can be only partially confirmed. **Reading but not listening improves significantly. Furthermore, our findings show significant**

improvement in productive skills on behalf of the FI+CLIL group, something which we had not hypothesised, as writing and particularly accuracy, significatively progress and so do lexico-grammatical abilities. This is in contrast with findings published in previous studies. Therefore, with the present study we have discretely contributed to show how under CLIL conditions certain aspects of language competence which did not seem to prove clear gains in previous studies can also be developed. These would be the case for productive skills (writing), and formal aspects such as accuracy (also in writing) or lexico-grammatical abilities.

As regards **the second subquestion**, which enquired wether age had an impact on the potential benefits learners may achieve in either of the two contexts compared, FI and FI+CLIL, results tend to confirm that **age does have an impact in the sense that the older, the better in contexts of CLIL learning.** This is in contrast with empirical data collected in the rest of Spain such as studies by Lasagabaster (2008), or Navés and Victori (2010) in that they seemed to find a combined positive effect of early foreign language learning and CLIL. The first part of **Hypothesis 1b,** which predicted that at different ages and with a similar number of hours, younger learners receiving FI+CLIL would benefit more than elder learners only receiving FI because 'the earlier the better' for CLIL instruction, can thus not be confirmed. Despite the statistically significant difference in favour of the FI+CLIL younger students in the written syntactic complexity area, there are also statistically significant differences in favour of the FI older students in written accuracy and in the listening skill. In the same vein, sometimes younger FI+CLIL students show a tendency towards improvement, but the same occurs with older FI participants. We have argued that this could be due to **maturational constraints**. **Older learners are reportedly faster and better learners than younger ones in most aspects of acquisition due to their higher capacity for abstraction and logical thinking** in EFL contexts of learning. Therefore, as has just been summarized, the present research contributes with empirical data to the belief that in formal settings older learners are faster and better in most aspects of acquisition.

On the other hand, the second part of **Hypothesis 1b** predicted that at the same age and with a different number of hours, learners with a higher total number of hours receiving FI+CLIL would benefit more than learners with fewer hours only receiving FI. In this case our findings show that FI+CLIL participants scored higher than FI learners

in all areas but written fluency. Although showing a tendency to improvement in some areas, differences turned out to be significant in holistic measures concerning writing and in reading skills. This finding was quite expected since not only CLIL students had enjoyed 280 hours of extra exposure, but also these were CLIL hours.

Regarding the **third research subquestion**, which asked whether gender had an impact in the benefits obtained with either a FI context of learning, or a FI+CLIL context, results confirm that female participants are better than male participants in both contexts of acquisition.

As a consequence, **Hypothesis 1c,** which predicted that female participants in the FI+CLIL group would not significantly outperform their male peers in the linguistic progress achieved after a one year CLIL treatment as they reportedly did in FI, can not be confirmed**. On the contrary, our results show that the CLIL approach does not seem to vanish the differences observed in traditional foreign language teaching contexts when gender is considered: contrary to expectations, female participants still outperformed their male counterparts not only in a FI context but also in a CLIL context.** Concerning this last subquestion, this research has contributed to add empirical data to those studies that examine the gender issue in order to analyze in depth the reasons for these gender-based differences.

As a final point, going back to the main research question, we can conclude that the effectiveness of a **CLIL** context of learning in this dissertation is confirmed but that it **does not suffice to improve the participants' overall linguistic competence** as, whereas some levels of language competence make substantial progress in the CLIL context examined, some other levels do not seem to follow the same path. As a result of this imbalance and of the contradictory results also found within the existing empirical evidence in this field of research (an issue further tackled below in the *Limitations and Further Research* section within this chapter), it is interesting to see what is there beyond CLIL. This will allow us to explore new possibilities geared to achieve linguistic progress in all abilities in parallel. This takes us to further consider other contexts of learning which can complement the CLIL context. Following Pérez-Vidal (2011), the *Combination of contexts* hypothesis states that students should be offered the possibility of experiencing three contexts of learning, that is FI, Stay Abroad (SA)

and CLIL in combination, at subsequent educational levels. As has been already mentioned in the Discussion chapter (chapter 5),

> it is by reaching an upper intermediate level through formal instruction, that CLIL approaches to education can be beneficial to learners to improve in receptive skills, general fluency, vocabulary, and self-regulatory abilities. Subsequent to FI and CLIL, a SA period in the target language country can prove most fruitful, specially if of an adequate length (…) while doing some academic work (Pérez-Vidal, 2011:11).

This hypothesis opens a new multifaceted dimension which combines attention to form, semiimersion and immersion contexts. This formula seems to combine all the ingredients necessary to achieve competence in all abilities and hence success in foreign language acquisition. Whether it is confirmed or not is, with no doubt, of undeniable interest for research, and has clear practical evidence.

The next section presents the limitations of the present investigation as well as the issues derived from this research study that deserve, to our mind, further research.

Limitations and further research

First of all, a limitation of the present investigation is that participants did not fulfill a sociolinguistic questionnaire. As explained in section 3.3.1 (Context and Participants), they were selected because they represented the first group undertaking a CLIL programme and the last group only undertaking FI before CLIL started in a state-run (ordinary governtment-supported) school in Barcelona (Spain). Having been placed together in this school since nursery, participants had all started learning English at the age of 6 (Grade 1), following the official curriculum in Catalonia, so both groups shared a common age of onset of exposure to English as their L3 with the same teachers. The majority of these learners followed extra-curricular classes. Despite sharing all these variables, we believe that it would have been desirable for all the participants to have fulfilled a sociolinguistic questionnaire in order to increase the reliability of the study.

Secondly, the most commonly used measures in the field of SLA were used in the present study to account for the participants' performance in the different L3 areas. However, in our view, the lack of progress of

certain students in some areas might be due to the tests designed for this dissertation. This would be the case of the test administered to measure the listening skill: a dictation. As explained in section 3.3.5 (Analyses/Measures) as well as in the discussion chapter (5.1 Linguistic Progress and Skill Development), it has to be born in mind that the evaluation of the dictations caused some scoring problems and thus the results obtained are not very reliable. Therefore, in a future investigation a more fine-grained instrument to measure this skill should be used.

In the third place, as has been already mentioned in chapter 5 (Discussion), future research should investigate the extent to which the results obtained are long-lasting. In our view this should be done specially concerning those areas where unexpected results were found such as the listening skill or accuracy in the written skill. While it should be noted that some of the findings in the research here discussed coincide with claims made for CLIL in similar contexts and at other latitudes in the continent, it is also true that contradictory results in this field of research are also very frequent. As Pérez-Vidal (forthcoming) explains, some possible explanations of contradictory CLIL research findings are:

- It is difficult to have valid control groups of FI with the same number of hours as the CLIL group. Hence, the CLIL group often has the "extra" hours.
- The groups compared are not really "comparable". CLIL groups often include "selected" students.
- Adequate instruments to measure conflated content and language are lacking.
- Most CLIL classrooms are experimental. Hence results should not be extrapolated.

Except for the second one, these are all situations that occur in the present investigation. Thus, it is evident that more research is needed in this field to obtain more accurate and precise results and to be able to infer any generalised outcomes related to European CLIL programmes. We defend a vision of CLIL as a long-term programme and not as an experimental project. In other words, a carefully planned programme similar to those successfully implemented in immersion settings in Canada or the USA.

Fourthly, given that this dissertation has shown that age plays a role in the acquisition of a foreign language in a class setting, it would be also interesting for future research to investigate the role that the learning context plays on the FL development on another population of adults. In so doing, there would be more evidence on the effects of age in a CLIL setting and consequently different kinds of programmes for participants at different ages could be developed so that they could all make the most of their CLIL and FI courses.

In the fifth place, in the present study data have been analysed from certain angles only. Thus the skills examined were reading, listening, writing and lexico-grammatical abilities. Other interesting areas such as speaking, pronunciation or pragmatics have not been tackled due to lack of time and space. Bearing in mind that the speaking skill is more practised in a CLIL context than in a FI context, analysing these areas would no doubt reveal very interesting findings.

In addition, further research should also focus on the issue of onset level. On the basis of the existing empirical evidence from the field of SLA, it is by reaching an upper intermediate level through formal instruction, that CLIL approaches to education can be beneficial to learners in order to improve in receptive skills, general fluency, vocabulary, and self-regulatory abilities (Pérez-Vidal, 2011: 11). This, as explained in chapter 5 (Discussion), would be a possible explanation for the finding concerning the first part of our second hypothesis (contrary to expectations, younger participants following a CLIL programme were not superior to older participants following FI only) if we consider that they had not reached this threshold level yet. As a consequence, given the importance of the participants' proficiency level, future research should aim at throwing some light on, first, whether a threshold level can indeed be identified and, second, what is its role in CLIL experiences so that their benefits could be maximized.

As a final point, further research should also take into account pedagogical implications.

I would here like to finish on a personal note. As explained in the introduction of this study, being myself a teacher, apart from other reasons, this research was motivated by a desire to discover what classroom conditions are most likely to facilitate acquisition. It is important here to remember that **our results confirm the effectiveness of a CLIL programme on the development of a FL and that they**

are consistent with the belief that *the older, the better*. Thus, taking these findings into account, if they happen to coincide with the results in other similar previous research studies, primary and secondary schools should consider them and consequently offer CLIL programmes. In this case, all students should have the possibility to participate in CLIL settings along with their FI classes. By doing so, when mature enough they would reach an upper intermediate level through FI which would enable them to draw on the learning opportunities they have in the FI context for greater improvement through CLIL.

These and many other questions remain to be answered and further studied in a field that promises to grow as the number of multilinguals worldwide keeps growing. CLIL is about innovation, and undoubtedly challenges have to be faced with any innovative approach. However, in the light of the results so far obtained, the firm base on which this new approach stands seems to be worth the effort for the hope of multilingualism.

As Lorenzo, *et al.* (2010) explained, it is possible that, in the long term, CLIL-type initiatives might contribute to the formulation of a common European ideology of language. Such a paradigm would, of necessity, be rooted in the historical tradition of educational multilingualism in the continent. Where it was once believed that the quintessential cultural endeavour of Europe across time lay in the search for the perfect language (Eco 1995), this quest is now considered utopian and dated; nowadays the goal has become the propagation of plurilingual competences and multicultural values and CLIL may well have a significant contribution to make in this endeavour.

7. REFERENCES

Ackerl, C. (2007). Lexico-grammar in the essays of CLIL and non-CLIL students: error analysis of written production. *ViewZ (Vienna English Working Papers)*, 16/3, 6-11.

Alcon, E. (2009). Focus on Form, learner uptake and subsequent lexical gains in learners' oral production. *International Review of Applied Linguistics in Language Teaching*, 47/3-4, 347-365.

Alcon, E. and Safont, P. (2001) Occurrence of exhortative speech acts in ELT materials and natural speech data: A focus on request, suggestions and advice realization strategies. *Studies in English Language and Linguistics* 3, 5-22.

Alejo, R. and MacArthur, F. (2009). Linguistic, cognitive and affective outcomes of CLIL experiences: The case of bilingual education programmes in Extremadura (Spain). Paper presented at the Gourt Conference 2009: Implicit and Explicit Conditions, Processes, and Knowledge in SLA & Bilingualism. Georgeton University, Washington, DC.

Allen, P., Swain, M., Harley, B., and Cummins, J. (1990). Aspects of classroom treatment: Towards a more comprehensive view of second language education. In Harley, B., Allen, P., Cummins, J., Swain, M. *The Development of Second Language Proficiency* (57-81). Cambridge: Cambridge University Press.

Allen, L. (2000). Form-meaning connections and the French causative: An experiment in processing instruction. *Studies in Second Language Acquisition*, 22(1), 69-84.

Allwright, R.L. (1984). The Importance of Interaction in Classroom Language Learning. *Applied Linguistics,* 5/2, 156-171.

Anderson, J.R. (1983). *The Architecture of Cognition.* Cambridge, Mass: Harvard University Press,

Anderson, J.R. and Fincham, J.M. (1994). Acquisition of procedural skills from examples. *Journal of Experimental Psychology: Learning, Memory and Cognition*, 20 (6), 1322-40.

References

Anderson, J.R., Fincham, J.M. and Douglass, S. (1997). The role of examples and rules in the acquisition of a cognitive skill. *Journal of Experimental Psychology: Learning, Memory and Cognition*, 23 (4), 932-45.

Artigal, J.M. (1993). Catalan and Basque immersion programmes. In Baetens Beadsmore, H. *European Models of Bilingual Education*, 30-53. Clevedon: Multilingual Matters.

August, D. and Hakuta, K. (1997). *Improving Schooling for Language-Minority Children: A Research Agenda.* Washington, DC: National Academy Press.

Autonomous University at Barcelona. (2008) *Pla de llengües* (unpublished manuscript).

Baetens Beadsmore, H. (1986). *Bilingualism: Basic Principles.* Clevedon: Multilingual Matters.

Baetens Beadsmore, H. (1993). *European Models of Bilingual Education.* Clevedon: Multilingual Matters.

Baetens Beardsmore, H. (1995). The European school experience in multilingual education. In Skuttnab-Kangas, T. *Multilingualism for All*, 21-68. Lisse: Svets & Zeitlinger.

Baetens Beardsmore, H. (1998). Remarks made at the October CEILINK conference. Strasbourg' in Marsh, D. and Marsland, B. *CLIL Initiatives for the Millenium. Report on the CEILINK Think-Tank.*, 24-30. Jyväskylä: University of Jyväskylä.

Baker, O. (1993). *Foundations of Bilingual Education and Bilingualism.* Clevedon: Multilingual Matters.

Ben-Zeev, S. (1977). Mechanisms by which childhood bilingualism affects understanding of language and cognitive structures. In Hornby, P.A. *Bilingualism,* 29-55. New york: Academic Press.

Berg, G. (1993). *"Mir we¨lle bleiwe, wat mir sin": Soziolinguistische und sprachtypologische Betrachtungen zur luxemburgischen Mehrsprachigkeit*. Tubingen: Niemeyer

References

Bernhardt, E. (1992). *Life in Language Immersion Classrooms*. Clevedon: Multilingual Matters.

Bhatia, T. and Ritchie, W. (2004). *The Handbook of Bilingualism*. United Kingdom: Blackwell Publishers.

Bialystok, E. (1985). The compatibility of teaching and learning strategies. *Applied Linguistics* 6: 155-262.

Bialystok, E. (1991). *Language Processing in Bilingual Children.* Cambridge: Cambridge University Press.

Bialystok, E. (1991). Metalinguistic dimensions of bilingual language proficiency. In Bialystok, E. *Language Processing in Bilingual Children,* 113-140. Cambridge: Cambridge University Press.

Bialystok, E. (2004). Language and literacy development. In Bhatia, T. and Ritchie, W. *The Handbook of Bilingualism.* United Kingdom: Blackwell Publishers.

Bialystok, E. (2005). Consequences of bilingualism for cognitive development. In Kroll, J.F., De Groot, A.M.B. Lexical access in bilingual production. *Handbook of Bilingualism: Psycholinguistic Approaches.* New York: Oxford University Press, 417-432.

Bialystok, E., Craik, F.I.M., Klein, R., Viswanathan, M. (2004). Bilingualism, aging, and cognitive control: Evidence from the Simon Task. *Psychology and Aging* 19, 290-303.

Birdsong, D. (1999). *Second Language Acquisition and the Critical Period Hypothesis.* Mahwah: Lawrence Erlbaum Associates.

Bloomfield, L. (1933*). Language*. London: Allen and Unwin.

Bonnet, A. (2004). *Chemie im Bilingualen Unterricht. Kompetenzerwerb durch Interaktion*. Opladen: Leske + budrich.

Brinton, D., Snow, M. and Wesche, M. (1989). *Content-based Second Language Instruction.* NY: Newbury Press.

Butler, Y. and Hakuta, K. (2004). Bilingualism and second language acquisition. In Bhatia, T. and Ritchie, W. *The Handbook of*

Bilingualism. United Kingdom: Blackwell Publishers.

Canale, M. and Swain, M. (1980). Theoretical bases of communicative approaches to second language teaching and testing. *Applied Linguistics* 1, 1-47.

Carroll, J. (1991). Cognitive abilities in foreign language aptitude: then and now. In Parry, T. and Stansfield, C. *Language Aptitude Reconsidered*. Englewood Cliffs, N.J.: Prentice Hall.

Casanave, C.P. (1994). Language development in students' journal. *Journal of Second Language Writing* 3, 179-201.

Casas, M. (2007). *Competències bàsiques per parlar i escriure ciència a l'educació primària.* Barcelona: Associació de Mestres Rosa Sensat.

Celaya, M.L., Pérez-Vidal, C., Torras, M.R. (2001). Matriz de criterios de medición para la determinación del perfil de competencia lingüística escrita en inglés lengua extranjera. *RESLA* 14, 87-89.

Cenoz, J. (2003). The additive effect of bilingualism on third language acquisition: A review. *The International Journal of Bilingualism* 7.1, 71-88.

Cenoz, J. and Genesee, F. (1998). *Beyond Bilingualism: Multilingualism and Multilingual Education*. Clevedon: Multilingual Matters.

Cenoz, J., Hufeisen, B., Jessner, U. (2001a). *Third Language Acquisition in the School Context.* Special Issue of *International Journal of Bilingualism and Bilingual Education* 4.

Cenoz, J., Hufeisen, B., Jessner, U. (2001b). *Looking beyond Second Language Acquisition: Studies in Tri- and Multilingualism.* Tübingen: Stauffenburg.

Cenoz, J. and Jessner, U. (2000). *English in Europe: The Acquisition of a Third Language.* Clevedon: Multilingual Matters.

Cenoz, J. and Lindsay, D. (1994). Teaching English in primary school: A project to introduce a third language to eight-year-olds.

Language and Education 8, 201-210.

Cenoz, J. and Valencia, J.F. (1994). Additive trilingualism: Evidence from the Basque Country. *Applied Psycholinguistics* 15, 195-207.

Centre de Recursos de Llengües Estrangeres (2006). *CLIL in Catalonia, from Theory to Practice*. APAC, 2005. Department of Education: Generalitat de Catalunya.

Chamot, A. and O'Malley, J. (1994). Language learner and learning strategies. In Ellis, N. *Implicit and Explicit Learning of Languages*, 371-392. London: Academic Press.

Chomsky, N. (1959). Review of *Verbal Behaviour* by Skinner, B.F., *Language* 35, 1-47.

Christian, D. and Genesee, F. (2001). *Bilingual Education.* TESOL Quarterly.

Clément, R. and Gardner, R.C. (2001). Second language mastery. In Robinson, W.P. and Gile, H. *The New Handbook of Language and Social Psychology*. New York: John Wiley, 489-504.

Clyne, M. (1997). Some of the things trilinguals do. *International Journal of Bilingualism* 1 (2), 95-116.

Clyne, M. (2003). *Dynamics of language contact*. Cambridge: Cambridge University Press.

Codó, E., Masats, D., Feixes, M., Espinet, M. and Couso, D. (2007). Analysing the level of complexity of students' written responses: A comparison between first and foreign language productions. In Rust, C. *Improving Student Learning through Teaching Book 14*, 158-170. Oxford: Oxford Centre for Staff and Learning Development.

Cohen, A. (1989). Attrition in the productive lexicon of two Portuguese third language speakers. *Studies in Second Language Acquisition* 11, 135-149.

Cohen, A. and Weaver, S. (2006). *Styles and Strategies-based Instruction: A Teacher's Guide.* Minneapolis, MN: Center for

References

Advanced Research on Language Acquisition, University of Minnesota.

Collins L., Halter R.H., Lightbown P.M. and Spada, N. (1999). Time and the distribution of time in L2 instruction. *TESOL Quarterly* 33: 655-680.

Consell Superior d'Avaluació del Sistema Educatiu (2005). *La situació de la llengua anglesa al batxillerat a Catalunya.* Barcelona: Generalitat de Catalunya, Department d'Educació.

Consell Superior d'Avaluació del Sistema Educatiu (2008a). *PISA 2006 Resultats de l'alumnat a Catalunya Avaluació de l'educació secundària obligatòria 2006.* Barcelona: Generalitat de Catalunya, Department d'Educació.

Consell Superior d'Avaluació del Sistema Educatiu (2008b). *Avaluació de l'Educació Secundària Obligàtoria 2006. Estudi sociodemogràfic i lingüístic als centres d'educació secundària de Catalunya.* Barcelona: Generalitat de Catalunya, Department d'Educació.

Conselleria d'Educació i Cultura (2008). Resolució de la Consellera d'Educació i Cultura de 19 de juny de 2008. *Butlletí Oficial de les Illes Balears* 121.

Conselleria d'Educació i Cultura (2009). Ordre de 17 de juny de 2009. *Butlletí Oficial de les Illes Balears* 98.

Cook, V. (1997). The consequences of bilingualism for cognitive processing. In de Groot, J. and Kroll, J.F. *Tutorials in Bilingualism: Psycholinguistic Perspectives*, 279-300. Mahwah: Lawrence Erlbaum Associates.

Costa, A., Hernández, M., Sebastián-Gallés, N. (2008). Bilingualism aids conflict resolution: Evidence from the ANT task. *Cognition* 106, 59-86.

Council of Europe (2001). *Common European Framework of Reference for Languages: Learning, Teaching, Assessment.* Cambridge: Cambridge University Press.

References

Coyle, D. (1998). Planning for effective classrooms: supporting students in content and language integrated learning contexts. (manuscript).

Coyle, D. (2000). Meeting the challenge: developing the 3Cs curriculum. In Green, S. *New Perspectives on Teaching and Learning Modern Languages.* Clevedon: Multilingual Matters.

Coyle, D. (2007). Content and language integrated learning: towards a connected research agenda for CLIL pedagogies. *International Journal Bilingual Education Bilingual* 10: 543-562.

Coyle, D., Hood, P., and Marsh, D. (2010). *Content and Language Integrated Learning.* Cambridge: Cambridge University Press.

Cummins, J. (1976). The influence of bilingualism on cognitive growth: A synthesis of research findings and explanatory hypotheses. *Working Papers on Bilingualism* 9, 1-43.

Cummins, J. (1981). The role of primary language development in promoting educational success for language minority students. In California Statement of Education (ed.). *Schooling and Language Minority Students: A Theoretical Framework* (3-49). Los Angeles: Evaluation, Dissemination and Assessment Center, California State University.

Cummins, J. (2000). *Language, Power and Pedagogy: Bilingual Children in the Crossfire*. Clevedon, UK: Multilingual Matters.

Cummins, J. (2001). Instructional conditions for trilingual development. *International Journal of Bilingual Education and Bilingualism* 4 (1), 61-75.

Cummins, J. and Swain, M. (1986). *Bilingualism in Education.* London: Longman.

Dafouz, E. and Guerini, M. (2009). *CLIL across Educational Levels*, 3-17. Richmond publishing. Madrid: Santillana.

Dafouz, E., Núñez, B. and Sancho, C. (2008). Analysing sentence in a CLIL university context: Non-native speaker use of personal pronouns and modal verbs. *The International Journal of*

References

Bilingualism and Bilingual Education 10/5, 647-622.

Dalton-Puffer, C. (2005). Negotiating interpersonal meanings in naturalistic classroom discourse: Directives in content and language integrated classrooms. *Journal of Pragmatics 37: 1275-1293.*

Dalton-Puffer, C. (2007). *Discourse in Content and Language Integrated Learning.* Amsterdam: John Benjamins.

Dalton-Puffer, C. (2008). Outcomes and processes in content and language integrated learning (CLIL): Current research from Europe. In Delanoy, W. and Volkmann, L. *Future Perspectives for English Language Teaching.* (139-57).

Dalton-Puffer, C. and Nikula, T. (2006). Pragmatics of content-based instruction: teacher and student directives in Finnish and Austrian classrooms. *Applied Linguistics* 27: 241-267.

Dalton-Puffer, C. and Smit, U. (2007). *Empirical Perspectives on CLIL Classroom Discourse.* Frankfurt, Vienna: Peter Lang.

Davies, M. (2006). Paralinguistic focus on form. In *TESOL Quarterly*, 40/4, 841-855.

Davis, K. A. (1994). *Language Planning in Multilingual Settings: Policies, Communities and Schools in Luxembourg.* Amsterdam: Benjamins.

Day, E. and Shapson, S. (1996). *Studies in Immersion Education.* Clevedon: Multilingual Matters.

De Angelis, G. and Selinker, L. (2001). Interlanguage transfer and competing linguistic systems in the multilingual mind. In Cenoz, J., Hufeisen, B., Jessner, U. *Cross-linguistic Influence in Third Language Acquisition: Psycholinguistic Perspectives* (42-58). Clevedon: Multilingual Matters.

De Bot, K. (1992, 1996). A bilingual production model: Levelt's 'Speaking' model adapted. *Applied Linguistics* 13, 1-24.

References

De Bot, K. (2004). The multilingual lexicon: Modelling selection and control. *International Journal of Multilingualism* 1.1, 17-32.

DeKeyser, R. (1993). The effect of error correction on L2 grammar knowledge and oral proficiency. *The Modern Language Journal* 77, 501-14.

DeKeyser, R. (1997). Beyond explicit rule learning: Automatizing second language morphosyntax. *Studies in Second Language Acquisition*, 19 (2), 195-221.

DeKeyser, R. (1998). Beyond focus on form: cognitive perspectives on learning and practicing second language grammar. In Doughty, C. and Williams, J. *Focus on Form in Classroom Second Language Acquisition.* Cambridge: Cambridge University Press.

DeKeyser, R. (2000/02). The robustness of critical period effects in second language acquisition. *Studies in Second Language Acquisition*, 22, 499-533.

Dekeyser, R. (2001). Automaticity and automatization. In Robinson, P. *Cognition and Second Language Instruction, 125-51.* New York*:* Cambridge University Press.

Dekeyser, R. (2003). Implicit and explicit learning. In Doughty, C. and Long, M. *The Handbook of Second Language Acquisition.* United Kingdom: Blackwell Publishers, 313-348.

Dekeyser, R. (2007a). *Practice in a Second Language. Perspectives from Applied Linguistics and Cognitive Psychology*. New York: Cambridge University Press.

Dekeyser, R. (2007b). Skill Acquisition theory. In Williams, J. and VanPatten, B. *Theories in Second Language Acquisition: an Introduction,* 97-113. Mahwah: Lawrence Erlbaum.

Dekeyser, R. and Sokalski, K. (1996). The Differential Role of Comprehension and Production Practice. *Language Learning* 46: 613-42.

Dekeyser, R. and Sokalski, K. (2001). The Differential Role of Comprehension and Production Practice. In Ellis, R. *Form-*

Focused Instruction and Second Language Learning, 81-112. United Kingdom: Blackwell Publishers.

Dentler, S., Hufeisen, B. and Lindemann, B. (2000). *Tertiär- und Drittspachen: Projekte und Empirische Untersuchungen.* Tübingen: Stauffenburg.

Dewaele, J.M. and Furnham, A. (2000). Personality and speech production: A pilot study of second language learners. *Personality and Individual Differences,* 28, 355-365.

Dewaele, J.M., Housen, A., and Wei, L. (2003). *Bilingualism: Beyond Basic Principles.* Clevedon: Multilingual Matters.

Dooly, M. (2008). *How we're going about it. Teachers' Voices on Innovative Approaches to Teaching and Learning Languages.* Cambridge: Cambridge Scholar Press.

Dörnyei, Z. (2000). *Motivational Strategies in the Language Classroom.* Cambridge: Cambridge University Press.

Dörnyei, Z. (2001). *Teaching and Research Motivation.* Harlow: Longman.

Dörnyei, Z. (2005). *The Psychology of the Language Learner: Individual Differences in Second Language Acquisition.* Mahwah, N.J.: Lawrence Erlbaum.

Doughty, C. (2001). Cognitive underpinnings of focus on form. In Robinson, P. *Cognition and Second Language Instruction* (206-57). Cambridge: Cambridge University Press.

Doughty, C. (2003). Instructed SLA: constraints, compensation and enhancement. In Doughty, C. and Long, M. *The Handbook of Second Language Acquisition*. United Kingdom: Blackwell Publishers.

Doughty, C. and Williams, J. (1998). *Focus on Form in Classroom Second Language Acquisition*. Cambridge University Press.

Doughty, C. and Long, M. (2003). *The Handbook of Second Language Acquisition*. United Kingdom: Blackwell Publishers.

Dulay, H., Burt, M. and Krashen, S.D. (1982*). Language Two*. New York: Oxford University Press.

Eco, U. (1995). *The Search for the Perfect Language (The Making of Europe).* Blackwell.

Eder, C. (1998). English across the curriculum. Diplomarbeit. University of Klagenfurt.

Edwards, A.P. and Westgate, D.P.G. (1994). *Investigating Classroom Talk*. London and Washington D.C: Falmer Press.

Egiguren, I. (2006). *Atzerriko hizkuntza goiztiarraren eragina gaitasun eleaniztunean.* Unpublished PhD dissertation. Vitoria-Gasteiz: University of the Basque Country.

Ehlich, K. and Rehbein, J. (1986). *Muster und Institution. Untersuchungen zur Schulischen Kommunikation*. Tübingen: Narr.

Ehrman, R. (1996). *Understanding Second Language Learning Difficulties*. Thousand Oaks, CA: Sage.

Ellis, N.C. (1993). Rules and instances in foreign language learning: Interactions of explicit and implicit knowledge. *European Journal of Cognitive Psychology*, 5, 289-318.

Ellis, R. (1988). The role of practice in classroom language learning. *AILA Review* 5: 20-39.

Ellis, R. (1992). *Second Language Acquisition and Language Pedagogy*. Clevedon, UK: Multilingual Matters.

Ellis, R. (1993). The structural syllabus and second language acquisition. *TESOL Quarterly*, 27, 91-113.

Ellis, R. (1994a). A theory of instructed second language acquisition. In Ellis, N.C. *Implicit and Explicit Learning of Languages*, 79-114. London: Academic Press.

References

Ellis, R. (1994b). *The Study of Second Language Acquisition*. Oxford: Oxford University Press.

Ellis, R. (1997). *Second Language Acquisition Research and Language Teaching*. Oxford: Oxford University Press.

Ellis, R. (1999a). *Learning in a Second Language through Interaction*. Amsterdam: John Benjamins.

Ellis, R. (1999b). Input-based approaches to teaching grammar: a review of classroom oriented research. *Annual Review of Applied Linguistics*, 19, 64-80.

Ellis, R. (2000). *Second Language Acquisition*. Oxford: Oxford University Press.

Ellis, R. (2001a). *Form-Focused Instruction and Second Language Learning*. United Kingdom: Blackwell Publishers.

Ellis, R. (2003). *Task-based Language Learning and Teaching*. Oxford: Oxford University Press.

Ellis, R. (2004). Individual differences in language learning. In Elder, C. and Davies, A. *Handbook of Applied Linguistics*. Oxford: Blackwell.

Ellis, R. (2008). *The Study of Second Language Acquisition* (second edition*)*. Oxford: Oxford University Press.

Ellis, R., Barkhuizen, G. (2005). *Analysing Learner Language.* Oxford: Oxford University Press.

Ellis, R., Basturkmen, H. and Loewen, S. (2002). Doing focus on form. System 30: 419-32.

Ellis, R., Tamaka, Y. and Yamasaki, A. (1994). Classroom Interaction, comprehension and the acquisition of L2 word meanings. *Language Learning* 44/3, 449-491.

Erlam, R. (2005). Language aptitude and its relationship to instructional effectiveness in second language acquisition. *Language Teaching Research 9*: 147-72.

References

Escobar, C. (2002). Attention and the processing of comprehensible input in communication tasks among secondary learners. In Nizegorodcew, A. *Beyond L2 Teaching*. Krakow: Jagiellonian University Press: 45-52.

Escobar, C. (2004). Content and Language Integrated Learning: Do they learn content? Do they learn language?. *Proceedings of the XXI International Conference AESLA*. 27-38.

Escobar, C. (2006). Algunes formes d'intervenció en la formació plurilingüe. In Nussbaum, L. and Unamuno, V. *Usos i competències multilingües d'escolars d'origen immigrant*. Bellaterra, Servei de Publicacions de la UAB: 135.

Escobar, C. (2007). *New Challenges in Foreign Language Education: Content and Language Integrated Learning and intercultural Language Teaching.* Symposium. University of Lleida.

Escobar, C. (2009). Enseñar en ingles. Cuando la lengua de la escuela es diferente de la lengua familiar. In *Cuadernos de pedagogía* 395, 46-76. Barcelona: Wolters Kluwer.

Escobar, C. and Nussbaum, L. (in press). *Aprender en otras lenguas. CLIL in the Classroom. Aprendre dans Autres Langues. L'AICLE a l'aula*. Bellaterra (Barcelona): ICE UAB.

Escobar, C. and Nussbaum, L. (2008). Tareas de intercambio de información y procesos de aprendizaje en el aula AICLE. In Camps, A. and Milian, M.. *Miradas y voces. Investigaciones sobre la educación lingüística y literaria en entornos plurilingües.* Barcelona: Graó, 159-178.

Escobar, C. and Pérez -Vidal, C. (2004). Teacher education for the implementation of CLIL. In Wilkinson, R. (ed.) *Meeting the Challenge of a Multilingual Higher Education.* Netherlands: Maastricht University.

Escobar, C. and Sánchez, A. (2009). Language learning through tasks in a CLIL Science classroom. *Porta Linguarum,* 11, 65-83.

European Commission (1995) 449 *White Paper on Education and Learning*. Brussels: European Comission.

References

European Commission (2003) 590 *Promoting Language Learning and Linguistic Diversity: An Action Plan 2004-2006*. Brussels: European Comission.

European Commission (2006). *Europeans and their Languages. Special Eurobarometer 243*. Brussels: European Comission.

European Commission (2008). *A Rewarding Challenge. How the Multiplicity of Languages could strengthen Europe.* Group of Intellectuals for Intercultural Dialogue chaired by Mr. Amin Maalouf. Brussels: European Comission.

European Council (2007). *Languages of Schooling. Towards a Framework for Europe*. Strasbourg: European Council.

Eurydice (2005). *Key Data on Teaching Languages at School in Europe.* Brussels: European Commission.

Eurydice (2006). *Content and Language Integrated Learning (CLIL) at School in Europe. Spain National Description* 2004/05, Brussels.

Eurydice (2008). *Key Data on Teaching Languages at School in Europe. Special Eurobarometer 243*. Brussels: European Commission.

Evnitskaya, N. and Aceros, J.C. (2008). "We are a good team": El contrato didáctico en parejas de aprendices de lengua extranjera. *Resla* ,21, 45-70.

Faingold, E. (1999). The re-emergence of Spanish and Hebrew in a multilingual adolescent. *International Journal of Bilingual Education and Bilingualism* 2, 283-295.

Feixas, M., Codó, E., Couso, D., Espinet, M. and Masats, D. (2009). Enseñar en inglés en la universidad: reflexiones del alumnado y el profesorado en torno a dos experiencias AICLE. In Roig, R., Blasco, J., Cano, M., Gilar, R., Lledó, A. and Mañas, C. *Investigar desde un Contexto Educativo Innovador.* Alicante: Editorial Marfil.

References

Foster, P. and Skehan, P. (1996). The influence of planning and task type on second language performance. *Studies in Second Language Acquisition* 18, 299-324.

Frigols, M.J. (2007). CLIL implementation in Spain: An approach to different models. Conselleria de Educación de la Generalitat Valenciana. Valencia.

Gajo, L. (2007). Knowledge and subject knowledge: How does bilingualism contribute to subject development?. *International Journal on Bilingual Education and Bilingualism* 10, 5: 563-581.

Gallaway, G. and Fichards, B.J. (1994). *Input and Interaction in Language Acquisition*. Cambridge: Cambridge University Press.

García-Mayo, M.P. and García-Lecumberri, M.L. (2003). *Age and the Acquisition of English as a Foreign Language.* Clevedon: Multilingual Matters.

Gardner, R.C. (2007). Motivation and second language acquisition. Manuscript. *Porta Linguarum* 8, 9-20.

Gass, S. (1997). *Input, Interaction, and the Second Language Learner.* Mahwah, NJ: Lawrence Erlbaum.

Gass, S. (2003). Input and interaction. In Doughty, C. and Long, M. *The Handbook of Second Language Acquisition,* 224-55. United Kingdom: Blackwell Publishers.

Gass, S., Mackey, A. and Pica, T. (1998). The role of input and interaction in second language acquisition. *The Modern Language Journal*, 82/3, 299-307.

Genesee, F. (1987). *Learning through two Languages. Studies of Immersion and Bilingual Education*. Cambridge MA: Newbury House.

Genesee, F. (1994). *Integrating Language and Content: Lessons from Immersion.* [Educational Practice Report 11]. Santa Cruz, CA: National Center for Research on Cultural Diversity and Second Language Learning.

References

Genesee, F. (2008). Insights from immersion research. Plenary talk delivered at the "CLIL Fusion. Multilingual Mindsets in a Multicultural World" Conference. Tallin, Estonia, 24 October 2008.

Green, D. (1986). Control, activation and resource: A framework and a model for the control of speech in bilinguals. *Brain and Language* 27, 210-223.

Green, D. (1998). Mental control of the bilingual lexico-semantic system. *Bilingualism: Language and Cognition* 1, 67-81.

Grenfell, M. (2002). *Modern Languages Across the Curriculum.* London: Routledge.

Grosjean, F. (1985). The bilingual as a competent but specific speaker-hearer. *Journal of Multilingual and Multicultural Development* 6, 467-477.

Grosjean, F. (1992). Another view of bilingualism. In Harris, R. *Cognitive Processing in Bilinguals* (51-62). Amsterdam: North Holland.

Grosjean, F. (1998). Studying bilinguals: Methodological and conceptual issues. *Bilingualism: Language and Cognition* 1, 131-149.

Grosjean, F. (1999). Individual bilingualism. In Spolsky, B. *Concise Ecyclopedia of Educational Linguistics*, 284-290. London: Elsevier.

Grosjean, F. (2001). The bilingual's language modes. In Nicol, J. *One Mind, Two Languages: Bilingual Language Processing.* Oxford: Blackwell, 1-25.

Grosjean, F. (2008). *Studying Bilinguals.* Oxford: Oxford University Press.

Hajer, M. (2000). Creating a language-promoting classroom: Content-area teachers at work. In Hall, J.K. and Verplaetse, L.S. *Second and Foreign Language Learning through Classroom*

Interaction, 265-285. Malwah, NJ: Lawrence Erlbaum Associates.

Hakuta, K. (1986). *Mirror of Language. The Debate on Bilingualism.* New York: Basic Books.

Halbach, A. (2009a). The Primary School Teacher and the Challenges of Bilingual Education. In Dafouz and Guerrini, M. (pp. 19-26). *CLIL across Educational Levels*. Richmond publishing. Madrid: Santillana.

Halbach, A. (2009b). Researching Bilingual Education in the Comunidad de Madrid. Paper presented at the *UAM CLIL Research Seminar*.

Hall, J.K. and Verplaetse, L.S. (2000). *Second and Foreign Language Learning through Classroom Interaction*. Malwah, NJ: Lawrence Erlbaum Associates.

Hamers, J.F. and Blanc, M.H. (2000). *Bilinguality and Bilingualism, 2nd ed.* Cambridge, UK: Cambridge University Press.

Harley, B. (1979). French gender "rules" in the speech of English-dominant, French-dominant, and monolingual French-speaking children. *Working Papers in Bilingualism,* 19, 129-156.

Harley, B. (1980). Interlanguage units and their relations. *Interlanguage Studies Bulletin* 5, 3-30.

Harley, B. (1984). How good is their French? *Language and Society,* 12 (Winter), 55-60.

Harley, B. (1986). *Age in Second Language Acquisition*. Clevedon: Multilingual Matters.

Harley, B. (1992). Patterns of second language development in French immersion. *Journal of French Language Studies* 2, 159-183.

Harley, B. (1998). The role of form-focused tasks in promoting child L2 acquisition. In Doughty, C. and Williams, J. *Focus on Form in Classroom Second Language Acquisition*. Cambridge University Press.

References

Harley, B. and King, M. (1989). Verb lexis in the written compositions of young L2 learners. *Studies in Second Language Acquisition* 11, 415-439.

Harley, B., Allen, P., Cummins, J., Swain, M. (1990). *The Development of Second Language Proficiency*. Cambridge: Cambridge University Press.

Haunold, C. (2006). English as a medium of instruction in Austrian secondary education. MA thesis. University of Vienna.

Hatch, E. (1978). Discourse analysis and second language acquisition. In Hatch, E. *Second Language Acquisition: A Book of Readings.* Rowley, MA: Newbury House.

Hellekjaer, G.O. (2006). *English for Occupational Purposes: the Case for CLIL*. Downloaded from the World Wide Web on February 6th 2007 at http://www.cinop.nl/ezine/mvt/nb_07-13/pp_hellekjaer.ppt#256,1.

Herdina, P. and Jessner, U. (2002). *A Dynamic Model of Multilingualism. Perspectives of Change in Psycholinguistics.* Clevedon: Multilingual Matters.

Hufeisen, B. and Lindemann, B. (1998). *L2-L3 und ihre Zwischensprachliche Interaktion: Zu Individueller Mehrsprachigkeit und Gesteuertem Lernen.* Tübingen: Stauffenburg.

Hufeisen, B. and Lindemann, B. (2005). *Gesamtsprachencurriculum – Integrierte Sprachendidaktik – Common Curriculum: Theoretische Überlegungen und Beispile der Umsetzung* (Giessener Beiträger zur Fremdsprachenforschung). Tübingen: Gunter Narr.

Hughes, A. (2003). *Testing for Language Teachers*. Cambridge: Cambridge University Press.

Hulstjin, J. (1989). Learning words meaning from context: An experimental comparison of the effects of four oriented cues. *Artikelen* 34: 13-25.

References

ISEI-IVEI (2007). *Alumnado trilingüe en secundaria: una nueva realidad*. On WWW at http://www.isei-ivei.net/cast/pub/Alumnado-triling-final.pdf.

Izumi, J. (2002). Output, input, enhancement, and the noticing hypothesis: An experimental study on ESL relativization. *Studies in Second Language Acquisition*, 24 (4), 541-77; 25 (1), 1-36.

Izumi, J. (2003). Comprehension and production processes in second language learning: In search of the psycholinguistic rationale of the output hypothesis. *Applied Linguistics*, 24 (2), 168-96.

Izumi, J., Bigelow, M., Fujiwara, M. and Fearnow, S. (1999). Testing the output hypothesis. *Studies in Second Language Acquisition*, 21, 421-452.

Jessner, U. (1999). Metalinguistic awareness in multilinguals: Cognitive aspects of third language learning. *Language Awareness* 8.3 & 4, 201-209.

Jessner, U. (2006). *Linguistic Awareness in Multilinguals: English as a Third Language*. Edinburgh: Edinburgh University Press.

Jessner, U. (2008). A DST-Model of multilingualism and the role of metalinguistic awareness. *Second Language Development as a Dynamic Process*. Special Issue of *Modern Language Journal* 92.2.

Jessner, U. (2008). Teaching third languages: Findings, trends and challenges. *Language Teaching*, 41:1, 15-56. Cambridge University Press.

Jiang, N. (2000). Lexical representation and development in a second language. *Applied Linguistics* 21, 47-77.

Johnson, K. (1988). Mistake correction. *ELT Journal* 42: 89-101.

Johnson, K. (1996). *Language Teaching and Skill Learning*. Oxford: Blackwell.

References

Johnson, K., Swain, M. (1997). *Immersion Education: International Perspectives.* Cambridge: CUP.

Jones, W.R. and Stewart, W.A.C. (1951). Bilingualism and verbal intelligence. *British Journal of Psychology* 4, 3-8.

Jourdenais, R., Ota, M., Stauffer, S., Boyson, B. and Doughty, C. (1995). Does textual enhancement promote noticing? A think aloud protocol. In Schmidt, R. (1995). *Attention and Awareness in Foreign Language Learning*, 183-216. Hawaii: University of Hawaii Press.

Juan-Garau, M. and Marcos-Llinàs, M. (submitted). Effects of Language Anxiety on Learners of Spanish as a FL. *The Modern Language Journal*.

Juan-Garau, M. (submitted). Oral fluency development in secondary-education CLIL learners. Paper submitted to *Vienna English Working Papers*. Special Isssue: *Research on CLIL 3*.

Juan-Garau, M., Marcos-Llinàs, M. and Salazar-Noguera, J. (2008). Enseñanza-aprendizaje de la lengua inglesa a través de contenidos en las Islas Baleares. In J. Salazar-Noguera, J. and Juan-Garau, M. (eds) *Aprendizaje integrado de lengua inglesa y contenidos multiculturales online*. Palma: Edicions UIB.

Junta de Andalucía (2005). Plan de Fomento del Plurilingüismo en Andalucía. Una política lingüística para la sociedad andaluza. Consejería de Educación de la Junta de Andalucía (2004). In: http://www.juntadeandalucia.es/boja/boletines/2005/65/d/5.htm. Last accessed 29th October 2009

Kerka, S. (2000). *Incidental Learning: Trends and Issues Alert N.18*. Columbus, OH: ERIC Clearinghouse.

Klapper, J. (1996). *Foreign-language Learning Through Immersion: Germany's Bilingual- Wing Schools.* Wales. Edwin Mellen Press.

Klieme, E. (2006). *Zusammenfassung zentraler Ergebnisse der DESI Studie*. Deutsches Institut für Internationale Pädagogische Forschung.

(www.dipf.de/desi/DESI_Ausgewaehlte_Ergebnisse.pdf) (last accessed 20 Sept 2006).

Kobayashi Y. (2002). The role of gender in foreign language learning attitudes: Japanese female students' attitudes towards learning English. *Gender Education* 14: 181-197.

Krashen, S.D. (1981). *Second Language Acquisition and Second Language Learning.* Oxford: Pergamon.

Krashen, S.D. (1982). *Principles and Practice in Second Language Acquisition.* Englewood Cliffs, NJ: Prentice-Hall.

Krashen, S.D. (1985)*. The Input Hipothesis: Issues and Implications*. New York: Longman.

Krashen, S.D. (1998). Comprehensible output*? System 26: 175-82.*

Krumm, H-J. (2005). Von der additiven zur curricularen Mehrsprachigkeit: Über die Notwendigkeit der Einbeziehung von Minderheiten-, Migranten- und Nachbarsprachen. In Hufeisen & Lutjeharms, 27-36.

Lambert, W. (1974). Culture and language as factors in learning and education. In Aboud, F. and Meade, R. *Cultural Factors in Learning and Education.* Bellingham, WA: Western Washington State University.

Lambert, W. (1981). Bilingualism and language acquisition. In Winitz, H. *Native Language and Foreign Language Acquisition.* New York: New York Academy of Sciences.

Lambert, W. and Tucker, R. (1972). *Bilingual Education of Children; The St. Lambert Experiment.* Rowley, MA: Newbury House.

Lapkin, S., Swain, M., Kamin, J. and Hanna, G. (1980). *Report on the 1979 Evaluation of the Peel County Late French Immersion Program, Grades 8,9,10, 11, and 12.* Toronto: Ontario Institute for Studies in Education.

Larsen-Freeman, D. (1997). Chaos/complexity science and second language acquisition. *Applied Linguistics* 18: 141-65.

References

Larsen-Freeman, D. and Long, M. (1991/1999). *An Introduction to Second Language Acquisition Research.* London: Longman.

Lasagabaster. D. (2000). Three languages and three linguistic models in the Basque educational system. In Cenoz, J. and Jessner, U. *English in Europe. The Acquisition of a Third Language*, 179-197. Clevedon: Multilingual Matters.

Lasagabaster. D. (2005). La presencia de tres lenguas en el currículo: multilingüismo en los contextos canadiense y español. *Revista de educación* 337, 405-426.

Lasagabaster, D. (2008). Foreign language competence in content and language integrated courses. *The Open Applied Linguistics Journal* 1, 31-42.

Lasagabaster, D. (2009). The implementation of CLIL and attitudes towards trilingualism. *ITL, International Journal of Applied Linguistics* 157, 23-43.

Lasagabaster, D. (in press). English achievement and students motivation in CLIL and EFL settings. *Innovation in Language Learning and Teaching.*

Lasagabaster, D. and Sierra, J.M. (2009). Language attitudes in CLIL and traditional EFL classes. *International Journal of CLIL Research* 1/2, 4-17.

Lasagabaster, D. and Sierra, J.M. (2010). Immersion and CLIL in English: more differences than similarities. *ELT Journal* 29/10, 1-9.

Lasagabaster, D., and Ruiz de Zarobe, Y. (2010). CLIL in a bilingual community: The Basque Autonomous Community. In Ruiz de Zarobe, Y. and Lasagabaster, D. *CLIL in Spain: Implementation, Results and Teacher Training.* Newcastle, UK: Cambridge Scholars Publishing.

Lasagabaster, D., and Ruiz de Zarobe, Y. (2010). Ways forward in CLIL: Provision issues and future planning. In Ruiz de Zarobe, Y. and Lasagabaster, D. *CLIL in Spain: Implementation, Results*

and Teacher Training. Newcastle, UK: Cambridge Scholars Publishing.

Lauridsen, K. and Toudic, D. (2008). *Languages at Work in Europe.* Göttingen: V&R unipress.

Lazaraton, A. (2001). Qualitative research methods in language test development and validation. In *European Language Testing in a Global Context*, 51-73. Proceedings of the ALTE Barcelona Conference. July 2001.

Lazaruk, W. (2007). Linguistic, Academic, and Cognitive Benefits of French Immersion. *The Canadian Modern Language Review* 63, 605-628.

Leow, R.P. (2001). Do learners notice enhanced forms while interacting with the L2? An online and offline study of the role of written input enhancement in L2 reading. *Hispania,* 84/3, 496-509.

Lewis, G. (1976). Bilingualism and bilingual education: The ancient world to the Renaissance. In Fishman, J. *Bilingual Education: An International Sociological Perspective*. Rowley, MA: Newbury House, 150-200.

Lewis, M. (1997). *Implementing the Lexical Approach: Putting Theory into Practice*. Hove: Language Teaching Publications.

Loewen, S. (2004). Uptake in incidental focus on form in meaning-focused ESL lessons. *Language Learning,* 54/1, 153-188.

Long, M.H. (1983). Native speaker / non-native speaker conversation and the negotiation of comprehensible input. *Applied Linguistics*, 4(2), 126-41.

Long, M.H. (1985). Input and second language acquistition theory. In Gass, S. and Madden, C. *Input in Second Language Acquisition*. Rowley: MA, Newbury House.

Long, M.H. (1991). Focus on form: a design feature in language teaching methodology. In de Bot, K., Ginsberg, R. and

References

Kramsch, C. *Foreign Language Research in Cross-cultural Perspective*. Amsterdam: John Benjamin.

Long, M.H. (1996). The role of the linguistic environment in second language acquisition. In Ritchie, W.C. and Bhatia, T.K. *Handbook of Second Language Acquisition*, 413-68. San Diego, CA: Academic Press.

Long, M.H. and Robinson, P. (1998). Focus on form: theory, research and practice. In Doughty, C. and Williams, J. *Focus on Form in Classroom Second Language Acquisition*. Cambridge University Press.

Lorenzo, F., Casal, S., and Moore, P. (2008). Los centros bilingües de la Comunidad Autónoma de Andalucía: gestión, competencias y organización [unpublished report].

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*, 20/11, 1-25.

Lyster, R. (1987). Speaking Immersion. *The Canadian Modern Language Review,* 43/4, 701-17.

Lyster, R. (1994). The effect of functional-analytic teaching on aspects of French immersion students' sociolinguistic competence. *Applied Linguistics,* 15, 263-287.

Lyster, R. (1998a). Negotiation of form, recasts, and explicit correction in relation to error types and learner repair in immersion classrooms. *Language Learning*, 48(2), 183-218.

Lyster, R. (1998b). Recasts, repetition and ambiguity in L2 classroom discourse. *Studies in Second Language Acquisition*, 20(1), 51-81.

Lyster, R. (2002). Negotiation in immersion teacher-student interaction. *International Journal of Educational Research* 37, 237-253.

References

Lyster, R. (2004). Differential effects of prompts and recasts in form-focused instruction. *Studies in Second Language Acquisition,* 26/3: 399-432.

Lyster, R. (2007)*. Learning and Teaching Languages Through Content. A Counterbalanced Approach*. Amsterdam/ Philadelplia: John Benjamins Publishing Company.

Lyster, R. and Ranta, L. (1997). Corrective feedback and learner uptake. *Studies in Second Language Acquisition*, 20, 37-66.

Llinares, A. (2006). A pragmatic analysis of children's interlanguage in EFL preschool contexts. *Intercultural Pragmatics,* 3-2, 171-193.

Llinares, A. (2007a). Young learners' functional use of the L2 in a low-immersion EFL context. *ELT Journal 61*, 39-45.

Llinares, A. (2007b). Classroom bilingualism at an early age: Towards a more natural EFL context. In Pérez Vidal, C., Juan-Garau, M. and Bel, A. *A Portrait of the Young in Multilingual Spain*. Bristol: Multilingual Matters.

Llinares, A. and Dafouz, E. (2010). Content and language integrated programmes in the Madrid region: Overview and research findings. In Ruiz de Zarobe, Y. and Lasagabaster, D. *CLIL in Spain: Implementation, Results and Teacher Training.* Newcastle, UK: Cambridge Scholars Publishing.

Llinares, A. and Whittaker, R. (2006). Oral and written production in social science. *ViewZ (Vienna English Working Papers)* 15/3, 28-32.

Llinares, A. and Whittaker, R. (2009). Writing and speaking in the history class: a comparative analysis of CLIL and first language contexts. In Dalton-Puffer, C., Nikula, T. and Smit, U. (eds) *Language Use in Content-and-Language Integrated Learning (CLIL)*. Amsterdam: John Benjamins.

MacIntyre, P. and Gardner, R. (1994). The subtle effects of induced anxiety on cognitive processing in the second language. *Language Learning*, 44, 283-305.

References

Mackey, A. (2006). Feedback, noticing and instructed second language learning. *Applied Linguistics*, 27/3, 405-430.

Mackey, A., Philp, J., Egi, T., Fujii, A., and Tatsumi, T. (2002). Individual differences in working memory, noticing of interactional feedback and L2 development. In Robinson, P. *Individual Differences in L2 Learning*. Amsterdam: John Benjamins.

Macnamara, J. (1967). The bilingual's linguistic performance: a psychological overview. *Journal of Social Issues*, 23, 59-77.

MacWhinney, B. (1997). Second language acquisition and the competition model. In de Groot, A.and Kroll, J.F. *Tutorials in Bilingualism: Psycholinguistic Perspectives*, 113-42. Mahwah, NJ: Lawrence Erlbaum.

Maljers, A., Marsh, D., Wolff, D. (2007). *Windows on CLIL. Content and Language Integrated Learning in the European Spotlight*. Graz, Austria: European Centre for Modern Languages.

Marsh, D. (2002). *Content and Language Integrated Learning. The European Dimension.* Jyväskylä. University of Jyväskylä Press.

Marsh, D. (2008). Language awareness and CLIL. In Cenoz, J. and Hornberger, NH. *Encyclopedia of Language and Education. Knowledge about Language*, 2nd edition, vol 6, New York: Springer Science + Business Media LLC, 233-246.

Marsh, D. and Marsland, B. (1999). *Distance in-Service Education for Enhancing Second Language Learning. DIESeLL.* Jyväskylä: University of Jyväskylä.

Marsh, D., Maljers, A. and Hartiala, A. (2001). *The CLIL Compendium. Profiling European CLIL Classrooms.* Jyväskylä: University of Jyväskylä.

Marsh, D., Maljers, A. and Hartiala, A. (2001). *Profiling European CLIL Classrooms: Languages Open Doors.* Jyväskylä. University of Jyväskylä and European Platform for Dutch Education.

Available in: http://www.cec.jyu.fi/kasvatusjaopetus/clil/clilprofiling

Marsh, D. and Wolff, D. (2006). *Diverse contexts – Converging Goals: CLIL in Europe*. Frankfurt: Peter Lang.

Marsol, A. (2008). Language use in the interactional space: A comparison of CLIL and traditional English language instruction. MA thesis, University of Barcelona.

Matiasek, S. (2005). English as the language of instruction in Austrian chemistry lessons: instance of explicit language teaching. MA thesis, University of Vienna.

McCabe, A., Llinares, A. and Whittaker, R. (2008). Introducing and tracking participants in CLIL history classes. Paper presented at 20th European Systemic Functional Linguistics Conference, University of Helsinki, Finland (11th to 13th of June).

McLaughlin, B. (1987). *Theories of Second Language Learning*. London: Edward Arnold.

McLaughlin, B. (1990). 'Conscious' versus 'unconscious' learning. *TESOL Quarterly* 24, 617-634.

McLaughlin, B. and Heredia, J. (1996). Information-processing approaches to research on second language acquisition and use. In Ritchie, W. and Bhatia, T. *Handbook of Second Language Acquisition.* New York, NY: Academic.

McLaughlin, B. and Nayak, N. (1989). Processing a new language: Does knowing a new language make a difference? In Dechert, H. and Raupach, M. *Interlingual Processes,* 5-16. Tübingen: Gunter Narr.

McNamara, T.F. (1996). *Measuring Second Language Performance*. London and New York: Longman.

McNamara, T.F. (2009). Personal communication at the University of Vienna. Vienna.

McNamara, T.F. and Roever, C. (2006) *Language Testing: The*

Social Dimension. Oxford: Blackwell.

McNamara, T.F. (2000) *Language Testing*. Oxford: Oxford University Press.

Meara, P. (1994). The year abroad and its effects. *Language Learning Journal*, 10, 32-38.

Mehotcheva, T. (2010). *After the fiesta is over. Foreign Language Attrition of Spanish in Dutch and German Erasmus Students.* Unpublished PhD Dissertation.

Meissner, F-J. (2004). Transfer und transferieren: Anleitungen zum Interkomprehensionsunterricht. In Klein, H. and Rutke, D. *Neuere Forschungen zur Europäischen Intercomprehension,* 39-66. Aachen: Shaker

Merisuo-Storm, T. (2007). Pupils' attitudes towards foreign-language learning and the development of literacy skills in bilingual education. In *Teaching and Teacher Education* 23/2, 226-235.

Met, M. (1994). Teaching content through a second language. In Genesee, F. *Educating Second Language Children* (159-182). New York: Cambridge University Press.

Mewald, C. (2004). Paradise lost and found: a case study of content based foreign language education in Lower Austria. Unpublished PhD Thesis. Norwich: University of East Anglia.

Mitchell, R. and Myles, F. (1998). *Second Language Learning Theories*. London: Hodder Arnold.

Mitchell, R. and Myles, F. (2004). *Second Language Learning Theories* (2nd edition). London: Hodder Arnold.

Mohan, B. and Beckett, G. (2003). A functional approach to research on content-based language learning: recasts in causal explanations. *The Modern Language Journal*, 87/3, 421-432.

Mohanty, A. and Perregaux, C. (1997). Language acquisition and bilingualism. In Berry, J., Dasen, P., and Saraswathy, S. *Handbook of Cross-Cultural Psychology. Vol 2: Basic*

Processes and Human Development, 217-53. Boston, MA: allyn and Bacon.

Moore, P. (2007). Enhancing classroom discourse: a modelling potential for content teachers. In Lorenzo, F. *et al*. *Models and Practice in CLIL*, 141-152. La Rioja: Revista Española de Lingüística Aplicada.

Moore, P. (2009). *On the Emergence of L2 Oracy in Bilingual Education: A Comparative Analysis of CLIL and Mainstream Learner Talk*. Sevilla: Universidad de Sevilla. Unpublished Doctoral Dissertation.

Mougeon, R. and Rehner, K. (2001). Acquisition of sociolinguistic variants by French immersion students: *The Case of Restrictive Expressions, and More*. The Modern Language Journal 85, 398-415.

Muñoz, C. (2000). Bilingualism and trilingualism in school students in Catalonia. In Cenoz, J. and Jessner, U. *English in Europe. The Acquisition of a Third Language*, 157-178. Clevedon: Multilingual Matters.

Muñoz, C. (2006). The effects of age on foreign language learning: The BAF project. In Muñoz, C. *Age and the Rate of Foreign Language Learning.* Clevedon: Multilingual Matters.

Muñoz, C. (2007). CLIL: Some thoughts on its psycholinguistic principles. In *Modelos Y Prácticas En Eicle. Revista Española de Lingüística Aplicada (RESLA).*

Muñoz, C. and Navés, T. (2007). Spain. In Maljers, A., Marsh, D., Wolff, D. *Windows on CLIL. Content and Language Integrated Learning in the European Spotlight.* Graz: European Platform for Dutch Education and European Centre for Modern Languages.

Muranoi, H. (2007). Output practice in the L2 classroom. In DeKeyser, R. *Practice in a Second Language. Perspectives from Applied Linguistics and Cognitive Psychology*, 51-84. Cambridge: Cambridge University Press.

References

Naiman, N. (1995). *The Good Language Learner*. Clevedon: Multilingual Matters.

Naiman, N., Fröhlich, M., Stern, H., and Todesco, A. (1978). The good language learner. *Research in Education Series* 7. Toronto: The Ontario Institute for Studies in Education. Reprinted in 1996 by Multilingual Matters.

Nation, R., and McLaughlin, B. (1986). Novices and experts: An information processing approach to the good language learner problem. *Applied Psycholinguistics* 7, 41-56.

Navés, T. and Muñoz, C. (1999) Implementation of CLIL in Spain. In Marsh, D. and Langé, G. (eds) *Implementing Content and Language Integrated Learning: A Research Driven TIE-CLIL Foundation Course Reader*. Jyväskylä: University of Jyväskylä.

Navés, T. (2006). *The Long Terms Effects of an Early Start on Writing in Foreign Language Learning*. Ph.D. [Dissertation]. Barcelona: University of Barcelona.

Navés, T. Miralpeix, I. and Celaya M.L. (2005). Who transfers more... and what? Crosslinguistic influence in relation to school grade and language dominance in EFL. *International Journal of Multilingualism* 2 (2), 113–134.

Navés, T. and Victori, M. (2010). CLIL in Catalonia: An overview of research studies. In Ruiz de Zarobe, Y. and Lasagabaster, D. *CLIL in Spain: Implementation, Results and Teacher Training*. Newcastle upon Tyne: Cambridge Scholars Publishing.

Nayak, N., Hansen, N., Krueger, N., and McLaughlin, B. (1990). Language-learning strategies in monolingual and multilingual adults. *Language Learning*, 40 (2), 221-244.

Nikula, T. (1997) Terminological considerations in teaching content through a foreign language. In Marsh, D., Marland, B. and Nikula, T. *Aspects of Implementing Plurilingual Education: Seminars and Field Notes,* 5-8. The Continuing education center: University of Jyväskylä.

References

Nikula, T. and Marsh, D. (1997). *Language and Content Integrated Learning in the Primary and Secondary School Sector.* Helsinki: National Board of Education.

Norris, J. and Ortega, L. (2000). Effectiveness of L2 instruction: A research synthesis and quantitative meta-analysis. *Language Learning*, 50, 417-528.

Norris, J. and Ortega, L. (2003). Defining and measuring SLA. In Doughty, C. and Long, M. *Handbook of Second Language Acquisition*. Oxford: Blackwell.

Nussbaum, L. (1990) Plurilingualism in foreign language classrooms in Catalonia. In *Network on Code-Switching and Language Contact. Papers for the Workshop on Impact and Consequences: Broader Considerations* (pp. 141-165). Strasbourg: ESF.

Nyholm, L. (2002). Läsa på engelska. Språkvård 3:38.

Oliver, R. and Mackey, A. (2003). Interactional context and feedback in child ESL classrooms. *The Modern Language Journal*, 87/4, 519-533.

Oller, J.W. (1979). *Language Tests at School.* London: Longman.

Oller, J.W. (1980). *Research in Language Testing.* Rowley, Mass: Newbury House Publication.

Oller, J.W. (1993). *Methods that Work.* Boston, Mass: Heinle & Heinle.

Onwuegbuzie A., Bailey P. and Daley C. (2001). Cognitive, affective, personality, and demographic predictors of foreign language achievement. *Journal Educational Research* 94: 3-15.

Oxford R.L. (1993). Instructional implications of gender differences in language learning styles and strategies. *Applied Language Learning* 4: 65-94.

Oxford, R.L. and Anderson, N.J. (1995). A crosscultural view of learning styles. *Language Teaching,* 28, 201-215.

References

Paivio, A. (1986). *Mental Representations: A Dual Coding Approach.* Oxford: Oxford University Press.

Pavlenko A, Piller, I. (2008). Language education and gender. In: May S, Hornberger NH. *Encyclopedia of Language and Education*, 2nd edition, volume 1, Language policy and political issues in education, New York: Springer Science+Business Media LLC 57-69.

Peal, E. and Lambert, W. (1962). The relation of bilingualism to intelligence. *Psychological Monographs,* 76, 1-23.

Pérez-Vidal, C. (1996). *La adquisición del inglés de un niño bilingüe (catalán/ingles).* Barcelona: Universidad de Barcelona.

Pérez-Vidal, C. (1997). Language teacher training and bilingual education in Spain. In Van de Craen, P. and Wolff, D. *Language Teacher Training and Bilingual Education' Technical report. Thematic Network Project n.9 Bilingual Education. Evaluation Conference.* Lille: University of Lille III.

Pérez-Vidal, C. (1999). Teacher education: The path towards CLIL. In Marsh, D. and Marsland, B. (eds.). *CLIL Initiatives for the Millenium: Report on the CEILINK Think-Tank*, 16-18.Jyväskylä, University of Jyväskylä, Continuing Education Centre.

Pérez-Vidal, C. (2001a). An introductory note: Input and interaction in language acquisition. In Safont, P. and Campoy, M.C. (eds.). *Oral Skills: Resources and Proposals for the Classroom* 19-32. Barcelona: Universitat Jaume I.

Pérez-Vidal, C. (2001b). Perfiles del Aprendizaje Integrado de Contenidos Curriculares y Lenguas Extranjeras (AICLE) en Europa: Las lenguas abren puertas. In Marsh, D., Maljers, A., and Hartiala, A. (eds.) *Profiling European CLIL Classrooms: Languages Open Doors*, 215-253. Jyväskylä, University of Jyväskylä & European Platform for Dutch Education.

Pérez -Vidal, C. (2001c). Multilingual Education: Problems and solutions in Catalonia. In Van de Craen, P. and Pérez-Vidal, C. *The Multilingual Challenge/Le défi Multilingue* Final Report 2000, 41-57. Barcelona: Printulibro.

References

Pérez-Vidal, C. (2004a). Lectures and talks given on Content and Language Integrated Learning in Europe: A European approach to education. In Pérez-Vidal, C. *Content and Language Integrated Learning in Europe: Teaching Materials for the Secondary School Classroom*, 15-39. Barcelona: European Commission, Pompeu Fabra University.

Pérez-Vidal, C. (2004b). *Content and Language Integrated Learning in Europe: Teaching Materials for the Secondary School Classroom*. Barcelona: European Commission, Pompeu Fabra University.

Pérez-Vidal, C. (2005a). 'Teachers'strategies in content and language integrated learning classrooms" Paper presented at the 5 International Symposium on Bilingualism. Barcelona, March.

Pérez-Vidal, C. (2005b). Lectures and Talks given on Content and Language Integrated Learning: A European approach to education (Chpt 2) In *Content and Language Integrated Learning (CLIL). Teaching materials for use in the Secondary School Classroom*. Barcelona: European Commision, Pompeu Fabra University.

Pérez-Vidal, C. (2007a). The European framework of reference: the Portfolio. In Usó, J. and Ruiz-Madrid, N. *Pedagogical Reflections on Learning Languages in Instructed Settings.* Cambridge: Cambridge Scholar Press.

Pérez-Vidal, C. (2007b). Multilingualism in the classroom: A question of context, contact, and tact. Symposium on: *New Challenges in Foreign Language Education: Content and Language Integrated Learning and Intercultural Language Teaching.* University of Lleida.

Pérez-Vidal, C. (2007c). The need for focus on form in content and language integrated approaches: an exploratory study. In *Modelos Y Prácticas En Eicle. Revista Española de Lingüística Aplicada (RESLA).*

Pérez-Vidal, C. (2008). La enseñanza bilingüe: El Enfoque Integrado de Contenidos y Lenguas en Europa. In *Aula de Innovación*

Educativa. Editorial GRAÓ.

Pérez-Vidal, C. (2009a). The integration of content and language in the classroom: A European approach to education (the second time around). In Dafouz, E. and Guerini, M. *CLIL Across Educational Levels*, 3-17. Richmond publishing. Madrid: Santillana.

Pérez-Vidal, C. (2009b) The effects of context and input on language acquisition: A Study Abroad (SA) perspective. *Eurosl'19Cork. Language Learning Roundtable.*

Pérez-Vidal, C. (2011). Language acquisition in three different contexts of learning: formal instruction, stay abroad, and semi-immersion (CLIL). In Ruiz de Zarobe, Y., Sierra, J.M. and Gallardo, F. *Content and Foreign Language Integrated Learning: Contributions to Multilingualism in European Contexts.* Spain: Peter Lang Publications.

Pérez-Vidal, C. (forthcoming). Does Content and Language Integrated Learning (CLIL) Education Suffice to learn a Foreign Language? In Abello-Contesse, C. and Chacón, R. *Studies in Applied English Linguistics / Estudios de lingüística inglesa aplicada (ELIA) 10.*

Pérez-Vidal, C. and Campanale, N, (2006). *Content and Language Integrated Learning (CLIL). Teaching Materials for Use in the Secondary School Classroom.* Barcelona: European Commision/Universitat Pompeu Fabra.

Pérez -Vidal, C. and Escobar, C. (2002). The case of Spain. In Grenfell, M. (ed.) *Modern Languages Across the Curriculum,* 114-130. London: Routledge.

Pérez -Vidal, C. and Escobar, C. (2004). Teacher education for the implementation of CLIL in the school system. In Wilkinson, R. (ed.) *Meeting the Challenge of a Multilingual Higher Education.* Netherlands: Maastricht University.

Pérez-Vidal, C., Escobar, C. and Rebull, E. (2005). Multilingualism in the classroom in Catalonia/El multilingüisme a les aules a

Catalunya. (Audiovisual documentary): Comisión europea/Universidad Pompeu Fabra.

Pérez -Vidal, C., Juan-Garau, M., Bel, A. (2008). *A Portrait of the Young in the New Multilingual Spain*. Clevendon: Multilingual Matters.

Pérez-Vidal, C. and Juan-Garau, M. (2010). To CLIL or not to CLIL? From bilingualism to multilingualism in Catalan/Spanish communities in Spain. In Ruiz de Zarobe, Y. and Lasagabaster, D. (eds) *CLIL in Spain: Implementation, Results and Teacher Training.* CUP: Cambridge Scholars

Pérez-Vidal, C., Torras, M.R., Celaya, M.L. (2000). Age and EFL Written Performance by Catalan/Spanish Bilinguals. *Spanish Applied Psycholinguistics*, 4-2, 267-290. University of Illinois.

Pérez, P. and Zayas, F. (2007). *Competencia en comunicación lingüística*. Madrid: Alianza.

Plan de fomento del Plurilingüismo. Una política lingüística para la sociedad andaluza (2005). Comunidad Autónoma de Andalucía.

Polytechnic University of Barcelona (2009). *English at UPC* (unpublished manuscript).

Pompeu Fabra University (2007). *Plan of Action for Multilingualism.* Barcelona: Pompeu Fabra University.

Pons, E. and Sorolla, N. (2009). *Informe de la Llengua Catalana (2005-2007)*. Barcelona: Observatori de la Llengua Catalana - CRUSCAT. Dossier 13.

Poulisse, N. (1997). Language production in bilinguals. In de Grot, A. and Kroll, J. *Tutorials in Bilingualism: Psycholinguistic Perspective* (201-224). Mahwah, NJ: Lawrence Erlbaum.

Räsänen, A. (1993). Adjustments in teaching styles. *Council of Europe: Language Learning for European Citizenship*, 95 Report 12ª. Strasbourg: council of Europe.

References

Rebuffot, J. (1993). *Le Point sur l'Immersion au Canada*. Montreal, QC: Éditions CEC.

Rehner, K. and Mougeon, R. (1999). Variation in the spoken French of immersion students: To 'ne' or not to 'ne', that is the sociolinguistic question. *The Canadian Modern Language Review* 56, 124-154.

Reilly, T. and Medrano, P. (2009). MEC/British Council Bilingual Project. Twelve years of education and a smooth transition into secondary. In Dafouz, E. and Guerrini, M. (pp. 59-70). *CLIL across Educational Levels*. Richmond publishing. Madrid: Santillana.

Richards, J. and Rodgers, T. (2001). *Approaches and Methods in Language Teaching.* Cambridge: Cambridge University Presss.

Rieder, A. and Hüttner, J. (2007). *EFL Narrative Competences of Secondary School Pupils: the Influence of Intensified Primary School Instruction.* Ms. Vienna University.

Rifkin B. (2005). A ceiling effect in traditional classroom foreign language instruction: data from Russian. *Modern Language Journal* 89: 3-18.

Ringbom, H. (1987). *The Role of the First Language in Foreign Language Learning.* Clevedon: Multilingual Matters.

Robinson, P. (1997). Individual differences and the fundamental similarity of implicit and explicit adult language learning. *Language Learning* 47*:* 45-99.

Robinson, P. (2002). *Individual Differences and Instructed Language Learning*. Amsterdam, John Benjamins.

Roquet, H. (2005). Bilingual reading: an essential factor to writing acquisition in a third language. Master Thesis (unpublished). Barcelona: Universitat Pompeu Fabra.

Roquet, M., Escobar,C., and Cuscó, I. ' Is it possible to teach history in English to students with a limited command of the language?'. Action-research internal report (unpublished).

References
Rovira i Virgili University (2008). *Pla de política lingüística de la Universitat Rovira i Virgili per al 2009-2011* (unpublished manuscript).

Ruiz de Zarobe, Y. (2007). CLIL in a bilingual community: similarities and differences with the learning of English as a foreign language. Paper presented at CLIL Symposium, Viena 20-22 September 2007.

Ruiz de Zarobe, Y. (2008). CLIL and foreign language learning: A longitudinal study in the Basque Country. *International CLIL Research Journal.* 1 (1): 60-73.

Ruiz de Zarobe, Y. and Jiménez Catalán, R. (2009). *Content and Language Integrated Learning. Evidence from Research in Europe*. Clevedon: Multilingual Matters.

Ruiz de Zarobe, Y. and Lasagabaster, D. (2010). *CLIL in Spain: Implementation, Results and Teacher Training*. Newcastle, UK: Cambridge Scholars Publishing.

Ruiz de Zarobe, Y., Sierra, J.M. and Gallardo del Puerto. F. (2011). *Content and Foreign Language Integrated Learning: Contributions to Multilingualism in European Contexts.* Bern, Berlin, Bruxelles, Frankfurt am Main, New York, Oxford, Wien: Peter Lang (Linguistic Insight series).

Saer, D.J. (1923). The effect of bilingualism on intelligence. *British Journal of Psychology* 14, 25-38.

Safont-Jordà, M.P. (2005). *Third Language Learners: Pragmatic Production and Awareness*. Clevedon: Multilingual Matters.

Sagasta, M. (2003). Acquiring writing skills in a third language: Positive effects of bilingualism. *International Journal of Bilingualism* 7 (1), 27-42.

Salaberry, M. (1997). The role of input and output practice in second language acquisition. *The Canadian Modern Language Review* 53, 422-451.

References

Salazar-Noguera, J. and Juan-Garau, M. (eds) (2008) *Aprendizaje integrado de lengua inglesa y contenidos multiculturales online.* Palma: Edicions UIB.

San Isidro, X. (2009a). *CLIL: Integrando linguas "a través" do currículo*. Xunta de Galicia: Consellería de Educación e Ordenación Universitaria.

San Isidro, X. (2009 b). Galicia, CLIL success in a bilingual community. In Marsh, D. *et al. CLIL Practice: Perspectives from the Field,* 5-13. Finland: University of Jyväskylä.

San Isidro, X. (2010). Análisis del modelo CLIL gallego. In Escobar, C., Evnitskaya, N., Moore, E. and Patiño, A. *AICLE/CLIL/EMILE. Educació Plurilingüe: Experiencias, Research & Politiques.* Barcelona: Publicaciones UAB.

Sanz, C. (2000). Bilingual education enhances third language acquisition: Evidence from Catalonia. *Applied Psycholinguistics* 21, 23-44.

Sanz, C. (2003). Computer delivered implicit vs. explicit feedback in processing instruction. In VanPatten, B. *Processing Instruction: Theory, Research and Commentary* (Second Language Acquisition Research). Mahwah, N.J.: Lawrence Erlbaum.

Sanz, C. (2008). Predicting enhanced L3 learning in bilingual contexts: The role of biliteracy. In Pérez-Vidal, C; Juan-Garau, M. and Bel, A. *A Portrait of the Young in the New Multilingual Spain,* 220-240. Clevedon: Multilingual Matters.

Schmid, S. (1995). Multilingualer Fremdsprachenunterricht: Ein didaktischer Versuch mit Lernstrategien. *Multilingua* 15, 55-90.

Schmidt, R. (1990). The role of consciousness in second language learning. *Applied Linguistics*, 12/2, 17-46.

Schmidt, R. (1995). *Attention and Awareness in Foreign Language Learning*. Hawaii: University of Hawaii Press.

Schmidt, R. (2001). Attention. In Robinson, P. *Cognition and Second Language Instruction.* New York, NY: Cambridge University Press, 3-32.

References

Schmidt R., Boraie D. and Kassagby, O. (1996). Foreign language motivation: international structure and external connections. In: Oxford R.L. *Language Learning Motivation: Pathways to the New Century*. Hawaii: University of Hawai 9-20.

Schmitt, N. (2002 1st ed., 2010). *An Introduction to Applied Linguistics*. London: Hodder Education.

Segalowitz, N. (2003). Automaticity and second languages. In Doughty, C. and Long, M. *The Handbook of Second Language Acquisition*. United Kingdom: Blackwell Publishers, 382-408.

Seidlhofer, B. (2005). Standard future or half-baked quackery? Descriptive and pedagogic bearings on the globalisation of English. In Gnutzmann, C. and Intemann, F. *The Globalisation of English and the English Language Classroom*. Tübingen: Narr, 159-173.

Selinker, L., Swain, M. and Dumas, G. (1975). The interlanguage hypothesis extended to children. *Language Learning* 25, 139-152.

Sharwood, M. (1993). Input enhancement in instructed SLA: theoretical bases. *Studies in Second Language Acquisition*, 15, 165-179.

Sheen, Y. (2006). Corrective feedback, individual differences and the acquisition of articles by second language learners. Unpublished PhD thesis, University of Nottingham.

Shehadeh, A. (2003). Learner output, hypothesis testing, and internalizing linguistic knowledge. *System*, 31/2, 155-171.

Sierra, J. (1994). Modelos de ensenanza bilingüe: sus resultados y su futuro. In Siguán, M. *Las lenguas en la Europa Comunitaria: las lenguas de minorías.* Barcelona: ICE de la Universitat de Barcelona.

Singleton, D. (1989). *Language Acquisition. The Age Factor.* Clevedon: Multilingual Matters.

Singleton, D. (2005). The critical period hypothesis. A coat of many colours. *International Review of Applied Linguistics* 10: 209-31.

References

Singley, M.K. and Anderson, J.R. (1989). *The Transfer of Cognitive Skill.* Cambridge, MA: Harvard University Press.

Skehan, P. (1989). *Individual Differences in Second-Language Learning.* London: Edward Arnold.

Skehan, P. (1991). Individual differences in second language learning. *Studies in Second Language Acquisition* 13: 275-98.

Skehan, P. (1998). *A Cognitive Approach to Language Learning.* Oxford: Oxford University Press.

Skinner, B. (1957). *Verbal Behaviour*. New York: Appleton-Century-Crofts.

Snow, M., Met, M. and Genesee, F. (1989). A conceptual framework for the integration of language and content in Second/Foreign Language instruction. In *TESOL Quarterly* 23_2: 201-217.

Spielman, G. and Radnofsky, M. (2001). Learning language under tension: new directions from a qualitative study. *The Modern Language Journal*, 85, 259-78.

Spolsky, B. (1988). *Conditions for Second Language Learning: Towards a General Theory.* Oxford: Oxford University Press.

Sunderland J. (2000). Issues of language and gender in second and foreign language education. *Language Teaching* 33: 203-223.

Swain, M. (1984). A review of immersion education in Canada: Research and evaluation studies. In *Studies on Immersion Education: A Collection for U.S. Educators*. Sacramento: California State Department of Education. 87-112. Revised and printed in *ELT Documents* 119, 35-51.

Swain, M. (1985). Communicative competence: Some roles of comprehensible input and comprehensible output in its development. In Gass, S. and Madden, C. *Input in Second Language Acquisition* ,235-53. Rowley, MA: Newbury House.

References

Swain, M. (1993). The Output Hypothesis: Just speaking and writing aren't enough. *Canadian Modern Language Review*, 50, 158-164.

Swain, M. (1995). Three functions of output in second language learning. In Cook, G. and Seidlhofer, B. *Principles and Practice in Applied Linguistics*, 125-144. Oxford: Oxford University Press.

Swain, M. (1998). Focus on Form through conscious reflection. In Doughty, C. and Williams, J. *Focus on Form in Classroom Second Language Acquisition*, 64-81. New York: Cambridge University Press.

Swain, M. (2005). The Output Hypothesis: theory and research. In Hinkel, E. *Handbook on Research in Second Language Teaching and Learning*, 471-484. Mahwah, NJ: Lawrence Erlbaum.

Swain, M. and Lapkin, S. (1982). *Evaluating Bilingual Education: a Canadian Case Study.* Clevedon: Multilingual Matters.

Swain, M., Lapkin, S., Rowen, N. and Hart, D. (1990). The role of mother tongue literacy in third language learning. *Language, Culture and Curriculum* 3 (1), 65-81.

Swain, M. and Lapkin, S. (1991). Additive bilingualism and French immersion education: The roles of language proficiency and literacy. In Reynolds, A.G. *Bilingualism, Multiculturalism and Second Language Learning: The McGill Conference in Honour of Wallace E. Lambert,* 203-16. Hillsdale, NJ: Lawrence Erlbaum.

Swain, M. and Lapkin, S. (1991). Aspects of the sociolinguistic performance of early and late French immersion students. In Scarcella, R., Andersen, E., and Krashen, S. *Developing Communicative Competence in a Second Language* (41-54). New York: Newbury House.

Swain, M. and Lapkin, S. (1995). Problems in output and the cognitive processes they generate: A step towards second language learning. *Applied Linguistics,* 16(3), 371-391.

Swain, M. and Lapkin, S. (1998). Interaction and second language learning: Two adolescent French immersion students working together. *The Modern Language Journal 82/3*, 320-337.

References

Swan, D. (1996). *A Singular Pluralism: The European Schools 1984-1994*. Dublin: Institute of Public Administration.

Sylvén, L. (2004). Teaching in English or English teaching? On the effects of content and languag integrated learning on Swedish learners' incidental vocabulary acquisition. PhD dissertation. Göteborg University.

Takahashi, S. (2005). Pragmalinguistic awareness: is it related to motivation and proficiency? *Applied Linguistics 26*: 90-120.

Tavakoli, P. and Foster, P. (2008). Task design and second language performance. The effect of narrative type on learner output. *Language Learning* 58, 2, 439-473.

Thomas, J. (1988). The role played by metalinguistic awareness in second and third language learning. *Journal of Multilingual and Multicultural Development* 9, 235-246.

Ting, T. (2011). CLIL and Neuroscience: How are they related? In Ruiz de Zarobe, Y., Sierra, J.M. and Gallardo del Puerto. F. *Content and Foreign Language Integrated Learning: Contributions to Multilingualism in European Contexts*, 75-101. Bern, Berlin, Bruxelles, Frankfurt am Main, New York, Oxford, Wien: Peter Lang (Linguistic Insight series).

Tomlin, R. and Villa, V. (1994). Attention in cognitive science and second language acquisition. *Studies in Second Language Acquisition*, 16, 183-203.

Turnbull, M., Lapkin, S., Hart, D. and Swain, M. (1998). Time on task and immersion graduates French proficiency. In Lapkin, S. *French Second Language Education in Canada: Empirical Studies*. Toronto: University of Toronto Press, 31-55.

Ur, P. (1996). *A Course in Language Teaching*. Cambridge: Cambridge University Press.

References

Valdés, G. and Figueroa, R. (1994). *Bilingualism and Testing: A Special Case of Bias.* Norwood, NJ: Ablex Publishing.

Vallbona, A. (2009) Implementing CLIL methodology in two courses of primary education in a state-funded school: A case study. MA thesis, Bellaterra, UAB.

Valls-Ferrer, M., Roquet-Pugès, H., Pérez-Vidal, C. (2010). The Effect of Practice in Different Contexts of Language Acquisition: Contrasting FI, SA and CBLC. In Caballero, R., and Pinar, M.J. *Ways and Modes of Human Communication*. On-line serial 129 (219-229). Cuenca: Ediciones de la Universidad de Castilla La Mancha, AESLA.

Van de Craen, P. and Pérez-Vidal, C. (2001). *The Multilingual Challenge/Le Défi Multilingue Final Report 2000.* Barcelona: European Commission-Pompeu Fabra University.

Van de Craen, P. and Wolff, D. (1997). *Language Teacher Training and Bilingual Education' Technical Report. Thematic Network Project n.9 Bilingual Education. Evaluation Conference.* Lille: University of Lille III.

Van de Craen, P., Ceuleers, E., Lochtman, K., Allain, L., Monst, K. (2006). An interdisciplinary research approach to CLIL learning in primary schools in Brussels. In Dalton-Puffer, C. and Smit, U. (2007). *Empirical Perspectives on CLIL Classroom Discourse.* Frankfurt, Vienna: Peter Lang.

Van de Craen, P., Ceuleers, E., Mondt, K., Allain, L. (2008). European multilingual language policies in Belgium and policy-driven research. In Lauridsen, K. and Toudic, D. *Languages at Work in Europe.* Göttingen: V&R unipress: 139-151.

Van Leeuwen, Ch. (2006). From assessment anecdotes to learning practices. Eds. Van Leeuwen, Ch. and Wilkinson, B. Nijmegen: Valkhof Pers. 11-25.

Van Leeuwen, Ch., Zegers, V. and Wilkinson, B. (2003). *Multilingual Approaches in Higher Education: Challenges and Practices.* Nijmegen: Valkhof Pers.

References

VanPatten, B. (1990). Attending to form and content in the input. *Studies in Second Language Acquisition* 12: 287-301.

VanPatten, B. (1996). *Input Processing and Grammar Instruction. Theory and Research*. Norwood, NJ: Ablex.

VanPatten, B. (2003). *Processing Instruction. Theory, Research and Commentary*. Mahwah, NJ: Lawrence Erlbaum.

VanPatten, B. (2004a,b). Input processing in second language acquisition. In VanPatten, B. *Processing Instruction. Theory, Research and Commentary*. Mahwah, NJ: Lawrence Erlbaum.

VanPatten, B. and Cadierno, T. (1993a). Explicit instruction and input processing. *Studies in Second Language Acquisition*, 15, 225-241.

VanPatten, B. and Cadierno, T. (1993b). Input processing and second language acquisition: A role for instruction. *Modern Language Journal*, 77, 45-57.

VanPatten, B. and Oikkenon, S. (1996). Explanation versus structured input in processing instruction. *Studies in Second Language Acquisition*, 18, 495-510.

Vázquez, A. and Llinares, A. (2009). A Study of Linguistic Transfer in CLIL Students' Oral Discourse. Paper presented at XXVII AESLA International Conference, Ciudad Real (Spain).

Victori, M. and Vallbona, A. (2008). A case study on the implementation of CLIL methodology in a primary education school: results, benefits and challenges. Paper presented at the *CLIL-TBL Seminar*, UB, Barcelona, December 12.

Vila-Moreno, X. (2008). Language-in-education policies in the Catalan language area. In J. Cenoz, J. and Gorter, D. (eds) *AILA REVIEW. Multilingualism and Minority Languages: Achievements and Challenges in Education*. Amsterdam/Philadelphia: John Benjamins.

Villareal, O. and García-Mayo, M.P. (2009). Tense and agreement morphology in the interlanguage of Basque and Spanish

bilinguals: CLIL vs. Non-CLIL. In Ruiz de Zarobe, Y. and Jimenez Catalán, R. *Content and Language Integrated Learning. Evidence from Research in Europe*. Clevedon: Multilingual Matters.

Vollmer, H., Heine, L., Troschke, R., Coetzee, D. and Küttel, V. (2006). Subject-specific competence and language use of CLIL learners: The case of geography in grade 10 of secondary schools in Germany. Paper presented at the ESSE 8 Conference in London, 29 August 2006.

Walsh, S. (2006). *Investigating Classroom Discourse*. London: Routledge.

Washburn, L. (1997). English immersion in Sweden: A case study of Röllingby high school 1987-1989. Unpublished PhD thesis. Stockholm University.

Weir, C.J. (1998). *Communicative Language Testing with Special Reference to English as a Foreign Language.* Volume 11. Exeter Linguistic Studies. University of Exeter.

Wesche, M. (2002). Early French immersion: How has the original Canadian model stood the test of time? In: Petra Burmeister, P., Piske, T. and Rohde, A. *An Integrated View of Language Development*. Papers in honour of Henning Wode. Trier: Wissenschaftlicher Verlag Trier, 357-379.

Whittaker, R. and Llinares, A. (2009). CLIL in social science classrooms: analysis of spoken and written productions. In Ruiz de Zarobe, Y. and Jimenez Catalán, R. *Content and Language Integrated Learning. Evidence from Research in Europe*. Clevedon: Multilingual Matters.

White, L. (1991). Adverb placement in second language acquisition. Some effects of positive and negative evidence in the classroom. *Second Language Research*, 7, 133-61.

Williams, S. and Hammarberg, B. (1998) Language switches in L3 production: Implications for a polyglot speaking model. *Applied Linguistics* 19, 295-333.

References

Wilkinson, B. (2004). *Integrating Content and Language. Meeting the Challenge of Multilingual Higher Education.* Netherlands: Maastricht University.

Wilkinson, B., Zegers, V. and Van Leeuwen, Ch. (2006). *Bridging the Assessment Gap in English-Medium Higher Education.* Nijmegen: Valkhof Pers.

Wode, H. (1994). *Bilinguale Unterrichtserprobung in Schleswig-Holstein.* Band II Analytische Auswertungen. Kiel: l&f Verlag.

Wolff, D. (1998). CLIL Initiatives for the Milleniu. In Marsh, D. and Marsland, B. Report on the CEILINK Think-Tank. Jyväskylä : University of Jyväskylä, 20-36.

Wolfe-Quintero, K., Inagaki, S., Kim, H. (1998). *Second Language Development in Writing: Measures of Fluency, Accuracy and Complexity.* Hawai'i: University of Hawai'i at Manoa.

Ytsma, J. (2000). Trilingual Primary Education in Friesland. In Cenoz, J. and Jessner, U. *English in Europe: The Acquisition of a Third Language* (222-235). Clevedon: Multilingual Matters.

Zydatiß, W. (2006). Bilingualer Unterricht = höhere Sprachfähigkeit: Konsequenzen für die Bewertungspraxis an deutschen Schulen. In Timm, J. (Hg.): *Fremdsprachenlernen und Fremdsprachenforschung: Kompetenzen, Standards, Lernformen, Evaluation.* Festschrift für Helmut Johannes Vollmer. Narr, Tübingen, 373-388.

Zyzik, E. and Polio, C. (2008). Incidental focus on form in university Spanish literature courses. *The Modern Language Journal,* 92/1, 53-70.

Zobl, H. (1992). Prior Linguistic Knowledge and the conservation of the learning procedure: Grammaticality judgements of unilingual and multilingual learners. In Gass, S. and Selinker, L. *Language Transfer in Language Learning*, 176-196. Amsterdam: John Benjamins.

References

MAGAZINES

Resla (2007) n extra 1. Models and practice in CLIL. http://dialnet.unirioja.es (Monográfico de la Revista de la asociación española de lingüística aplicada –RESLA-, que recoge investigaciones y experiencias AICLE).

Cuadernos de pedagogía (2009) n 395. Enseñar en inglés. Cuando la lengua de la escuela es diferente de la lengua familiar. Barcelona: Wolters Kluwer.

WEBSITES

IES Icària, Barcelona
http://iesicaria.xtec.es

Equipo colaborativo CLIL-SI
http://www.clil-si.org

Portal CLIL de la Universitat de Barcelona
http://www.ub.edu/filoan/CLIL.html

EUROCLIC
http://www.euroclic.net

Eurydice
http://www.eurydice.org

LIST OF APPENDICES

Appendix 1: Tests administered

Appendix 2: School's CLIL programme development

Appendix 3: Rating scale used for assessment of the writing task
(adapted from Friedl/Auer 2007)

APPENDIX 1: Tests administered

NOM:...
DATA:....................
CURS:...
CLASSE:..................

A) MULTIPLE CHOICE

En aquest exercici heu d'escollir l'alternativa que us sembli correcta. Teniu tres opcions, però només una és la correcta. Disposeu de <u>15 minuts</u> per resoldre el exercici. No podeu fer cap pregunta. No contesteu si no sabeu la resposta.

Segur que et serà fàcil respondre aquestes preguntes.

1.   a)   Do you can change this 5 €note?
      b)   Have you change for this 5 €note?
      c)   Have you got change for this 5 €note?
      No, I'm sorry, I haven't got any change.

2.   Sally often _____ to do her homework.
      a)   forgetting
      b)   is forgetting
      c)   forgets

3.   How old is your sister?
      a)   She has got ten years old.
      b)   She is ten years old.
      c)   She is ten years.

4.   That green cup on the table is _____.
      a)   my
      b)   mine
      c)   you

5.   ___ Monday, I must go to school.
      a)   At
      b)   In
      c)   On

6.   Sarah plays tennis _____ the weekend.

     a)   in
     b)   on
     c)   at

7. What are you wearing?
     a)   I wearing green socks
     b)   I'm wearing green socks
     c)   Wearing green socks

8. There are _____ children in the park.
     a)   a lot of
     b)   much
     c)   so more

9. My car is _____ than yours.
     a)   the biggest
     b)   bigger
     c)   very big

10.   a)   Mary don't usually speak with people.
     b)   Mary do no usually speak with people.
     c)   Mary doesn't usually speak with people.

Intenta-ho ara amb aquestes.

11.   Did you have lunch at home yesterday?
     a)   No, I haven't.
     b)   No, I didn't.
     c)   No, I hadn't.

12. I visited London last year.
     a)   How long have you stayed there?
     b)   How long did you stay there?
     c)   How long you stay there?

13.   a)   What were you see in London last year?
     b)   What have you seen in London last year?
     c)   What did you see in London last year?

14. What is your father doing now?
     a)   He has just have lunch.
     b)   He is having lunch.

c)   He has lunch.

15.   Where do you live?
a)   At, 45 Green Street, London.
b)   In 45, Green Street, London.
c)   From 45, Green Street, London.

16.   My mother usually cooks dinner for us, but she's in hospital, so
a)   my father does it this week.
b)   has done it this week.
c)   is doing it this week.

17.   I really love her.
a)   She is the intelligentest person I know.
b)   She is the most intelligent person I know.
c)   She is the more intelligent person I know.

18.   I think this is _____ film I have ever seen.
a)   the worse
b)   the baddest
c)   the worst

19.   My mother is _____ than my father.
a)   happy
b)   more happy
c)   happier

20.   How _____ people are there in class?
a)   much
b)   more
c)   many

Aquestes són una mica més difícils però segur que en pots respondre alguna

21. Oh, dear! We don't have _____ money.
     a)   many
     b)   much
     c)   too many

22. There wasn't _____ in the building when the fire started.
     a)   nobody
     b)   somebody
     c)   anybody

23. a)   You mustn't smoke in class!
     b)   You don't smoke in class!
     c)   You can not smoke in class!

24. Why are you putting on your coat?
     a)   Because I took the dog for a walk.
     b)   Because I'll take the dog for a walk.
     c)   Because I'm going to take the dog for a walk.

25. Please, do not disturb me.
     a)   I am trying to study.
     b)   I try to study.
     c)   I will try study.

26. I   _____ to London.
     a)   never went
     b)   have never been
     c)   didn't ever go

27. a)   My friend John has shown me his car the day before
          yesterday.
     b)   My friend John showed me his car the day before yesterday.
     c)   My friend John shown me his car the day before yesterday.

28. a)   I think it is going to rain tomorrow.
     b)   I think it is raining tomorrow.
     c)   I think it rained tomorrow.

29. She learns everything very _____
     a)   easy.
     b)   easily.

      c)   easier.

30.   They invited us _____ dinner with them.
      a)   having
      b)   to have
      c)   have

B) CORRECT/ INCORRECT

En aquest exercici heu de decidir si les frases en anglès son correctes o no. Marqueu la vostra resposta amb una creu a la columna adient. <u>Teniu 10 minuts</u> per resoldre l' exercici. No podeu fer cap pregunta al professor.

	Correct	Incorrect
1. Brighton is on the coast of England		
2. How much it costs?		
9. She can't to ski.		
4. She's fourteen years.		
10. I am taller than my brother.		
6. Betty's father has got green eyes?		
7. The landscapes in Scotland are wonderfuls.		
8. Do they eat in the classroom?		
12. The cat is sleep under the chair.		
14. She play the piano very well.		
11. I am very hungry right now.		
15. My parents go to the theatre every Saturday		
17. Did you understood what she said?		
5. My uncle born in Liverpool		

13. I don't agree with you.		
16. I have always liked Barcelona.		
3. He is such a rich man!		
18. When was the first time you see her?		
19. We don't have many time left.		
20. I am going to write a novel some day.		

C) COMPOSITION

Mireu aquest dibuix amb molta cura i escriviu petits textos segons les intruccions que us donem. Teniu <u>20 minuts</u> per escriure allò que pugueu.

Imagina el diàleg entre els policies, el nen i la mare:

_____
_____
_____
_____
_____
_____
_____
_____
_____
_____
_____
_____
_____
_____
_____
_____
_____
_____
_____
_____
_____
_____
_____
_____

Per què ha passat això que has imaginat en el diàleg?

_____
_____
_____
_____
_____
_____
_____
_____
_____
_____
_____
_____

Com creus que acabarà aquesta situació?
_____

_____
_____
_____
_____
_____
_____
_____
_____
_____
_____
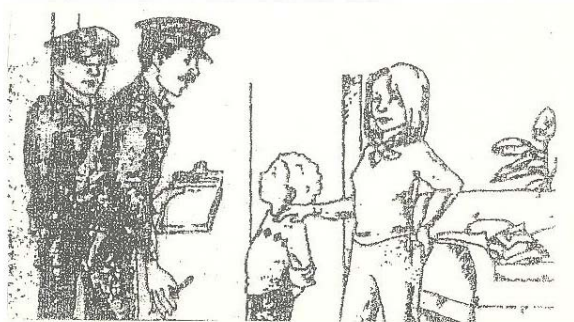
D) CLOZE

Ara us proposem una activitat de triar la paraula correcta per a cada espai. Us aconsellem que abans de començar feu una lectura ràpida de tot el text per saber més o menys de què va. Cal que marqueu l' opció vàlida per a cada forat en el full de respostes. Teniu 15 minuts per fer aquesta activitat. Molta sort!!

Natural disasters such as volcanic eruptions, fires, floods and avalanches happen every year, somewhere in the world. But (**1**) is another, and perhaps more dangerous, (**2**) disaster, which we do not hear about very (**3**). A tsunami is (**4**) huge wave that can cause terrible damage (**5**) destruction.

Tsunami is a Japanese word that (**6**) 'harbour wave'. But why do tsunamis (**7**)? Tsunamis are usually caused by earthquakes at the bottom of the sea. At first, the (**8**) in the sea is quite small, but it moves very (**9**). When the wave gets close to the coast, the ocean floor makes it grow enormously. By the time it reaches the (**10**) it has become huge. Some tsunamis can be 30 metres (**11**). These giant waves can hit Japan, Indonesia, Central (**12**) and South America.

In 1960, there was an earthquake that measured 8.6 on the Richter (**13**) at the (**14**) of the Pacific Ocean. It started a giant wave that (**15**) the coast of Chile and Peru. (**16**) result was that between 1,500 and 2,000 people were killed, and whole villages were (**17**). The wave then travelled (**18**) 15 hours to Hawaii, to the town of Hilo, and another 200-300 people lost their (**19**). It was only one wave, but it caused $500 million of damage.

Tsunamis do not happen very often, but Hawaii now has a 'tsunami watch' station that (**20**) for the next one to come. The station was opened to warn people and give them time to protect themselves against the killer waves.

1)	a. there	b. they	c. he	d. it
2)	a. nature	b. natural	c. naturally	d. naturally
3)	a. days	b. times	c. often	d. sometimes
4)	a. an	b. a	c. ones	d. another
5)	a. of	b. and	c. the	d. because
6)	a. mean	b. signifies	c. means	d. tells
7)	a. happens	b. happen	c. exists	d. go

8)	a. wave	b. water	c. animals	d. one
9)	a. quick	b. speedy	c. rapid	d. quickly
10)	a. sea	b. cost	c. coastal	d. coast
11)	a. high	b. height	c. tall	d. of high
12)	a. america	b. America	c. American	d. USA
13)	a. scale	b. ladder	c. scalate	d. speed
14)	a. under	b. bottom	c. down	d. inferior
15)	a. attacked	b. went	c. hit	d. arrived
16)	a. A	b. The	c. After	d. One
17)	a. destroy	b. have destroyed	c. destroyed	d. was destroyed
18)	a. for	b. since	c. after	d. ago
19)	a. lifes	b. lives	c. livings	d. live
20)	a. are watching	b. have watched	c. is watched	d. watches

READING / CLOZE

Natural disasters such as volcanic eruptions, fires, floods and avalanches happen every year, somewhere in the world. But **there** is another, and perhaps more dangerous, **natural** disaster, which we do not hear about very **often**. A tsunami is **a** huge wave that can cause terrible damage **and** destruction.

Tsunami is a Japanese word that **means** 'harbour wave'. But why do tsunamis **happen**? Tsunamis are usually caused by earthquakes at the bottom of the sea. At first, the **wave** in the sea is quite small, but it moves very **quickly**. When the wave gets close to the coast, the ocean floor makes it grow enormously. By the time it reaches the **coast** it has become huge. Some tsunamis can be 30 metres **high**. These giant waves can hit Japan, Indonesia, Central **America** and South America.

In 1960, there was an earthquake that measured 8.6 on the Richter **scale** at the **bottom** of the Pacific Ocean. It started a giant wave that **hit** the coast of Chile and Peru. **The** result was that between 1,500 and 2,000 people were killed, and whole villages were **destroyed**. The wave then travelled **for** 15 hours to Hawaii, to the town of Hilo, and another 200-300 people lost their **lives**. It was only one wave, but it caused $500 million of damage.

Tsunamis do not happen very often, but Hawaii now has a 'tsunami watch' station that **watches** for the next one to come. The station was

opened to warn people and give them time to protect themselves against the killer waves.

E. DICTATION

Ara farem un dictat. Has d'escoltar amb molta cura i anar escrivint allò que dictarà el professor cada vegada. El professor s'aturarà uns segons entre bloc de paraules i bloc de paraules.

Si no ho entens tot, no et preocupis. Apunta les paraules que entenguis. Si et queda alguna paraula pots deixar un espai i mirar d'omplir-lo la segona vegada que ho llegeixi.

Ànims i endavant

ANTARCTICA

_____
_____
_____
_____
_____
_____
_____
_____
_____
_____
_____
_____
_____

_____
_____
_____
_____
_____
_____
_____
_____
_____
_____

DICTATION

ANTARCTICA

A lot of explorers have visited the South Pole, but it's very cold and very dangerous. The British explorer, Scott, wanted to go to the Pole. He arrived there in January 1911, but he was too late. Another explorer from Norway arrived a month earlier. Scott and all his men died in Antarctica.

The continent is also very popular with scientists. The first winter camp opened in 1899 and today some scientists live and work in Antarctica all year round. In the last 30 years, people have started going to Antarctica for holidays. Now there are about 10,000 visitors every year. Nearly half of the visitors are American, but people come from Europe and Japan, too.

APPENDIX 2: School's CLIL programme development

Table 23 below displays how the school's new CLIL programme progressively involved all learners as of their 3rd grade, year by year.

Shaded cells indicate collected data and red shells represent the grades and cohorts analysed for the purpose of the present disssertation. These were the participants analysed in order to measure pupils from the control group comparable to learners with similar ages who had taken part in the CLIL programme. Further research can be conducted since data from the following academic years (T3: 2006-07, T4: 2007-08) are also available.

Table 23. Experimental design and data collection times of the CLIL project.

			T1	T2	T3	T4
Cohort	**2002-03**	**2003-2004**	**2004-05**	**2005-06**	**2006-07**	**2007-08**
0 (control group)	6p (11-12) FI:980h CLIL: 0h	1s (12-13) FI:1120h CLIL: 0h	2s (13-14) FI:1260h CLIL: 0h	3s (14-15) FI:1400h CLIL: 0h	4s (15-16) FI:1540h CLIL: 0h	1B (16-17) FI:1680h CLIL: 0h
1 (CLIL starts 02-03)	5p (10-11) FI:840h CLIL: 70h	6p (11-12) FI:980h CLIL:140h	1s (12-13) FI:1120h CLIL:210h	2s (13-14) FI:1260h CLIL:280h	3s (14-15) FI:1400h CLIL:350h	4s (15-16) FI:1540h CLIL:420h
2 (CLIL starts 03-04)	4p (9-10) FI:700h CLIL: 0h	5p (10-11) FI:840h CLIL: 70h	6p (11-12) FI:980h CLIL:140h	1s (12-13) FI:1120h CLIL:210h	2s (13-14) FI:1260h CLIL:280h	3s (14-15) FI:1400h CLIL:350h
3 (CLIL 02-03: 1st pure group)	3p (8-9) FI:560h CLIL: 70h	4p (9-10) FI:700h CLIL:140h	5p (10-11) FI:840h CLIL:210h	6p (11-12) FI:980h CLIL:280h	1s (12-13) FI:1120h CLIL:350h	2s (13-14) FI:1260h CLIL:420h
4 (CLIL starts 03-04)	2p (7-8) FI:420h CLIL: 0h	3p (8-9) FI:560h CLIL: 70h	4p (9-10) FI:700h CLIL:140h	5p (10-11) FI:840h CLIL:210h	6p (11-12) FI:980h CLIL:280h	1s (12-13) FI:1120h CLIL:350h
5 (CLIL starts 04-05)	1p (6-7) FI:280h CLIL: 0h	2p (7-8) FI:420h CLIL: 0h	3p (8-9) FI:560h CLIL: 70h	4p (9-10) FI:700h CLIL:140h	5p (10-11) FI:840h CLIL:210h	6p (11-12) FI:980h CLIL:280h
6 (CLIL starts 05-06)	N5 (5-6) FI:140h CLIL: 0h	1p (6-7) FI:280h CLIL: 0h	2p (7-8) FI:420h CLIL: 0h	3p (8-9) FI:560h CLIL: 70h	4p (9-10) FI:700h CLIL:140h	5p (10-11) FI:840h CLIL:210h
7 (CLIL starts 06-07)	N4: (4-5) FI:0h CLIL: 0h	N5 (5-6) FI:140h CLIL: 0h	1p (6-7) FI:280h CLIL: 0h	2p (7-8) FI:420h CLIL: 0h	3p (8-9) FI:560h CLIL: 70h	4p (9-10) FI:700h CLIL:140h

Source: personal

More specifically, this table shows the type of learning context (FI=Formal instruction of English as a foreign language, CLIL=Content and language integrated learning), the number of hours received at the end of each year of both FI and CLIL, and for each cohort of pupils the school year (N5: Nursery, 3p: 3rd year of primary education –Grade 3, 1s: 1st year of secondary education –Grade 7th, 1B: 1st year of Baccalaureate –Grade 11) and their age. It should be noted that each year of CLIL involves 70 extra hours, to the FI hours.

APPENDIX 3: Rating scale used for assessment of the writing task (adapted from Friedl/Auer 2007)

Task fulfilment:content and relevance; text format, length and register

5   Task fully achieved, content entirely relevant; appropriate format, length and register
4   Task almost fully achieved, content mostly relevant; mostly appropriate format, length and register
3   Task adequately achieved, some gaps or redundant information, acceptable format, length and register
2   Task achieved only in a limited sense, frequent gaps or redundant information, often inadequate format, length and register
1   Task poorly achieved, major gaps or pointless repetition; inadequate format, length and register
0   Not enough to evaluate

Organisation: Structure, paragraphing, cohesion and coherence, editing and punctuation

5   Clear overall structure, meaningful paragraphing; very good use of connectives, no editing mistakes, conventions of punctuation observed
4   Overall structure mostly clear, good paragraphing, good use of connectives, hardly any editing mistakes, conventions of punctuation mostly observed
3   Adequately structured, paragraphing misleading at times, adequate use of connectives; some editing and punctuating errors
2   Limited overall structuring, frequent mistakes in paragraphing, limited use of connectives; frequent editing and punctuation errors
1   Poor overall structuring, no meaningful paragraphing, poor use of connectives; numerous editing and punctuation errors
0   Not enough to evaluate

Grammar: Accuracy/ errors, variety of structures, readiness to use complex structures

5   Accurate use of grammar and structures, hardly any errors of agreement, tense, word order, articles, pronouns, etc.; meaning

clear, great variety of structures, frequent use of complex structures

4   Mostly accurate use of grammar and structures, few errors of agreement etc.; meaning mostly clear; good variety of structures, readiness to use complex structures

3   Adequate use of grammar and structures; some errors of agreement etc.; meaning sometimes not clear; adequate variety of structures; some readiness to use complex structures

2   Limited use of grammar and structures; frequent errors of agreement etc.; meaning often not clear; limited variety of structures; limited readiness to use complex structures

1   Poor use of grammar and structures; numerous errors of agreement etc.; meaning very often not clear; poor variety of structures

0   Not enough to evaluate

Vocabulary: Range and choice of words, accuracy, spelling, comprehensibility

5   Wide range of vocabulary; very good choice of words; accurate form and usage; hardly any spelling mistakes; meaning clear.

4   Good range of vocabulary; good choice of words; mostly accurate form and usage, few spelling mistakes; meaning mostly clear.

3   Adequate range of vocabulary and choice of words; some repetitions; some errors of form and usage; some spelling mistakes; meaning sometimes not clear; some translation from mother tongue

2   Limited range of vocabulary and choice of words; frequent repetitions; frequent errors of form and usage; frequent spelling mistakes; meaning often not clear; frequent translation from mother tongue

1   Poor range of vocabulary and choice of words; highly repetitive; numerous errors of form and usage; numerous spelling mistakes; meaning very often not clear; mainly translation from mother tongue.

0   Not enough to evaluate