Learning English Grammar in Content and Language Integrated Learning: Comparing the Grammatical Proficiency of CLIL Students and Students Receiving Mainstream EFL Instruction

Pro Gradu Thesis
Department of Modern Languages
University of Helsinki
20.10.2013

Laura Korpela
Contents

1. Introduction ......................................................................................................................... 1
   1.1. The Objectives of The Present Study ............................................................................. 3
   1.2. Chapter Overview .......................................................................................................... 4
2. Bilingualism .......................................................................................................................... 6
3. Different Forms of CLIL and the Terminology Used in This Study ......................... 8
   3.1. The Terminology of the Present Study .......................................................................... 9
   3.2. Immersion Education: the Original Form of CLIL ................................................... 10
   3.3. CLIL: General Outlines and Key Dimensions ............................................................ 11
   3.4. Immersion and CLIL: The Main Differences ................................................................. 13
   3.5. CLIL in Finland ............................................................................................................. 13
   3.6. CLIL in Töölö Primary School: Bilingual Classes ...................................................... 15
4. The Acquisition of L2 Grammar in CLIL and Mainstream Foreign Language Education .................................................................................................................. 17
   4.1. Acquisition, Learning and Different Types of L2 Knowledge ................................. 19
       4.1.1. The Advantages of Explicit L2 Grammar Instruction ............................................... 20
       4.1.2. Form-Focused Instruction in the CLIL Setting ......................................................... 21
   4.2. Developmental Patterns of L2 Grammar and the Interlanguage Theory ... ............... 23
   4.3. Learner Errors and Error Categories .......................................................................... 25
   4.4. Töölö Primary School Curriculum and the Contents and Goals of Grammar Instruction in CLIL and Mainstream EFL Programme ..................................................... 28
   4.5. Research Findings on the Learning Outcomes of CLIL .............................................. 31
5. Research Design ................................................................................................................... 34
   5.1. Research Questions ....................................................................................................... 35
   5.2. Data Collection ............................................................................................................. 35
   5.3. The Informants ............................................................................................................. 38
   5.4. Data Analysis ............................................................................................................... 40
6. Results and Analysis .......................................................................................................... 42
   6.1. The Implicit Knowledge of Grammar: Grammaticality Judgments ...................... 42
   6.2. Knowledge of Different Grammatical Categories ...................................................... 46
   6.3. CLIL and Mainstream Students’ Productive Skills Compared .............................. 47
       6.3.1. Translation of Questions from Finnish into English ................................................ 48
6.3.2. The Types of Errors Produced: Translation of Questions .................. 50
6.3.3. The Types of Errors Produced: Written Composition ........................ 56

7. Reliability and Validity .............................................................................. 62

7.1. The Quantitative Content .................................................................... 62
7.2. The Qualitative Content: Descriptive Error Analysis ......................... 63

8. Discussion and Conclusion ..................................................................... 66

8.1. The Main Findings .............................................................................. 66
8.2. Implications for Teaching and Further Research ................................. 72

References .................................................................................................... 73

Appendix 1: The Grammar Test ................................................................. 84
Appendix 2: The errors produced in the written composition ................... 92
Appendix 3: Tables ....................................................................................... 96

Tables and Figures

Figure 1: The sources and types of learner errors in the L2. ....................... 26
Table 1: The grammatical content covered in formal EFL lessons by the 5th grade in CLIL and by the 6th grade in mainstream classes (Töölö Primary School Curriculum 2005). ................................................................. 30
Table 2: The sentences in the grammaticality judgment part and the grammatical items measured by each sentence. ................................................................. 36
Table 3: The informants: class, gender and year of birth .............................. 38
Table 4: English mark in the previous certificate ......................................... 39
Table 5a: Implicit knowledge of English grammar: the mean totals from the grammaticality judgments ........................................................................... 42
Table 5b: Implicit knowledge of English grammar: the distribution of total scores. ........................................................................................................ 43
Table 6: Implicit knowledge of the English grammar: the distribution of incorrect answers ..................................................................................... 44
Table 7a: The distribution of mean totals from the cloze passage. ............... 46
Table 7b: The distribution of mean totals in different grammatical categories. .. 47
Table 8: The mean totals from the translation of questions .......................... 49
Figure 2: The modified categories used in the error analysis ...................... 51
Table 9a: The errors produced by the mainstream students in translation .... 53
Table 9b: The errors produced by the CLIL students in translation .............. 54
Figure 3: The translation of questions: the number of errors produced per category. ................................................................. 54
Figure 4: The translation of questions: the frequency of error types within the group................................................................................................................................. 56
Figure 5: Written composition: the number of errors produced per category..... 58
Figure 6: Written composition: the frequency of error types within the group... 58
Table 10: Correlations between the students’ previous English marks and the test results................................................................. 96
1. Introduction

Plurilingualism, intercultural knowledge and intercultural communication skills are the major competences for future citizens of plurilingual Europe and the globalized world. The key issue is fluent and functional foreign language proficiency. In order to function as mobile European citizens, young people need to develop appropriate linguistic knowledge and skills to operate successfully in a foreign language across diverse settings.

The political and educational interest to improve the quality of foreign language education, to promote plurilingualism and to better respond to the demands of globalizing world has led to the implementation of content and language integrated learning (CLIL) all around Europe and the world in the past decades. CLIL refers to “a number of different models of instruction involving two languages as a medium and/or target of instruction” (Järvinen 1999: 15). The foreign language is both the medium and the target of learning: CLIL students learn a foreign language by using it to study other academic subjects. CLIL has been found to be effective in all sectors of education from primary through to adult and higher education.

In Finland, CLIL began with Swedish immersion programmes in the 1980s and English CLIL programmes in the 1990s. The 2000s and 2010s saw the rise of CLIL programmes in other major world languages like Spanish, Russian and Chinese (cf. Mustaparta & Tella 1999; Nikula & Marsh 1997; Vuorinen 2009). In 2003, CLIL was identified as a priority area for development in the European Commissions’ Action Plan for Promoting Language Learning and Linguistic Diversity (European Commission 2003: 8): “Content and Language Integrated Learning (CLIL), in which pupils learn a subject through the medium of a foreign language, has a major contribution to make to the Union’s language learning goals. It can provide effective opportunities for pupils to use their new language skills now, rather than learn them now for use later. It opens doors on languages for a broader range of learners, nurturing self-confidence in young learners and those who have not responded well to formal language instruction in general education. It provides exposure to the language without requiring extra time in the curriculum”.
According to a wide base of research, the outcomes of CLIL are mainly positive. CLIL does not have a negative effect on first language development, content learning or cognitive development, and provides an effective and authentic setting for foreign language learning: CLIL students generally acquire a wider vocabulary range, higher fluency, better comprehension and production skills, and better motivation and confidence to speak the foreign language than students receiving formal, ‘mainstream’ foreign language instruction 1 (e.g. Cummins 1995; Genesee 1987; Jäppinen 2002, 2003, 2005; Järvinen 1999; Nikula & Marsh 1997: 83—101, 114—115, 1999c: 72, 1999d: 83; Seikkula-Leino 2002; Swain 1996; Swain & Lapkin 1982; Vuorinen 2009). CLIL students’ implicit knowledge of the accurate forms of the target language is as good as or even superior to that of the mainstream students 2 (Järvinen 1999), and CLIL students also have the courage to use different forms of the language even before they have been formally taught (Nikula & Marsh 1997: 85).

It has been found, however, that CLIL students do not necessarily learn to master the target language grammar. Research has shown that the productive skills of CLIL students do not match up to their receptive skills, and that students attending mainstream foreign language lessons produce more accurate target language than CLIL students (Genesee 1987; Kokkila 1996: 105—107; Nikula & Marsh 1997: 90—91; Stanutz 1992; Swain 1996: 94).

Deficient mastery of the grammar can, first of all, make it more difficult for the CLIL students to express content mastery or deal with complex subject matter (cf. Cummins & Swain 1986: 37—56; Doughty & Williams 1998: 2; Nikula & Marsh 1999c: 75, 1999d: 80—81; Swain 1996: 92—93; Swain & Lapkin 1982: 68). Being a CLIL teacher myself, I have noticed that the discrepancy between content mastery and grammatical proficiency can be particularly disadvantageous and frustrating in an exam situation: the student knows the

---

1 In this study, the terms ‘mainstream foreign language (L2) education’, ‘mainstream foreign language (L2) lessons’ and ‘mainstream foreign language (L2) instruction’ all refer to the formal foreign language instruction that takes place on isolated lessons. The terms ‘formal’, ‘instructed’ and ‘mainstream’ will be used interchangeably depending on the context.

2 The students who attend the formal and instructed, mainstream foreign language lessons.
answer, but does not know how to formulate it in English. The student might even know the right vocabulary, but struggles with verb tenses, prepositions and word order. Sometimes even a minor preposition or wrong word order might change the meaning radically, when the teacher has difficulty deciding whether the student lacks the content knowledge or the grammatical proficiency to formulate a correct answer. The lack of precision in scientific argumentation and the inability to use the signalling systems for nuances can have a great impact on assessment (Nikula & Marsh 1999c: 75).

It should also be noted that in all sectors and stages of schooling, foreign language proficiency is measured by exams that place great emphasis on grammatical proficiency, one example being the Finnish matriculation examination. In the academic world, 'knowing a language' is more or less the same as performing well in a test. Insufficient mastery of grammar may, therefore, have a negative effect on academic performance, both in the CLIL programme and at later stages of education. Furthermore, the errors may fossilize and be difficult to correct later on, so the effects of insufficient acquisition of the foreign language grammar may be far-flung.

1.1. The Objectives of The Present Study

Earlier research on both immersion and CLIL has shown that students who attend formal foreign language instruction outperform CLIL students in grammatical accuracy when they have to produce the target language themselves. There are also contradictory findings: in Järvinen's (1999) study, CLIL students produced more complex, accurate and versatile sentences than mainstream students. However, many of these studies have been conducted in CLIL and immersion settings that involve little or no explicit foreign language instruction. Recently there have been a number of promising findings on the outcomes of integrating explicit grammar instruction or form-focused communicative tasks to various content-based and meaning-focused approaches (DeKeyser 2003; Doughty & Varela 1998; Ellis 2002; Lyster 2004; Norris and Ortega 2000; Sheen 2005; Spada & Lightbown 2008). As a consequence, instructed foreign language lessons have been added to many CLIL programmes in Finland and abroad. Despite the supportive evidence, more research is needed.
on the foreign language learning outcomes in these settings, especially in the Finnish educational context. This study joins this discourse and seeks to re-evaluate the earlier findings.

The purpose of this study is to evaluate the learning of the English grammar in a CLIL programme in Töölö Primary School, Helsinki, that combines content-based, implicit foreign language acquisition on content lessons with explicit foreign language instruction on formal English as an A1 lessons. The research builds on a comparison of the grammatical proficiency of English between fifth grade CLIL students and sixth grade mainstream students who receive only formal English instruction. The main aim is to analyse, describe and interpret what differences there are in the students’ grammatical proficiency, if any, and what types of errors the students produce in written English, both in translation and a written composition.

The data will be collected by means of an English language test involving grammaticality judgments, a cloze passage (‘fill in the blanks’) exercise involving various structures, translation of questions from Finnish into English, and a written composition. The test will be administered to fifth-grade CLIL students and sixth-grade mainstream students in Töölö Primary School. The data will be analysed mainly quantitatively, drawing on statistical methods and descriptive errors analysis.

This is a case study, so the results can be generalized only if similar conditions apply. Nevertheless, this study can provide valuable information for CLIL teachers, including myself, about the nature of language learning in CLIL, about the possible challenges in achieving the linguistic learning goals of CLIL, and about the grounds for improving the teaching practices of CLIL when it comes to teaching and learning the foreign language grammar.

1.2. Chapter Overview

Chapters 2 through to 5 present the theoretical framework of this study. Chapter 2 outlines the different definitions of bilingualism and chapter 3 presents the various forms of CLIL as well as the terminology used in this study. Chapter 3
also describes the CLIL approach used in Töölö Primary School. Chapter 4 deals with the acquisition, learning and teaching of foreign language grammar, the types of knowledge involved in the mastery of a language, different types of learner errors, and the contents and goals of grammar instruction in CLIL and mainstream foreign language education in Töölö Primary School. Lastly, chapter 4 gives an overview of the most relevant research findings that this study builds on. Chapters 5 through to 7 concern the research design, results and analysis as well as issues related to reliability and validity. Chapter 8 summarizes the main findings and gives implications for further research.
2. Bilingualism

The goal of CLIL is always some degree of bilingualism (cf. Räsänen 1994: 16). However, there is no simple definition of bilingualism available. There are at least three perspectives from which to look at the concept: the level of proficiency in both languages, the functions achieved by both languages in different contexts, and the bilingual person’s receptive and productive abilities in both languages.

One way to define bilingualism is in terms of proficiency in the two languages. But which level of proficiency should bilinguals have? It is very rare for bilinguals to achieve ambilingualism, an equal level of proficiency in both languages. Even balanced bilingualism or equilingualism with equal competence in both languages is rare (Baker 2007: 9). In fact, it is more typical for bilingual people not to possess the same level or type of competence in both languages (cf. Li 2008: 8). Therefore, defining bilingualism in terms of proficiency alone is problematic. Accordingly, Butler and Hakuta (2006: 115) define bilinguals as people who are able to communicate in two languages by speaking and writing regardless of the level of proficiency.

Another definition looks at bilingualism from a functionalist perspective, placing emphasis on the contexts where a bilingual person uses the two languages (Baetens Beardsmore 1991). Functional bilingualism can be interpreted in two ways. From the most minimalistic point of view, a person is functionally bilingual if he or she can accomplish a limited set of activities in a second language (e.g. in a work context). From a maximalist point of view, a person is functionally bilingual if he or she can carry out any activity in a given linguistic context on a satisfactory level.

A third aspect draws on the receptive and productive skills in the two languages. Receptive bilinguals understand the spoken and/or written forms of two languages but do not necessarily use both languages themselves, whereas productive bilinguals understand, speak and write both languages (Baetens Beardsmore 1991).
Lambert (1977: 18—19) makes a further distinction between *additive* and *subtractive bilingualism*. In additive bilingualism, the acquisition of a second language does not have a negative effect on first language proficiency, whereas in the latter case, the second language gradually replaces the first language in different contexts. Subtractive bilingualism may take place, for example, in the case of immigrant children if they learn the language of the majority without receiving any educational support for their first language development. However, if the society assigns positive values to both languages, the more likely result is additive bilingualism (Lambert 1977).

The desired learning goal in CLIL is not native-language proficiency but functional and additive bilingualism (Baetens Beardsmore 1991; Genesee 1988; Lambert 1977: 18—19). This will be discussed in more detail in the next chapter.
3. Different Forms of CLIL and the Terminology Used in This Study

Content and language integrated learning, CLIL, is an umbrella term for “a number of different models of instruction involving two languages as a medium and/or target of instruction” (Järvinen 1999: 15). Teaching fulfils CLIL requirements if at least 25 % of the learning discourse is carried out in a foreign language (Nikula & Marsh 1999b: 23). CLIL may be introduced at any educational level from kindergarten to post-secondary and it provides learning environments that involve considerable exposure to a foreign language as well as opportunities to use the foreign language in meaningful communication (Nikula & Marsh 1999b: 25). The foreign language is, in other words, both the medium and the target of learning.

The ultimate linguistic learning goal of CLIL is typically functional and additive bilingualism (Baetens Beardsmore 1991; Genesee 1988; Lambert 1977: 18—19): the students do not need to gain native-like proficiency in the foreign language, but they need to develop the skills to use it across different contexts so that the foreign language does not have a negative effect on first language proficiency (cf. Byram 1997: 11—12; Genesee 1988).

Confusingly, a great variety of terms have been used for CLIL since its first introduction as French immersion in Canada in the 1960s. Alongside CLIL, the terms include immersion, bilingual education, content-based instruction, content-area teaching, and language- and content-enhanced teaching, to name but a few. During the past decades, CLIL, bilingual education and immersion have become well-established and the most commonly encountered terms in international discussion. Bilingual education and CLIL have been often used as synonyms, while immersion has been regarded as a special, original form of CLIL with its own methodological principles. The term ‘bilingual education’, however, has often been considered confusing and misleading as ‘bilingual’ can simultaneously refer to either methodological principles, the students’ linguistic repertoire at the start of the programme, or the linguistic goal of the programme. To clarify the terminology issue, nowadays CLIL is the recommended umbrella term for all
language and content integrated learning contexts in Europe (cf. Järvinen 1999: 16; Nikula & Marsh 1999b: 21). Immersion is, however, where it all started, and thus it will be briefly presented in section 3.2.

3.1. The Terminology of the Present Study

In this study, CLIL is regarded as the umbrella term for various types of educational models that involve a foreign language as a medium of content instruction. CLIL is also used to describe the educational model followed in Töölö Primary School where this study was conducted. The adjectives ‘content-based’ and ‘meaning-focused’ are also used to refer to CLIL lessons and the type and focus of learning activities taking place on these lessons. The students attending the CLIL programme are called ‘CLIL students’, whereas the students attending the formal, instructed mainstream EFL (‘English as a Foreign Language’) lessons are called ‘mainstream students’. Furthermore, the terms ‘mainstream foreign language (L2) education’, ‘mainstream foreign language (L2) lessons’ and ‘mainstream foreign language (L2) instruction’ all refer to the formal, explicit foreign language instruction that takes place on isolated lessons. The terms ‘formal’, ‘instructed’ and ‘mainstream’ may be used interchangeably when referring to these isolated lessons where the L2 is explicitly taught and learned as opposed to being implicitly acquired in content-based lessons.

The terms L1 (mother tongue, first language) and L2 (a learned or acquired second or foreign language) are also used to describe the students’ linguistic repertoire. The term ‘grammatical proficiency’ refers to the overall grammatical knowledge and skills that the students possess. This proficiency includes different types of knowledge, such as implicit and explicit knowledge, procedural and declarative knowledge and receptive and productive skills, both written and spoken. This study is concerned with grammatical proficiency in written L2 production. These concepts will be described in more detail in chapter 4.

In the following sections, I will briefly describe the general methodological outlines and goals of immersion and CLIL as well as present the CLIL programme followed in Töölö Primary School. I will start with immersion as it is the original form of CLIL, and then carry on with CLIL.
3.2. Immersion Education: the Original Form of CLIL

Immersion education is the first and so far the most researched form of CLIL. The first immersion programme was established in Quebec, Canada, in the 1960s, when a few English-speaking middle class parents demanded the school district administrators set up a French immersion kindergarten class (Baker 2007: 245; Swain & Johnson 1997). Today, immersion education has spread around the world.

Immersion is an umbrella term for various educational models where foreign language is used in content instruction for at least 50% of the time. The immersion approaches differ in two respects: the age at which the immersion is started and the amount of time spent in immersion (Baker 2007: 245). As for age, the immersion may start either at kindergarten or early school years (early immersion), at nine to ten years old (delayed or middle immersion), or at secondary level (late immersion) (Baker 2007: 245). As for time, while partial immersion provides around 50% immersion throughout the school years, in total immersion the immersion language is used exclusively for the first two or three years of schooling, gradually diminishing to 80% immersion per week for the next four years, finishing with approximately 50% immersion (Baker 2007: 245).

The first Canadian programme followed the early total immersion approach, which has subsequently become the most popular bilingual education model in Canada and around the world. This is mostly because of the substantial amount of research evidence about its the successful outcomes and its great adaptability to different societal, cultural, political and educational environments. The Canadian immersion model has been used in countries like Australia, the Basque Country, Catalonia, Ireland, Wales, Singapore, and Finland, where the first immersion programmes were launched in the 1980s with Swedish as the immersion language (Baker 2007: 245; Cummins & Swain 1986; Järvinen 1999: 16) (see section 3.5. for more on CLIL in Finland).

The immersion students have a common first language and the same level of proficiency in the target language at the beginning of immersion. The goal of immersion is additive bilingualism: becoming bilingual and bicultural without
negative effects on first language acquisition or proficiency (Genesee 1988; Järvinen 1999: 14; Lambert 1977). The development of the students’ first language is guaranteed by the dominant position of the language in the surrounding society (Laurén 1991b). The immersion teachers are usually bilingual so that they understand the students’ first language but only speak the immersion language to them. The students are encouraged but not forced to use the immersion language: the idea is that they will start using it once they feel confident enough.

3.3. **CLIL: General Outlines and Key Dimensions**

According to the European CLIL Compendium, CLIL is a generic term that refers to “a dual-focused educational context in which an additional language, thus not usually the first language of the learners involved, is used as a medium in the teaching and learning of non-language content” for at least 25% of the time (Marsh & Hartiala 2001: 17). Like immersion, CLIL is a generic term that allows for various educational models, methodological approaches and desired goals (Marsh 1999: 50). The educational models differ, for example, in terms of the starting age of CLIL, the amount of exposure to the foreign language, the required linguistic background from the learners, and the linguistic environment of the classrooms (monolingual, bilingual or multilingual). CLIL is taking place and has been found to be effective in all sectors of education from primary through to adult and higher education. Like in immersion, the goal of CLIL is additive and functional bilingualism.

The CLIL teachers are either bilingual or native speakers of the target language and depending on the CLIL model the school is following, the learners either have some knowledge of the target language or start from the very basics. All in all, according to the description of CLIL by the European Commission (2013), CLIL refers to a range of diverse methodologies that are adaptable to different learning environments and purposes. The key issue is that the learner is gaining new knowledge about the non-language subject while encountering, using and learning the foreign language. The methodologies and approaches used are usually linked to the subject area with the content leading the activities. CLIL
does not require extra teaching hours and it complements other subjects rather than competes with them (European Commission 2013).

**Key Dimensions**

The European CLIL Compendium (Marsh & Hartiala 2001) has presented the various reasons for implementing CLIL under five key dimensions involving culture, environment, language, content and learning. The schools can decide freely which dimensions will be given more emphasis than others and which learning and development outcomes their CLIL programmes focus on. The dimensions of CLIL are (Marsh & Hartiala 2001: 16):

1. **The Culture Dimension:** building intercultural knowledge and understanding, developing intercultural communication skills, learning about specific countries, regions and/or minority groups, introducing the wider cultural context.

2. **The Environment Dimension:** prepare for internationalization, access international certification, enhance school profile.

3. **The Language Dimension:** improve overall target language competence, develop oral communication skills, deepen awareness of both mother tongue and the target language, develop plurilingual interests and attitudes, introduce a target language, allow learners more contact with the target language.

4. **The Content Dimension:** provide opportunities to study content through different perspectives, access subject-specific target language terminology, prepare for future studies or working life.

5. **The Learning Dimension:** complement individual learning strategies, diversify methods and forms of classroom practice, increase learner motivation and confidence in both the language and the subject being taught.
### 3.4. Immersion and CLIL: The Main Differences

Immersion education and CLIL have similar goals and methods, but they differ in three main respects. First, a pre-requisite for successful immersion is a society that enables the learners to become functional bilinguals and in which the first language of the learners has a strong position (Nikula & Marsh 1999b: 26). Basically, this applies to societies with two official or main languages. This is not a pre-requisite for CLIL. Second, in immersion education the learners learn to read and write in the immersion language, whereas in CLIL they learn to read and write in their first language (Malmström 1993: 20—22). Third, in immersion classes approximately 50% of the teaching and learning discourse should take place in the immersion language while in CLIL the minimum requisite is only 25% (Nikula & Marsh 1999a: 52). Moreover, immersion has well-established methodological principles and goals whereas CLIL is an umbrella term for various educational models and goals, including immersion itself (Kangasvieri et al. 2011: 22; Marsh 1999: 50).

These differences considered, the term that best describes the educational model followed by the informants of this study is CLIL: the foreign language content varies between 30—50% per day and the students learn basic literacy skills in Finnish (see section 3.6. for more on CLIL in Töölö Primary School).

### 3.5. CLIL in Finland

In Finland, the CLIL experiments have adopted a multitude of names, including language shower (kielisuihku), immersion (kielikylpy), language class (kieliluokka), foreign language class (vieraskielinen luokka), bilingual class or bilingual teaching (kaksikielinen luokka tai kaksikielinen opetus), language-enriched teaching (kielirikasteinen opetus) or language-oriented teaching (kielipainotteinen opetus). These have slightly different emphasis, practices and goals, but the basic approach in all of these is content and language integrated learning (CLIL, sisältöpainotteinen kieltenopetus): teaching a curricular subject through the medium of a foreign language (e.g. Kangasvieri et al. 2011: 19—22, 55; Nikula & Marsh 1996, 1997; Räsänen 1994: 17; Vuorinen 2009).
The first larger-scale CLIL approach in Finland was Swedish immersion (ruotsin kielikylpyopetus), designed and organized in Vaasa in 1987 by university professor Christer Laurén together with researchers, local teachers, and city and school administration (Vuorinen 2009: 23—24). The approach was adopted from Canada and suited the political and educational grounds of Finland very well: both Canada and Finland have two official languages and knowledge of the two is required in administrative positions. In addition, the students in both countries were required to study the second official language of the country, in which very few succeeded well enough (Vuorinen 2009: 23—24). Thus, there was a need for a different, more effective method to teach the second language (cf. Laurén 1991a, 1991b, 1992; Swain & Lapkin 1982). Swedish immersion based on the Canadian model was the forerunner of many other models of CLIL that operate in Finnish schools today.

The number of Finnish schools that offer CLIL is about 4—8% (Vuorinen 2009: 24). The majority of schools that implement CLIL provide either Swedish immersion or CLIL in English or other world languages like Spanish, Russian or Chinese. The immersion approach (kielikylpy) in Finland only relates to Swedish language immersion programmes and differs from other types of CLIL in Finnish schools, mainly in terms of its political and educational aims as described above (cf. Swain & Johnson 1997). Most CLIL models follow the national curriculum, but some have adopted the International Baccalaureate Preliminary Years and Middle Years or Diploma curricula.

Hartiala (2000: 38) groups the linguistic goals of CLIL in Finland under three main categories: 1. to build self-confidence among student cohorts and boost interest in foreign language learning, 2. to enhance and activate existing foreign language knowledge, and 3. to substantially improve the learners’ language skills. As for the five key dimensions of CLIL presented in section 3.3., Hartiala (2000: 38) states that Finnish primary schools generally place greater emphasis on linguistic development (the language dimension) and intercultural understanding (the culture dimension) whereas secondary and upper secondary levels focus more on the content dimension.
3.6. CLIL in Töölö Primary School: Bilingual Classes

The current research data was collected in the bilingual classes in Töölö Primary School. The school has implemented CLIL in English since 1991 in the form of bilingual classes (*kaksikielinen opetus* or *kaksikieliset luokat*), meaning that the academic subjects are taught both in Finnish and English. The amount of English instruction should be 50% throughout grades one to six. In reality, however, the amount of English instruction rises grade by grade, starting with approximately 30% English and 70% Finnish instruction on the first grade and finishing with a 50% English and 50% Finnish situation on the sixth grade. Teachers have to ensure that English instruction does not hamper content learning, and since the students do not speak English at all at the start of the programme, it is often impossible to teach half of the content in English even though that is the general principle. However, to maximize the amount of input, English is used alongside Finnish in all teacher-student interaction, so the daily English input is still likely to reach 50% even on the first and second grade.

The students do not need to speak or write English before commencing the programme, but their *language aptitude* is tested before they are admitted to the programme to ensure that they are able to undertake the programme. Students are, thus, selected on the basis of their score in the language aptitude test. Involving both an underlying language learning capacity and the capacity to handle decontextualized language, language aptitude has been found to be one of the best predictors of L2 learning (cf. Ellis 1995: 522). Studying in both Finnish and English typically means much extra work and challenges for the student as well as requires ample support from parents. Practically, all CLIL students speak or at least understand Finnish at the start of the programme even if their first language is something different. They also learn to read and write in Finnish on the first grade. The first-graders also practice reading, writing and spelling in English, but on a very basic level suitable for second-language learners. Many students continue in English-speaking or CLIL classes in the secondary level.

The bilingual classes follow the school’s general curriculum and the general learning goals and principles are the same as for the mainstream Finnish-speaking classes. Student evaluation also follows the same general principles. It
is stated in the school curriculum that the language of instruction should not affect the learning outcomes or evaluation.

The CLIL teachers in Töölö are native speakers of Finnish and have a native-like proficiency in English. The teachers have much freedom in deciding which content to teach in Finnish, in English, or in both languages, as long as approximately half of the instruction is carried out in English. Generally, some contents are more suitable to learn in Finnish, such as Finnish history and the geography of Finland and the Nordic Countries. Mathematics is taught mainly in English, and only the most essential mathematical concepts are learned in both languages. All CLIL students attend Finnish as a mother tongue lessons and L2 speakers of Finnish take Finnish as a Second Language lessons once or twice a week.

English has an A1 language status in bilingual classes, meaning that besides content instruction in English, the students attend formal EFL lessons taught solely in the target language by a native speaker. At the time of the data collection, there was one native English-speaking teacher for the bilingual classes, teaching A1 English for all bilingual classes as well as physical education for girls and boys on grades 1—4 and boys on grades 5—6.

The general linguistic goal of the CLIL programme is to provide the student with a functional command of the English language and the ability to use it properly and concisely to convey meaning. This includes having knowledge about the language, listening attentively, talking to the point, reading with understanding, and writing fluently with accurate spelling and punctuation. The content of language learning is closely connected to the language skills needed in other academic subjects (e.g. mathematics, history and science). The students should learn the most essential concepts and contents in different academic subjects in both English and Finnish, so that they may continue their studies in either language.

(Adapted from Töölö Primary School Curriculum, 2005 and 2012).

---

3 ‘EFL’ stands for ‘English as a Foreign Language’. In this study, ‘EFL lessons’ or ‘EFL education’ refers to the formal, instructed English as an A1 language lessons attended by the students.
4. The Acquisition of L2 Grammar in CLIL and Mainstream Foreign Language Education

Foreign language education has adopted an increasingly communicative and intercultural emphasis during the past decades. Modern foreign language instruction draws primarily on the principles of communicative language teaching (CLT). The goal is to develop the students' communicative competence, which has four components: grammatical, discourse, sociolinguistic and strategic competence (Canale & Swain 1980; Canale 1983). Grammatical competence includes the knowledge of e.g. vocabulary, word formation, syntax, spelling, and pronunciation. Discourse competence involves knowledge of the conventions, coherence and cohesion of the language. Sociolinguistic competence refers to the knowledge required to use the language appropriately in different social contexts, such as form and function, registers and roles, while strategic competence is the skill to use efficient communication strategies.

Communicative language teaching is based on the view that language is learned primarily by taking part in meaningful interaction and carrying out authentic communicative tasks, defined by Nunan (1989) as a “piece of classroom work which involves learners in comprehending, manipulating, producing or interacting in the target language while their attention is principally focused on meaning rather than the form". Communicative tasks involve, for example, problem solving, group work and collaborative learning.

At the core of the CLT methodology are Krashen’s (1984) five hypotheses about second language acquisition4 (SLA):

1. Acquisition-Learning Hypothesis: language can be either acquired naturally or learned consciously. Acquisition produces implicit and

---

4The term second language acquisition or SLA is an established term in the field of applied linguistics. In this context, the term ‘second language’ typically refers to a language that has an official second language status in the society, while ‘foreign language’ points to a language that has no such status. In this study, the terms ‘second language’ and ‘foreign language’ are used interchangeably to refer to any additional language that the learner acquires or learns besides the first language.
procedural knowledge and takes place when a child learns the L1 or when a L2 is acquired in naturalistic settings. Learning takes place in formal settings, usually through instruction, and results in explicit and declarative knowledge of the language.

2. **Natural Order Hypothesis**: the rules and structures of L2 are learned in a certain order regardless of the order in which they are taught.

3. **Monitor Hypothesis**: conscious learning plays only a limited role in foreign language acquisition and performance. Instead, learning acts as a monitor of the input and an editor of the acquired competence of the language.

4. **Input Hypothesis**: language acquisition and automation takes place by comprehending and processing comprehensible linguistic input that is a little step beyond the learner’s current level, and as a result the learner moves from one proficiency level (i) to another (i+1).

5. **Affective Filter Hypothesis**: anxiety, communication apprehension and frustration act as an affective filter in the learning process and inhibit language learning or acquisition. The learner needs to feel confident and at ease when studying the language and have opportunities to practice the language in a safe atmosphere that does not trigger anxiety.

Krashen’s Input Hypothesis has been criticized for ignoring the learners’ need to actually produce the language in order to acquire it. There is a wide agreement between researchers that simple exposure to the foreign language is not enough for the development of L2 proficiency and grammatical accuracy: the learner also needs opportunities to produce meaningful output (*the output hypothesis*) (Swain 1985; Meriläinen 2008: 29; Nikula & Marsh 1997: 113, 1999d: 80).

Mainstream foreign language instruction and CLIL both stand on the theoretical principles of communicative language teaching. However, they differ in the way the L2 is taught and acquired and in the type of L2 knowledge targeted and produced (cf. Hartiala 2000: 53—54; Järvinen 1999: 29—49; Meriläinen 2008: 27). In the following section I will describe and contrast the main theoretical outlines of second language acquisition, especially from the point of view of learning the L2 grammar in CLIL and mainstream EFL instruction.
4.1. Acquisition, Learning and Different Types of L2 Knowledge

CLIL, as well as its other forms like bilingual education and immersion, is based on Krashen’s (1984) idea of informal *language acquisition* as opposed to formal *language learning*. The former refers to the naturalistic, informal and implicit acquisition of a foreign or second language through substantial exposure, challenging input and meaningful output, while the latter denotes the conscious and rule-based, explicit learning of a foreign or second language (Krashen 1984; see also Ellis 1995: 355—356; Järvinen 1999: 37—46). Formal language learning is thought to take place during *instructed* language lessons and generate *explicit knowledge* of the structures and rules of the language; by contrast, language acquisition takes place in *naturalistic* environments such as CLIL lessons and is believed to produce *implicit knowledge* of the language: the intuitive, hidden and unconscious L2 capacity that manifests itself only in actual performance (Ellis 1995: 355—356; Järvinen 1999: 37—46). Another distinction has been made between *declarative knowledge*, knowing about the language as a grammatical system, and *procedural knowledge*, knowing how to use the language functionally (Ellis 1995: 388; O’Malley & Chamot 1990: 73). In general, explicit and declarative knowledge are the focus of formal i.e. mainstream L2 instruction while naturalistic approaches such as CLIL and immersion emphasize implicit and procedural knowledge.

These naturalistic—instructed, informal—formal, meaning—form, implicit—explicit, procedural—declarative and acquisition—learning distinctions are, by and large, the major differences between L2 learning in CLIL contexts and mainstream L2 lessons. While mainstream students receive formal L2 instruction, explicitly *learn* the L2 rules and develop declarative knowledge, CLIL students mainly *acquire* the L2 through exposure to a large amount of comprehensible input and opportunities for producing output in meaningful communicative interaction (cf. Järvinen 1999: 29; Seikkula-Leino 2002: 38). This acquisition process is, according to Krashen’s Monitor Hypothesis (1984), implicit, natural and unconscious, and is executed by the innate language processor or learning device, the Universal Grammar (Chomsky 1996). The result is improved procedural knowledge of the L2 (cf. Järvinen 1999: 29).
4.1.1. The Advantages of Explicit L2 Grammar Instruction

There has been much disagreement as to whether formal, explicit instruction can facilitate the acquisition of L2 grammatical features and whether formally learned explicit knowledge can transform to implicit knowledge at all. According to the noninterface hypothesis (Krashen 1984), what is explicitly learned cannot become implicit knowledge. This view finds support in neurological evidence that the declarative, explicit knowledge and procedural, implicit knowledge are stored separately in the brain (Paradis 1994). Correspondingly, a vast number of researchers have concluded that explicit grammar instruction may promote metalinguistic knowledge but not the actual competence of the L2 (e.g. Eubank & Beck 1995: 40; Truscott 1998: 119; VanPatten & Cadierno 1993; Wong 2004). A study by Wong (2004), for example, indicated that students who received explicit grammar instruction alone performed no differently in production and comprehension tests than a control group; hence, explicit instruction is unnecessary.

By contrast, the interface hypothesis argues that explicit knowledge can become implicit knowledge if the learners have opportunities for a plenty of communicative practice (DeKeyser 1998; Ellis 2006: 97). The weak interface hypothesis adds that the conversion from explicit to implicit knowledge requires that there is a knowledge gap and the learner notices it; the targeted grammatical feature has to reside in the zone of proximal development (Ellis 1993). Both of these views are in favour of explicit grammar instruction.

Today there is both direct and indirect evidence supporting the benefits of explicit grammar instruction (Ellis 2006: 86). The results of studies conducted in the 1980s and early 1990s (e.g. Doughty 1991; Ellis 1990; Genesee 1987; Lightbown & Spada 1990; Long 1983; Pica 1983; White et al. 1991) suggested that instructed learners progressed more rapidly and achieved higher levels of proficiency than naturalistic learners, such as immersion and CLIL students. Likewise, more recent evidence (e.g. DeKeyser 2003; Doughty & Varela 1998; Ellis 2002; Lyster 2004; Norris and Ortega 2000; Sheen 2005) shows that grammar instruction contributes to both implicit, procedural knowledge and explicit, declarative knowledge of the language. CLIL and immersion teachers
have noticed that with an exclusive focus on meaning with no attention to the form, some language features are never acquired and some non-target forms persist for years (Spada & Lightbown 2008: 184). Accordingly, there have been a number of studies with promising results about integrating explicit grammar instruction or form-focused communicative tasks to content-based and meaning-focused approaches like CLIL and immersion (e.g. Doughty & Varela 1998; Lyster 2004; Sheen 2005; Spada & Lightbown 1993). However, Ellis (2006: 86) points out that some of the evidence is controversial. For example, many studies have measured L2 grammar learning by what Ellis (2006: 86) calls constrained constructed responses (fill in the blanks, sentence joining, etc.), which might favour instructed learners. When measured by means of free constructed responses such as communicative tasks, the evidence in favour of explicit grammar instruction is controversial (Ellis 2006).

4.1.2. Form-Focused Instruction in the CLIL Setting

Albeit controversial, the findings on the advantages of grammar instruction has led to the reintroduction of grammar to communicative L2 lessons by means of form-focused instruction (FFI), an umbrella term for any planned or incidental instructional activities that induce learners to pay attention to linguistic form within the communicative setting (Ellis 2001). Due to the promising results, today many meaning-focused approaches like CLIL and immersion involve form-focused components as well (De Graaff & Housen 2009: 736). This is also the case in Töölö Primary School where the present study was conducted. Hartiala (2000: 40) asserts that formal foreign language learning is actually essential in Finnish CLIL approaches in order to support the CLIL programme. This is because the foreign language has an add-on value in the CLIL context and is not generally learnt for its own sake. In Hartiala’s (2000: 40) view, CLIL should be complementing and not replacing formal L2 instruction.

Depending on the targeted language features, the learners’ age and abilities as well as the teachers’ characteristics and preferences, form-focused instruction in CLIL may take place either on isolated L2 lessons or be integrated in content-based lessons (Spada & Lightbown 2008: 200—201). Nowadays an increasing number of CLIL teachers and researchers emphasize the importance of planning...
lessons that include both content objectives and linguistic objectives (Spada & Lightbown 2008: 184). In a study of immersion classrooms, Allen et al. (1990: 74, quoted in Lyster 2011: 620) reported that interaction that occurs naturally might give little opportunities for CLIL students to produce a full range of target language forms. Therefore, they recommended that CLIL teachers “implement carefully planned and guided communicative practice that will push students towards the production of comprehensible output” (Allen et al. 1990: 76, quoted in Lyster 2011: 620).

To integrate form-focused activities in the CLIL classroom, the so-called noticing and awareness activities can be used. Noticing activities draw the learners’ attention to problematic L2 features through e.g. colour coding and boldfacing in written input or intonation and gestures in oral output (Lyster 2011: 618). They are followed by awareness activities that can be either receptive or productive and include e.g. rule-discovery tasks and comparison of language patterns that require some elaboration of the noticed forms. This elaboration then helps the learners restructure the declarative rules in their interlanguage. According to Lyster (2011: 618), noticing and awareness activities are aimed to strengthen the language learners’ metalinguistic awareness and linguistic sensitivity, which then help them extract linguistic information from content-based input and thus develop their interlanguage through content instruction.

According to Spada and Lightbown (2008: 198—201), learners with poorer metalinguistic skills in their L1 and those who score lower in a language aptitude test may benefit from more explicit and isolated L2 instruction that helps them identify some form—meaning connections, for example. In addition, language features that have low salience, low frequency or low communicative value may be taught and learned more effectively by means of isolated L2 instruction (Spada & Lightbown 2008: 197). However, isolated grammar lessons should never take place without follow-up communicative or content-based activities in either formal L2 lessons or content-based lessons (Ellis 2006: 101; Spada & Lightbown 2008: 201).

The CLIL programme in Töölö Primary School is a combination of explicit L2 instruction on formal English lessons with implicit L2 acquisition on content-
based lessons. Depending on the teacher, some form-focused activities may also take place on the content-based lessons. The main objective of the programme is that the CLIL learners develop both implicit and explicit, or procedural and declarative knowledge of the L2. After all, both types of knowledge are at work when the students comprehend and produce the L2 (cf. De Graaff & Housen 2009: 733).

4.2. Developmental Patterns of L2 Grammar and the Interlanguage Theory

Research on SLA has given evidence that the acquisition of L2 grammatical features follows similar developmental patterns (order and sequence of acquisition) that have been observed in L1 acquisition, so that one form or structure always precedes another (Ellis 1995: 20—21). Krashen’s Natural Order Hypothesis (1984) maintains the same idea: the acquisition of L2 grammar occurs in stages. This idea is based on the interlanguage theory by Selinker (1972). According to this theory, the L2 learning process involves a series of mental grammars that the learners construct, revise and restructure as they gradually become more proficient in the L2. In this process, they form mental representations (interlanguages) of L2 rules that they then test in L2 production. Sometimes this leads to overgeneralization, such as using the past tense marker -ed for irregular verbs or the plural ending -s for irregular plurals like ‘childrens’ (Ellis 1995: 30). Another result can be simplification or the reduction of the L2 system into a simpler form (Ellis 1995: 30).

According to Ellis (1995: 28—29), the learner’s L1 plays a major role in the development of interlanguage and sometimes causes interference or negative transfer, when a L1 rule or structure is erroneously applied to L2. Negative transfer can also manifest itself as avoidance of certain L2 forms, the overuse of certain forms or the simplification of the output. There is also positive transfer when the L1 facilitates L2 acquisition. All in all, transfer is thought to be a natural part of the L2 learning process.
Foreign language learners, neither CLIL nor mainstream students, seldom reach the same level of proficiency as native speakers (Ellis 1995: 353). According to Selinker (1972), this is due to fossilization: non-target forms become fixed in the interlanguage. Fossilized forms are most likely to appear in productive language use, such as speaking or free written production (Ellis 1995: 353). Research findings suggest, however, that fossilization can be prevented by paying attention to the language form in content-based lessons as well (e.g. Doughty & Varela 1998; Lyster 2004; Sheen 2005; Spada & Lightbown 2008).

The interlanguage theory has had a major impact on foreign language teaching practices and views of learner errors from the 1970s onwards: errors are no longer seen as hazardous deviations from the standard but an inevitable and natural part of the learning process (cf. Cook 1991: 7). In the light of the developmental patterns, errors reflect the interlanguage stage the learner has reached in the L2 acquisition continuum.

The CLIL approach is particularly meaning-focused and errors do not receive too much attention in teaching as long as they do not alter the meaning radically. Nikula and Marsh (1997: 91) have suggested that CLIL teachers should not ignore the learners’ errors completely but give corrective feedback or recasts instead. To avoid making CLIL learners apprehensive of errors, teachers should subtly help them notice the errors themselves by reformulating the learner’s erroneous utterance before carrying on. After all, a number of scholars (Cummins & Swain 1986: 37—56; Doughty & Williams 1998: 2; Nikula & Marsh 1999d: 80—81; Swain 1996: 92—93; Swain & Lapkin 1982: 68) have stated that ignoring the erroneous use of language repeatedly may do the CLIL learners a major disservice as it gives them a false idea of their language skills and makes it more difficult for them to express themselves and deal with complex subject matter later on, especially on higher grades. There is also danger that some errors may fossilize or become automatized and be hard to unlearn as a result.

The following section categorizes the types of errors L2 learners mainly produce. This categorization will serve as a backbone for the second part of the data analysis.
4.3. Learner Errors and Error Categories

An error can be defined as “a deviation from the norms of the target language” (Ellis 1995: 51). The point of reference is usually the standard written dialect. Learners make errors in both comprehension (receptive errors) and production (productive errors). The present study is concerned with productive errors in the students’ written English.

Errors are a natural part of the second language learning process (cf. Cook 1991: 7). Besides being productive or receptive, learner errors can be categorized into overt or covert (Ellis 1995: 52). An overt error is a clear deviation from the norm, such as “I runned all the way”, where as covert errors occur on the phrase level and can be deemed errors only from the context: for example, if the learner uses ‘it’ in the phrase “It was stopped” to refer to “wind” (Ellis 1995: 52). Learners can also make pragmatic errors that are structurally correct but not appropriate, such as saying “I want some juice” to ask politely for juice (Ellis 1995: 52).

In the field of SLA, the study of errors has traditionally been carried out by means of descriptive error analysis (cf. Ellis 1995: 58). The error studies have focused both on the types and sources of errors. Out of the many attempts to categorize errors for the purposes of analysis, this study relies on the categorization presented in Figure 1, modified from Ellis (1994: 58) on the basis of Lott (1983) and Richards (1971).
An important distinction has been made between *competence* and *performance* errors, or *errors* and *mistakes*: errors stem from a lack of L2 knowledge and competence whereas mistakes are caused mainly by *processing problems* and occur when learners fail to perform their L2 competence (Ellis 1995: 51, 58). Instead of accessing what they know, they fall back to non-standard alternative rules that are easier to remember and produce. In this study, spelling mistakes and incorrect word choices are also categorized as processing problems if they affect the grammaticality of the sentence. The spelling mistakes can, for example, derive from phonology: the learner knows the spoken form of the word and tries to produce its written form on the basis of pronunciation.

Another subcategory in performance consists of errors linked to *communication strategies*, or *strategic errors*. These occur when a learner uses strategies like paraphrasing or circumlocution to overcome the lack of knowledge (Ellis 1995: 58). The learner might, for example, use a word that closely resembles the correct translation or belongs to the same semantic category, but is unidiomatic in that particular context. However, the learner still manages to deliver the

**Figure 1:** The sources and types of learner errors in the L2.
meaning. This error category is problematic because communication strategies contribute essentially to fluency and are an integral part of the communicative competence. However, in the context of error analysis, they are treated as performance errors.

Competence errors have been divided into transfer from the L1 and intralingual errors (Ellis 1995: 58—59, on the basis of Lott 1983). There are three types of transfer errors:

1a. **Overextension of analogy**: the learner misuses an item because it shares features with an item in the L1, for example when Italian learners use ‘process’ to mean ‘trial’ because Italian ‘processo’ has this meaning.

1b. **Transfer of structure**: the learner uses some L1 feature (phonological, lexical, grammatical, or pragmatic) rather than that of the target language.

1c. **Interlingual errors**: occur when a particular distinction does not exist in the L1, for example Finnish learners saying ‘do a mistake’ instead of ‘make a mistake’ because there is no make/do distinction in Finnish. Pragmatic errors can also derive from interlingual transfer, such as the formerly presented ‘I want some juice’ instead of ‘I would like some juice’.

Furthermore, there are four types of intralingual errors (Ellis 1995: 59, on the basis of Richards 1971):

2a. **Overgeneralization**: occurs when the learner uses a deviant structure on the basis of other structures in the language, for example ‘He can sings’ or ‘Does he sings?’, while only ‘He can sing’, ‘He sings’ and ‘Does he sing?’ are grammatically correct.

2b. **Ignorance of rule restrictions**: a rule is applied to contexts where they do not apply, such as ‘He made me to rest’ on the basis of the regular pattern ‘verb + to (do something)’.

2c. **Incomplete application of rules**: refers to structural failures or omissions of constituents like auxiliary verbs or prepositions, such as asking ‘You like sing?’ instead of ‘Do you like to sing?’, or ‘I am go to school’ instead of ‘I am going to school’.
2d. **False concepts hypothesized**: the learner does not fully understand a distinction in the target language, such as using 'was' to mark past tense in 'Yesterday I was go to the cinema'.

Another and a simpler way to classify errors is to use grammatical or linguistic categories, such as verb forms, adjectives, articles and prepositions (Ellis 1995: 55). This type of classification will also be used in the data analysis when a more detailed analysis is not reasonable.

Error analysis has been deemed problematic because of difficulties in categorization: the categories either overlap or are too vague and broad. Despite these problems, however, error analysis can be advantageous for pedagogical and diagnostic purposes (Ellis 1995). This is why I chose to utilize error analysis in this study.

### 4.4. Töölö Primary School Curriculum and the Contents and Goals of Grammar Instruction in CLIL and Mainstream EFL Programme

In section 4.2. we learned that nowadays the CLIL approaches in Finland typically combine *naturalistic acquisition* and *formal or instructed learning* of the L2, thus focusing on both *implicit* and *explicit knowledge*, or in other words, *procedural* and *declarative knowledge* of the language (De Graaff & Housen 2009: 736). The CLIL programme in Töölö Primary School, likewise, involves both explicit EFL instruction on formal English lessons and implicit EFL acquisition on content-based lessons. The formal EFL lessons are adapted to the students' needs and designed to facilitate the learning of both content and language. Therefore, the English grammar on these lessons is selected on the basis of the discourse needed in content lessons: passive voice, present and past perfect and direct speech are taught as early as the fifth grade in CLIL because these structures are needed to describe a sequence of past events in history and a chain of consequences in science, for instance (cf. Doughty & Varela 1998).
The goals and contents of EFL learning in CLIL and mainstream lessons respectively are defined in the Töölö Primary School Curriculum (2005). On the first CLIL grade, the content of the formal EFL lessons ranges from learning basic vocabulary, spelling, reading, writing, and active listening to nursery rhymes, drama, and storytelling. Playing and games are an essential part of the lessons. Grammar is given more emphasis from the second grade onwards. The mainstream students start A1 English on the third grade.

Table 1 presents the grammatical content to be covered on the formal EFL lessons by the fifth CLIL grade and the sixth mainstream grade. Basically, by fifth grade the CLIL students have formally learned the same grammar topics as sixth grade mainstream students, but at a slower pace. The CLIL students have also studied present and past perfect, passive voice and direct speech to facilitate content learning in subjects like history and science. The CLIL students have, of course, been exposed to a greater amount of English input and thus their implicit knowledge of English might be better than that of mainstream students. Both groups should have, nevertheless, possessed more or less the same explicit knowledge of English grammar at the time when this study was conducted, and the small differences in curriculum contents were taken into account in the research design and analysis.
Table 1: The grammatical content covered in formal EFL lessons by the 5th grade in CLIL and by the 6th grade in mainstream classes (Töölö Primary School Curriculum 2005).

<table>
<thead>
<tr>
<th>Grammatical content covered in formal English lessons by</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Sixth grade in mainstream classes</strong></td>
<td><strong>Fifth grade in CLIL classes</strong></td>
</tr>
<tr>
<td>- correct word order</td>
<td>All of the contents listed on the left, plus the following:</td>
</tr>
<tr>
<td>- definite and indefinite articles: a, an, the, -</td>
<td>- present perfect and past perfect</td>
</tr>
<tr>
<td>- formation of questions and answers</td>
<td>- passive voice</td>
</tr>
<tr>
<td>- present tense: simple and continuous</td>
<td>- direct speech</td>
</tr>
<tr>
<td>- past tense: simple and continuous</td>
<td></td>
</tr>
<tr>
<td>- common irregular verbs (past tense)</td>
<td></td>
</tr>
<tr>
<td>- common auxiliary verbs (e.g. do/does, can)</td>
<td></td>
</tr>
<tr>
<td>- imperative form</td>
<td></td>
</tr>
<tr>
<td>- ‘be going to’ structure</td>
<td></td>
</tr>
<tr>
<td>- introduction to future tense: will, won’t</td>
<td></td>
</tr>
<tr>
<td>- comparison of adjectives</td>
<td></td>
</tr>
<tr>
<td>- common irregular plural nouns (e.g. children, men)</td>
<td></td>
</tr>
<tr>
<td>- personal pronouns</td>
<td></td>
</tr>
<tr>
<td>- possessive pronouns</td>
<td></td>
</tr>
<tr>
<td>- demonstrative pronouns (this, these, that, those)</td>
<td></td>
</tr>
<tr>
<td>- indefinite pronouns (e.g. everybody, anybody, somebody, nobody, everything, something)</td>
<td></td>
</tr>
<tr>
<td>- interrogative pronouns (who, whose, what)</td>
<td></td>
</tr>
<tr>
<td>- ‘there is/are’ structure</td>
<td></td>
</tr>
<tr>
<td>- common prepositions and prepositional structures, including common verb+preposition combinations</td>
<td></td>
</tr>
<tr>
<td>- common adverbs</td>
<td></td>
</tr>
<tr>
<td>- the genitive form (’s and of)</td>
<td></td>
</tr>
</tbody>
</table>
4.5. Research Findings on the Learning Outcomes of CLIL

The learning outcomes of CLIL, including immersion education\(^5\), have been widely researched ever since the introduction of Canadian immersion in the 1960s. The studies have mainly focused on content learning, cognitive development, first language development and foreign language acquisition in CLIL.

As for the content, research findings suggest that while it may slow down content learning at first, in the long run content learning in CLIL is equal or even superior to that of mother tongue instruction (see e.g. Genesee 1987; Nikula & Marsh 1999c: 72; Seikkula-Leino 2002; Swain & Lapkin 1982). According to Nikula and Marsh (1997: 70), this could be explained by the fact that CLIL teaching often needs to focus on the most essential topics of a subject, because CLIL takes more time than monolingual lessons. Teachers have also reported that students concentrate better when taught in a foreign language (Nikula & Marsh 1997: 70) Jäppinen (2002, 2003, 2005) has also found that CLIL does not hinder content learning or cognitive development, as long as the CLIL teachers make sure the students develop basic cognitive abilities and learn the basic conceptual framework in all subjects already at the early stages of CLIL (see also Cummins & Swain 1986: 50—52).

It has also been found that CLIL does not have a negative effect on first language development but instead, can actually raise metalinguistic awareness that can be advantageous for any language learning. Parents have reported that CLIL instruction has enriched their children’s first language vocabulary and increased the awareness of first language structures and semantics (see e.g. Cummins

\(^5\) It should be noted that some of the studies referenced here were conducted in immersion settings that differ from CLIL settings in three respects: the official status of the L2 in the surrounding society, the language used in learning basic literacy skills, and the amount of L2 usage in teaching (see section 3.4.). Therefore, the findings might not be directly comparable even though in SLA literature and research reports they are, more often than not, treated as if they were.
CLIL also has clear advantages when it comes to foreign language learning. According to research, CLIL students generally acquire a wider vocabulary range, higher fluency, better comprehension skills, and better motivation and confidence to speak the language in comparison to mainstream students who receive formal foreign language instruction (see e.g. Genesee 1987; Nikula & Marsh 1997: 83—101, 1999d; Swain 1996; Swain & Lapkin 1982). Järvinen (1999: 137—138) has also found that their implicit knowledge of the accurate forms of the target language is also as good as or even better than that of mainstream language students, and they have also been found to produce more complex, accurate and versatile sentences than mainstream students. CLIL students also have the courage to try to use different forms of the language even before they have been formally taught, for example verb tenses or passive structures (Nikula & Marsh 1997: 85). The linguistic advantages have been especially visible in schools where a substantial amount of teaching is carried out in the target language. However, researchers assert that even a short language shower could have a positive impact on student motivation and foreign language skills.

Despite the clear advantages in foreign language learning, the grammatical proficiency of CLIL and immersion students has been found to lag behind their general L2 proficiency level. Despite the positive findings in Järvinen’s study (1999), a number of studies in Finland and abroad have shown that mainstream, instructed L2 learners produce more accurate target language than CLIL students (e.g. Genesee 1987; Kokkila 1996; Nikula & Marsh 1997: 90—91; Stanutz 1992; Swain 1996: 94). For example, Kokkila (1996: 105—107) found that CLIL students spoke more fluently and used more communicative strategies than mainstream students, but the latter used more accurate forms. In addition, regardless of the amount of CLIL attended, the productive skills of CLIL students do not equal their comprehension (Järvinen 1999: 18—19; Kokkila 1996; Nikula & Marsh 1997: 90; Swain 1996: 94). In comparison to native speakers of the target language, CLIL and immersion students operate with simpler and
grammatically less redundant verb structures (Cummins & Swain 1986; Järvinen 1999: 19).

The explanation for these differences in grammatical proficiency may lie in the meaning-focused nature of CLIL that emphasizes the implicit, natural acquisition of the L2. Many of the aforementioned studies have been conducted in CLIL and immersion settings that involve little or no formal, explicit L2 instruction. Typically for naturalistic settings, CLIL teachers often ignore grammatical accuracy if the students only manage to get their messages across (Mustaparta & Tella 1999: 25; Nikula & Marsh 1999d: 80). Today, however, many findings speak in favour of integrating explicit, form-focused L2 instruction to the CLIL setting (e.g. DeKeyser 2003; Doughty & Varela 1998; Ellis 2002; Lyster 2004; Norris and Ortega 2000; Sheen 2005; Spada & Lightbown 2008: 184; see more in subsection 4.1.1). This could also have its downsides, because the emphasis on accuracy may, at worst, cause communication apprehension in the L2: a relatively large proportion of mainstream students have been found to avoid speaking the language because they are afraid of making mistakes and subjecting themselves to negative evaluation (cf. Horwitz et al. 1986; Horwitz 2001; Korpela 2010; Nuto 2003; Yli-Renko 1991, 1993). In this respect, naturalistic CLIL without a focus on forms could have more potential to develop the confidence to use the foreign language, especially when it comes to oral communication.

Nevertheless, deficient acquisition of the L2 grammar can make it more difficult for CLIL students to express content mastery or deal with more complex subject matter, especially at higher stages of education (cf. Cummins & Swain 1986: 37—56; Doughty & Williams 1998: 2; Nikula & Marsh 1999c: 75, 1999d: 80—81; Swain 1996: 92—93; Swain & Lapkin 1982: 68). This can have various negative effects on their academic achievement and performance in the long run.

---

6 It is worth noting, though, that CLIL does not even aim at native-like proficiency.
5. Research Design

In the light of earlier research findings on the deficiencies in CLIL students’ grammatical accuracy in the L2 as well as on the benefits of L2 grammar instruction in content-based approaches, it is interesting to evaluate what kind of foreign language proficiency is acquired in a CLIL programme that also includes formal foreign language lessons. In order to evaluate the outcomes of this type of CLIL programme, it is important to find out if there are considerable deficiencies in the CLIL students’ grammatical proficiency in comparison to mainstream students receiving formal EFL instruction.

The purpose of this study is to compare the grammatical proficiency of English between CLIL students and mainstream students in Töölö Primary School. The main aim is to analyse, describe and interpret what kind of differences there are in the grammatical proficiency of the two groups, if any, and what types of grammatical errors the students produce. Earlier research on both immersion and CLIL has shown that students who receive formal L2 instruction produce more grammatically accurate English than the CLIL students. However, these studies have mainly been conducted in CLIL settings with little or no formal, explicit foreign language instruction. Besides, some findings indicate that CLIL students actually gain better implicit knowledge and productive ability in the L2 than mainstream students. This study seeks to re-evaluate these findings.
5.1. Research Questions

The present study seeks answers to the following questions:

1. What differences can be observed in CLIL students’ and mainstream students’ grammatical proficiency of English?
   1.1. What differences, if any, are there in the CLIL students’ and mainstream students’ implicit knowledge of the English grammar?
   1.2. What differences, if any, are there in the CLIL students’ and mainstream students’ productive skills in written English?

2. What types of errors do CLIL students and mainstream students make in translation from Finnish into English and written composition in English?

5.2. Data Collection

The data was collected in April 2013 by means of a written English grammar test and a background questionnaire (see Appendix 1). The written grammar test consisted of four parts: grammaticality judgments, a cloze passage exercise, translation of questions, and written composition.

1. **Grammaticality judgments**: the first part of the test consisted of grammaticality judgments, generally considered a valid method for assessing the *implicit knowledge* of a foreign language (Järvinen 1999). The part consisted of ten sentences, all written three times as options a, b and c, only one of which was grammatically correct. The students had to tick the option they deemed correct. However, they did not have to spot the errors or correct the phrases. To add variation to the data, some phrases (#7, #8, #9) were purposefully difficult and contained grammatical items the students might not have explicitly learned yet. It was assumed, however, that they have most likely seen or heard these forms somewhere, so they would be able to tell which form “sounds right”: the teacher has been using them, they have seen them on the Internet, they have heard them on TV, and so on. These sentences especially measured the students’ implicit knowledge of the English grammar. Table 2 presents the correct sentences and

7 However, the method does have its limitations which are discussed in more detail in section 7.1.
the knowledge of the grammatical items measured by each sentence. The complete exercise, including the 20 erroneous sentences, can be seen in Appendix 1.

**Table 2:** The sentences in the grammaticality judgment part and the grammatical items measured by each sentence.

<table>
<thead>
<tr>
<th>Sentence</th>
<th>Grammatical items measured</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. b. I watched <strong>the film</strong> that you told me about last Saturday.</td>
<td>Use of definite article 'the'</td>
</tr>
<tr>
<td>2. b. I think the film was awesome but my sister <strong>didn’t like</strong> it.</td>
<td>Past tense: 3rd person sg. -s</td>
</tr>
<tr>
<td>3. a. <strong>Her</strong> taste in films is just too different from mine.</td>
<td>Possessive pronoun ‘her’</td>
</tr>
<tr>
<td>4. b. <strong>There were</strong> many frightening scenes in the film.</td>
<td>There was/were</td>
</tr>
<tr>
<td>5. a. Yesterday <strong>I came</strong> to school by bike.</td>
<td>Past tense: irregular verb ‘come/came’</td>
</tr>
<tr>
<td>6. a. My bike is broken but my <strong>dad said</strong> that he won’t buy <strong>me</strong> a new one before summer.</td>
<td>Past tense: 3rd person sg. –s Objective pronoun ‘me’</td>
</tr>
<tr>
<td>7. c. I don’t know why <strong>he said that</strong>.</td>
<td>Indirect speech</td>
</tr>
<tr>
<td>8. a. Maybe <strong>he hasn’t got enough money</strong>.</td>
<td>3rd person sg. ‘has got’</td>
</tr>
<tr>
<td>9. a. On my way to school I saw <strong>a dog walking</strong> down the street and <strong>jumping on people</strong>.</td>
<td>Non-finite clauses</td>
</tr>
<tr>
<td>10. a. I was <strong>late from school</strong> because I stopped to watch <strong>him</strong>.</td>
<td>Preposition structure ‘late from’</td>
</tr>
</tbody>
</table>

2. **Cloze passage:** the second part was a cloze passage exercise in which the students had to complete 24 gaps in a story by choosing the correct form from three options written in English. Another way to construct this exercise would have been a similar cloze passage with a Finnish translation of the word(s) that go in the gaps, but as this would have focused too much on testing the knowledge of vocabulary besides grammar itself, I chose to give the options in English. The forms to fill in the gaps consisted of verbs tenses and forms (irregular verbs, third person singular –s), various prepositions (verb+preposition structures, time and place), comparison of adjectives, nouns (irregular plurals), and personal pronouns (possessive, objective and indefinite forms). The exercise was
modified from ‘Part III: Structures’ in the National English Examination (valtakunnallinen englannin koe) for sixth graders in 2009.

3. **Translation of questions:** The third part contained five questions that the students had to translate from Finnish into English. The part tested production skills in written English by means of question formation. The scene was an imaginary interview of a film star:

2. Osaatko puhua espanjaa hyvin? (can – Spanish – well)
3. Mitä mieltä äitisi on tästä elokuvasta? (think about)
4. Missä sinä vietit eilisillan?
5. Oliko tämä sinun viimeinen elokuvasi?

In the first three questions English words were given to the students but not in the correct order. In the first question, all words were given but in incorrect order, in order to see if the student masters the word order even if he or she did not remember the words in English. The last two questions did not contain any tips. This part focused particularly on the knowledge of auxiliary verbs, past tense, prepositions and word order. The exercise was taken directly from the National English Examination for sixth graders in 2008.

4. **Written composition:** The fourth part was a short written composition with a minimum of 40 words, which tested production skills in written English. Twenty expressions in four categories (nouns, adjectives, past tense of verbs and prepositions of time) were given to the students in Finnish, and they had to use at least eight of them in their story, two from each category. This requirement ensured the stories would elicit enough grammatical data for analysis, yet the large number of expressions gave the students enough freedom of choice in case they would not know some of them in English, for example. The exercise was modified from the National English Examination for sixth graders in 2003.

The background questionnaire gathered information about the students’ year of birth, gender, class, mother tongue, the year they started studying English at school and their English mark in the previous certificate. All of these were control questions to assist in the selection of informants (see section 5.3.).
I administered the English grammar test and the background questionnaire during normal school days. The test took about 60 minutes to complete. The mainstream class took the test in the afternoon, starting at 13:30, so some students may have been a bit tired by the time they reached the written composition (part 4). This may have affected the results to some extent. Anyway, everyone managed to complete the test without interruptions.

5.3. The Informants

The informants were a group of 30 sixth and fifth grade students from Töölö Primary School in Helsinki. 14 informants (47 %, n = 30) were CLIL students who attended the bilingual fifth grade, while 16 informants (53%, n = 30) were sixth-grade mainstream students who attended mainstream EFL lessons. There were 11 females (36.7%, n = 30) and 17 males (56.7%, n = 30), and two informants (6.7%, n = 30) did not reveal their gender. All CLIL informants were born in 2001 and the mainstream informants mainly in 2000, except for two informants born in 1999 and two in 2001. Native speakers of English and students having studied or learned English formally before the start of primary school were excluded from the analysis. The informants were, therefore, chosen by purposeful selection, also known as discretionary sampling, purposeful sampling or criterion-based selection (cf. Maxwell 2005: 88, Tuomi & Sarajärvi 2003: 89). As opposed to random sampling, purposeful selection means that a number of unique, informative cases are deliberately selected for analysis to ensure they meet the research criteria (cf. Maxwell 2005: 88).

Table 3: The informants: class, gender and year of birth.

<table>
<thead>
<tr>
<th>Class</th>
<th>Number of informants</th>
<th>Gender</th>
<th>Year of birth</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>1999</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>2000</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>2001</td>
</tr>
<tr>
<td>6B mnstr.</td>
<td>16 (53%)</td>
<td>Female</td>
<td>Male 11</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5C CLIL</td>
<td>14 (47%)</td>
<td>Female</td>
<td>Male 6</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>2001 (n = 14)</td>
</tr>
<tr>
<td>Total:</td>
<td>30 (100%)</td>
<td>Female:</td>
<td>Male:</td>
</tr>
<tr>
<td></td>
<td></td>
<td>11 (36.7%)</td>
<td>17 (56.7%)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
The students were also asked to provide their previous English marks. Table 4 shows that while 5C is a more homogenous group in this respect – there are no marks below 8 (‘good knowledge’) – the students on 6B got more top marks: 37.5% of all the students on class 6B got a 10. However, the means (6B: 8.88, 5C: 8.79) imply that the groups have more or less equivalent English skills.

**Table 4:** English mark in the previous certificate.

<table>
<thead>
<tr>
<th>Mark</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
<th>Total</th>
<th>Mean</th>
</tr>
</thead>
<tbody>
<tr>
<td>Class</td>
<td>6B</td>
<td>2 (12.5%, n = 16)</td>
<td>4 (25%, n = 16)</td>
<td>4 (25%, n = 16)</td>
<td>6 (37.5%, n = 16)</td>
<td>16</td>
</tr>
<tr>
<td></td>
<td>5C</td>
<td>0 (0%, n = 14)</td>
<td>5 (35.7%, n = 14)</td>
<td>7 (50%, n = 14)</td>
<td>2 (14.3%, n = 14)</td>
<td>14</td>
</tr>
<tr>
<td>Total</td>
<td></td>
<td>2 (6.7%, n = 30)</td>
<td>9 (30%, n = 30)</td>
<td>11 (36.7%, n = 30)</td>
<td>8 (26.7%, n = 30)</td>
<td>30</td>
</tr>
</tbody>
</table>

The CLIL students were selected from one grade lower than the mainstream students for two reasons. First, having started English on the first grade, the fifth-grade CLIL students were two years ahead in their formal English studies than the sixth-grade mainstream students who had started English on the third grade. By selecting fifth-grade CLIL students and sixth-grade mainstream students this discrepancy was levelled by one year, or even more, as the EFL lessons during the first year of CLIL are very play-like compared to mainstream third grade EFL lessons, mostly because the first graders cannot even read or write Finnish. The second reason was that the fifth-grade CLIL students had used the same English textbook and workbook as the sixth-grade mainstream students during the formal EFL lessons. This guaranteed that both groups had done exactly the same grammar exercises and covered the same grammar topics before the data collection. The CLIL students had studied grammar a bit further (see Table 1 earlier), but this was taken into account in the test design and data analysis.
5.4. Data Analysis

The comparison of CLIL and mainstream students’ grammatical proficiency in English was based on the total scores they obtained in different parts of the test as well as on the categorization and analysis of the types of errors they made in the translation of questions and in the written composition.

First of all, I analysed the students’ performance in parts 1, 2 and 3 by quantitative methods with the help of the SPSS statistics programme. I entered the data into SPSS in April 2013 and conducted the analysis in May and June 2013. First, I checked the data for errors. Then, I compared the test results of classes 5C and 6B by calculating the frequencies, mean values, medians, standard deviations, variance, and minimum and maximum values for each item and/or grammatical category in parts 1, 2 and 3 of the test. I also compared the students’ total mean scores for parts 1, 2 and 3.

I verified the reliability and validity of the quantitative parts of test by calculating the correlation between the test results and the students’ past English marks as well as the intercorrelations for parts 1, 2 and 3 of the test. A statistically significant positive correlation ($r = .511—.631, p < 0.01$; see Appendix 3) was found between the students’ past English marks and the results for parts 1, 2 and 3. The intercorrelations of the results for parts 1, 2 and 3 ranged between $r = .531—.635 (p < 0.01)$ (see Appendix 3), indicating high internal consistency.

Lastly, to obtain a more detailed picture of the CLIL and mainstream students’ grammatical proficiency and to compare the type of language they produce, I categorized the errors made in the translation of questions (part 3) and written composition (part 4) following the principles and practices of descriptive error analysis (EA) (Ellis 1995: 51—59). The analysis was based on the error categorization presented in section 4.3. (Figure 1), with some modifications and additions. The extended categorization used here will be presented in subsection 6.3.2. (Figure 2). Despite the problems of categorization discussed in section 4.3., descriptive error analysis is regarded an advantageous tool in evaluating learner language for pedagogical and diagnostic purposes. In the analysis, an ‘error’ was seen as any written form that deviated from the norms of the target language.
The point of reference was the standard written dialect generally taught in school.

Researchers have pointed out that the method of data collection can have a marked impact on the results obtained. Clinical elicitation, which is used in this study, prompts different language use than naturalistic situations. Even free composition and translation have been noticed to bring about markedly different results: L1 influence, for example, is much more evident in translation tasks (Lococo 1976; Ellis 1995: 50). For this reason, I used both translation and free written composition to obtain as comprehensive data as possible. Besides, data collection in naturalistic situations would not have served the purposes of this study as effectively as clinical elicitation. Any incorrect hypotheses the learners have are more likely to appear in a written test, because it is a formal situation where language users typically aim at careful, correct language (Ellis 1995: 22).
6. Results and Analysis

The following sections will give an account of the major findings of the present study, starting with a comparison of CLIL and mainstream students’ implicit knowledge of the English grammar, moving on to a comparison and analysis of the total and item-specific mean scores from the cloze passage and the translation of questions, and finishing with an analysis of the errors made in the translation of questions and the written composition.

6.1. The Implicit Knowledge of Grammar: Grammaticality Judgments

The implicit knowledge of English grammar was measured by means of ten sentences all written as three options, only one of which was grammatically correct (see part 1, Appendix 1). The students had to tick the correct one. Each correct answer was worth one (1) point so that the total maximum score was 10 points. The mean totals and distributions of total scores are presented in Tables 5a and 5b.

Table 5a: Implicit knowledge of English grammar: the mean totals from the grammaticality judgments.

<table>
<thead>
<tr>
<th>Class</th>
<th>Mean</th>
<th>Variance</th>
<th>Std. Deviation</th>
<th>N</th>
</tr>
</thead>
<tbody>
<tr>
<td>6B mainstream</td>
<td>7.44</td>
<td>2.929</td>
<td>1.711</td>
<td>16</td>
</tr>
<tr>
<td>5C CLIL</td>
<td>8.07</td>
<td>2.071</td>
<td>1.439</td>
<td>14</td>
</tr>
<tr>
<td>Total</td>
<td>7.73</td>
<td>2.547</td>
<td>1.596</td>
<td>30</td>
</tr>
</tbody>
</table>
Table 5b: Implicit knowledge of English grammar: the distribution of total scores.

<table>
<thead>
<tr>
<th>Class</th>
<th>TOTAL SCORE: IMPLICIT KNOWLEDGE OF GRAMMAR</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>4/10</td>
<td>5/10</td>
</tr>
<tr>
<td>6B</td>
<td>Count</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>% within class</td>
<td>6.3%</td>
</tr>
<tr>
<td>5C</td>
<td>Count</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>% within class</td>
<td>0.0%</td>
</tr>
<tr>
<td>Total</td>
<td>% of all answers</td>
<td>3.3%</td>
</tr>
</tbody>
</table>

Judging by the mean scores alone (Table 5a), the CLIL students’ implicit knowledge of English grammar is somewhat better than that of mainstream students. The CLIL students’ skills are also more homogeneous: there is greater variance and standard deviation in 6B’s answers. The same tendency can be seen in the total score frequencies (Table 5b): the mainstream students got either high or low total scores with no results in between (0.0% got 7/10). The number of proficient students on both classes was more or less the same: a total of 68.9% of the mainstream students and 70% of the CLIL students got eight or more correct answers out of 10.

To see which sentences and grammatical items were the most problematic, the frequency (number of occurrences and percentage) of incorrect answers within class and within the whole data set was calculated for each sentence. The distributions are presented in Table 6 on the next page.
Table 6: Implicit knowledge of the English grammar: the distribution of incorrect answers.

<table>
<thead>
<tr>
<th>Sentence</th>
<th>Grammatical items measured</th>
<th>Number and % of incorrect answers: 6B mnstr. (n=16)</th>
<th>Number and % of incorrect answers: 5C CLIL (n=14)</th>
<th>Total number and % of incorrect answers (n=30)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. b. I watched the film that you told me about last Saturday.</td>
<td>Use of definite article</td>
<td>3 18.8% within class</td>
<td>3 21.4% within class</td>
<td>6 20.0% of all responses</td>
</tr>
<tr>
<td>2. b. I think the film was awesome but my sister didn’t like it.</td>
<td>Past tense: 3rd person sg. -s</td>
<td>1 6.3% within class</td>
<td>0 0.0% within class</td>
<td>1 3.3% of all responses</td>
</tr>
<tr>
<td>3. a. Her taste in films is just too different from mine.</td>
<td>Possessive pronoun ‘her’</td>
<td>1 6.3% within class</td>
<td>0 0.0% within class</td>
<td>1 3.3% of all responses</td>
</tr>
<tr>
<td>4. b. There were many frightening scenes in the film.</td>
<td>There was/were</td>
<td>2 12.5% within class</td>
<td>2 14.3% within class</td>
<td>4 13.3% of all responses</td>
</tr>
<tr>
<td>5. a. Yesterday I came to school by bike.</td>
<td>Past tense: irregular verb ‘come/came’</td>
<td>3 18.8% within class</td>
<td>0 0.0% within class</td>
<td>3 10.0% of all responses</td>
</tr>
<tr>
<td>6. a. My bike is broken but my dad said that he won’t buy me a new one before summer.</td>
<td>Past tense: 3rd person sg. –s</td>
<td>3 18.8% within class</td>
<td>2 14.3% within class</td>
<td>5 16.7% of all responses</td>
</tr>
<tr>
<td>7. c. I don’t know why he said that.</td>
<td>Indirect speech</td>
<td>5 31.3% within class</td>
<td>4 28.6% within class</td>
<td>9 30.0% of all responses</td>
</tr>
<tr>
<td>8. a. Maybe he hasn’t got enough money.</td>
<td>3rd person sg. ‘has got’ Phrase structure: placing ‘enough’ in a sentence</td>
<td>14 87.5% within class</td>
<td>9 64.3% within class</td>
<td>23 76.7% of all responses</td>
</tr>
<tr>
<td>9. a. On my way to school I saw a dog walking down the street and jumping on people.</td>
<td>Non-finite clauses</td>
<td>9 56.3% within class</td>
<td>6 42.9% within class</td>
<td>15 50.0% of all responses</td>
</tr>
<tr>
<td>10. a. I was late from school because I stopped to watch him.</td>
<td>Preposition structure ‘late from’ Objective pronoun ‘him’</td>
<td>0 0.0% within class</td>
<td>1 7.1% within class</td>
<td>1 3.3% of all responses</td>
</tr>
</tbody>
</table>

Table 6 clearly shows that structure #8 “Maybe he hasn’t got enough money” was, surprisingly, the most difficult one for both CLIL and mainstream students with as many as 23 (76.7%) incorrect answers out of 30. The students had either placed ‘enough’ at the end of the sentence (“he doesn’t have money enough”) or
ignored the third person singular -s ("he haven’t got enough money"). The former is most likely a result of transfer from Finnish while the latter is a common interlanguage feature among EFL learners.

Other grammatical items that caused problems were non-finite clauses (#9 "I saw a dog walking down the street and jumping on people"), with 15 (50.0%) incorrect answers out of 30, and indirect speech (#7 "I don’t know why he said that"), with 9 (30.0%) incorrect answers out of 30. Interestingly, neither of these forms had been explicitly taught to the students, as far as what can be inferred from the grammatical contents specified in the Töölö Primary School curriculum for English as an A1 language (2005 and 2012). In the light of these findings, formal, explicit instruction might indeed promote L2 acquisition, because the students were most at ease with structures that had been explicitly taught. This is in line with the output hypothesis (Swain 1985) and a multitude of findings on the benefits of formal L2 instruction (e.g. DeKeyser 2003; Doughty 1991; Doughty & Varela 1998; Ellis 1990, 2002; Genesee 1987; Lightbown & Spada 1990; Long 1983; Lyster 2004; Norris and Ortega 2000; Pica 1983; Sheen 2005; Spada & Lightbown 1993; White et al. 1991), but contrary to what for example Krashen (1984) and a number of other researchers (Eubank & Beck 1995: 40; Truscott 1998: 119; VanPatten & Cadierno 1993; Wong 2004) have concluded about the distinct nature of explicit and implicit L2 knowledge.

Despite some erroneous answers, the CLIL students outperformed the mainstream students in the majority of cases (7 out of 10). The CLIL students displayed a relatively higher percentage of incorrect answers for sentences #1, #4 and #10, but these differences are insignificant in such a small data set. The grammaticality judgments show that the CLIL students’ implicit knowledge of the English grammar is better than that of mainstream students. This finding is in line with Järvinen’s (1999: 136—137) study, for example, where she found that the implicit knowledge of the target language develops faster in CLIL than mainstream L2 instruction. The possible explanations to this finding will be discussed in more detail in chapter 8.
6.2. Knowledge of Different Grammatical Categories

The cloze passage exercise (part 2) measured the students’ grammatical proficiency in terms of five broad categories: nouns (irregular plurals), verbs (3rd person singular –s, irregular past tense), adjectives (comparison), pronouns (possessive, objective, indefinite), and prepositions (some common structures). Each category was measured with four gaps to fill in, making up a total of 20 items. Each right answer was worth one point; thus, the total score was 20.

Table 7a illustrates the students’ total scores from this exercise. Once again, the CLIL students (mean total 16.71) outperform the mainstream students (mean total 14.69). The minimum and maximum values, the standard deviation as well as the variance indicate that the mainstream students also possess a more diverse knowledge English: the lowest total score in this group was 5/20 with a standard deviation of 3.928 while the CLIL students have all scored between 11/20—20/20, the standard deviation being 2.555.

Table 7a: The distribution of mean totals from the cloze passage.

<table>
<thead>
<tr>
<th>Class</th>
<th>MEAN TOTAL (out of 20.00)</th>
<th>Minimum</th>
<th>Maximum</th>
<th>Median</th>
<th>Std. Deviation</th>
<th>Variance</th>
<th>N</th>
</tr>
</thead>
<tbody>
<tr>
<td>6B</td>
<td>14.69</td>
<td>5</td>
<td>19</td>
<td>15.00</td>
<td>3.928</td>
<td>15.429</td>
<td>16</td>
</tr>
<tr>
<td>5C</td>
<td>16.71</td>
<td>11</td>
<td>19</td>
<td>17.50</td>
<td>2.555</td>
<td>6.527</td>
<td>14</td>
</tr>
<tr>
<td>Total</td>
<td>15.63</td>
<td>5</td>
<td>19</td>
<td>16.00</td>
<td>3.459</td>
<td>11.964</td>
<td>30</td>
</tr>
</tbody>
</table>

A closer look at the grammatical categories (Table 7b) reveals that the comparison of adjectives caused the most difficulties, the total score being the lowest (mainstream students 2.38, CLIL students 3.00, mean total 2.67). The CLIL students demonstrate a better knowledge of all grammatical categories except for pronouns (CLIL students 3.29 < mainstream students 3.38). However, such a small difference (0.09) is statistically insignificant. Interestingly, the present data disagree with a number of earlier studies where instructed, mainstream L2 learners have systematically outperformed CLIL students in grammatical

Table 7b: The distribution of mean totals in different grammatical categories.

<table>
<thead>
<tr>
<th>Class</th>
<th>Nouns</th>
<th>Verbs</th>
<th>Adjectives</th>
<th>Pronouns</th>
<th>Prepositions</th>
</tr>
</thead>
<tbody>
<tr>
<td>6B</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean</td>
<td>2.94</td>
<td>2.81</td>
<td>2.38</td>
<td>3.38</td>
<td>3.19</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Minimum</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>Maximum</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td>Median</td>
<td>4.00</td>
<td>3.00</td>
<td>3.00</td>
<td>3.00</td>
<td>3.00</td>
</tr>
<tr>
<td>Std. Deviation</td>
<td>1.389</td>
<td>1.167</td>
<td>1.147</td>
<td>.619</td>
<td>.834</td>
</tr>
<tr>
<td>Variance</td>
<td>1.929</td>
<td>1.363</td>
<td>1.317</td>
<td>.383</td>
<td>.696</td>
</tr>
<tr>
<td>N</td>
<td>16</td>
<td>16</td>
<td>16</td>
<td>16</td>
<td>16</td>
</tr>
<tr>
<td>5C</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean</td>
<td>3.29</td>
<td>3.64</td>
<td>3.00</td>
<td>3.29</td>
<td>3.50</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Minimum</td>
<td>1</td>
<td>3</td>
<td>1</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>Maximum</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td>Median</td>
<td>4.00</td>
<td>4.00</td>
<td>3.00</td>
<td>3.00</td>
<td>4.00</td>
</tr>
<tr>
<td>Std. Deviation</td>
<td>.994</td>
<td>.497</td>
<td>.877</td>
<td>.611</td>
<td>.650</td>
</tr>
<tr>
<td>Variance</td>
<td>.989</td>
<td>.247</td>
<td>.769</td>
<td>.374</td>
<td>.423</td>
</tr>
<tr>
<td>N</td>
<td>14</td>
<td>14</td>
<td>14</td>
<td>14</td>
<td>14</td>
</tr>
<tr>
<td>Total</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean</td>
<td>3.10</td>
<td>3.20</td>
<td>2.67</td>
<td>3.33</td>
<td>3.33</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Minimum</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>Maximum</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td>Median</td>
<td>4.00</td>
<td>3.00</td>
<td>3.00</td>
<td>3.00</td>
<td>3.50</td>
</tr>
<tr>
<td>Std. Deviation</td>
<td>1.213</td>
<td>.997</td>
<td>1.061</td>
<td>.606</td>
<td>.758</td>
</tr>
<tr>
<td>Variance</td>
<td>1.472</td>
<td>.993</td>
<td>1.126</td>
<td>.368</td>
<td>.575</td>
</tr>
<tr>
<td>N</td>
<td>30</td>
<td>30</td>
<td>30</td>
<td>30</td>
<td>30</td>
</tr>
</tbody>
</table>

6.3. CLIL and Mainstream Students’ Productive Skills Compared

The students’ productive skills in English were measured by parts 3 and 4 of the test, which involved the translation of questions from Finnish into English and a written composition with a minimum of 40 words. The following subsections will present the results: first, I will examine the mean totals from the translation of
questions, and then carry on with the analysis of errors produced in the translation of questions and the written composition.

6.3.1. Translation of Questions from Finnish into English

The translation exercise (part 3) consisted of five questions that the students needed to translate from Finnish into English (see section 5.2. or Appendix 1). The acceptable translations are presented below. Any deviations from these were marked as errors.

1. Where do you live now? (6.0 points)
2. Can you speak Spanish well? (6.0 points)
3. What does your mom/mum/mother think about this film/movie? (9.0 points)
4. EITHER: Where did you spend last/yesterday night/evening? (7.0 points)
   OR: Where were you last/yesterday night/evening? (7.0 points)
5. Was this your last film/movie? (6.0 points)

The highest possible scores for each question are also presented above. Each correct word counted as 1.0 point. In addition, the student received 1.0 point for correct word order. Each error reduced the total score by 0.5-1.0 points depending on the gravity of the error. For example, minor spelling mistakes took off -0.5 points and incorrect or missing words -1.0 points.

Table 8 illustrates the mean totals for the whole exercise as well as for each question. As expected, the greatest differences in performance (CLIL students 8.75 > mainstream students 7.97) occurred in question 3, which was the longest sentence and the most complicated structure. The CLIL students got higher mean totals for all parts of the exercise (CLIL students 32.75 > mainstream students 30.56), indicating that the CLIL students’ productive skills are somewhat better. This agrees with Järvinen’s (1999) finding that CLIL students outperform mainstream students in target language production, but contrasts with a number of other studies where the CLIL students lag behind in grammatical accuracy (Genesee 1987; Kokkila 1996; Nikula & Marsh 1997; Stanutz 1992; Swain 1996).
Despite the higher mean totals obtained by the CLIL students, the groups’ total scores actually differ by as little as 2.187 points. This indicates that the mainstream students have acquired nearly as high proficiency as the CLIL students by attending EFL lessons twice a week for only four years. By contrast, the CLIL students have been exposed to a significantly greater amount of English at school, with 30—50% of content instruction in English and formal EFL lessons
two or three times a week for five years. In this respect, the productive skills of
the two groups are surprisingly equal. One plausible explanation for this could be
task familiarity: the mainstream students are used to translation exercises of this
type whereas the CLIL students are advised against thinking first in Finnish and
then translating into English. On the other hand, the exercise seems to have been
on the easier side, considering the fact that the total scores for the CLIL and
mainstream students ranged from 30.0 to 34.0 and 23.0 to 33.5 respectively. The
small range suggests the exercise might have not been distinguishing enough.

6.3.2. The Types of Errors Produced: Translation of Questions

To get a more comprehensive picture of the CLIL and mainstream students’
productive skills in written English, the errors that the students produced in the
translation of questions (part 3) and the written composition (part 4) were also
analysed and categorized. The exercises provided different data: while the
questions (part 3) tested the students’ ability to translate a set of pre-formulated
sentences from Finnish to English, the written composition (part 4) allowed for
more freedom of expression and thus elicited more diverse written language and
errors. For this reason, the errors produced in these exercises were also
examined separately. First, I will look into the erroneous translations of
questions.

A total of 28 translations were included in the analysis, 14 by CLIL and 14 by
mainstream students. The nature of the error analysis required two mainstream
answers to be left out so that the results would be comparable. The included
answers were randomly selected. Tables 9a and 9b list all erroneous translations
classified on the basis of error types and their descriptions. The error analysis
relied on the categorization presented earlier in Figure 1 (Ellis 1995: 58; Lott
1983; Richards: 1971), with one exception: the subcategory 2c Incomplete
application of rules was divided into 2c Omissions and 2e Structural failures for
the sake of accuracy (see Figure 2). Omissions refer to the omission of
grammatical or lexical constituents whereas structural failures involve such
errors as incorrect word order or failure to develop a grammatically correct
phrase structure (see Tables 9a and 9b for examples in the present data).
Figure 2: The modified categories used in the error analysis.

Before examining the data in more detail, I will give a brief explanation about the error categorization process. The numbering of translations (marked with #) refers to the list in Tables 9a and 9b.

Most errors, such as omissions and transfer of L1 structure, were easy to categorize. Minor spelling mistakes (in translations #4, #17, #19, #21, #28 and #32) and errors that did not have a marked effect on the grammaticality of the sentence, such as using the auxiliary verb ‘did’ instead of ‘do’ in translation #31 (“What did your mother think about this film?”), were categorized as occasional performance mistakes under category 3. Processing problems. The assumption was that the students knew the right form or spelling but failed to produce it in the test situation. An exception was translation #25 (“Was that your last movie?”) where the student used ‘that’ instead of ‘this’. This mistake was, likewise, regarded as a performance mistake, but a strategic one (4. Communication strategies), assuming that the student either mixed ‘this’ and ‘that’ or did not remember the correct word for Finnish ‘tämä’ (‘this’) in English
and replaced it with a word from the same grammatical category, objective pronouns (‘tuo’, ‘that’). The translation was, nonetheless, structurally correct.

Some errors were placed into more than one category, because their origin and type were ambiguous. For example, the errors in translation #3 (“Can you well Spanish?”) were originally placed into categories 2c. Omissions (of ‘speak’) and 2e. Structural failures (‘well’ placed incorrectly), but on a closer look, the translation could also have stemmed from the Finnish structure “Osaatko hyvin espanjaa?”, in which case the sentence would fall into error category 1b. Transfer of L1 structure. The sentence was eventually placed into all three categories.

The errors belonging to category 2d. False concepts hypothesized were, at first, difficult to distinguish. To facilitate the classification, I determined that the most important criterion would be the inability to understand a distinction in the target language. The erroneous translations #5 and #20 through to #24 fall into this category. The errors in these sentences stemmed mainly from the students’ inability to comprehend the distinction between the auxiliary verbs ‘are’, ‘do’, ‘is’ and ‘was’, or the inability to use an auxiliary verb that agrees with the subject of the phrase (“Was this your last movie?”). The students’ general performance in the test needed to be taken into consideration, however. An example of a case like this is translation #21 that contains two errors (“Do this you last film?”). The student clearly knew the distinction between ‘you’ and ‘your’ because he produced no such error in other parts of the test, so the ‘you’ in this sentence was categorized as an accidental mistake, i.e. ‘3. Processing problem’. However, the student who produced translation #5 (“What think you mom about this film?”) consistently used ‘you’ where ‘your’ was required, so the error was placed into the 2d. False concepts hypothesized category.
Table 9a: The errors produced by the mainstream students in translation.

| Error types (number and % within class 6B; total n of errors =45, n of responses =14): |
|---------------------------------|---------------------------------|
| **1b. Transfer of L1 structure (7; 15.6%, n=45)** | **2a. Overgeneralizations (4; 8.9%, n=45)** |
| 10. Where you were yesterday? (1) | 2. Can you speak Spanish? (1) |
| 13. Where you was yeastery evening? (1) | 12. Where did you spent your last night? (2) |
| 14. Where you ___ last night? (1) | 18. Where you was on last night? (1) |
| 15. Where you wer last evening? (1) | 16. Where you were last evening? (2) |
| 18. Where you was on last night? (1) | 21. Do this you last film? (1) |
| 2. Overgeneralization of verb form 'was' and preposition 'on' |  |
| **2b. Ignorance of rule restrictions (2; 4.4%, n=45)** |  |
| 7. What think about your mom are there film? (1) |  |
| 18. Where you was on last night? (1) |  |
| 2b. Omissions (12; 26.7%, n=45) | 2d. False concepts hypothesized (6; 13.3%, n=45): |
| 1. Where do you live? (1) | 5. What think you mom about this film? (1) |
| 3. Can you well Spanish? (1) | 20. Are this you last movie? (1) |
| 5. What think you mom about this film? (1) | 21. Do this you last film? (1) |
| 7. What think about your mom are there film? (1) | 22. Are this your last film? (1) |
| 8. What your mom think about this film? (5) | 23. Are this your last movie? (1) |
| 9. What you think about your mom this movie? (1) | 24. Is this your last movie? (1) |
| 11. Where you last night? (1) |  |
| 2c. Structural failures (4; 8.9%, n=45): |  |
| 3. Can you well Spanish? (1) |  |
| 5. What think you mom about this film? (1) | 7. What think about your mom are there film? (1) |
| 9. What you think about your mom this movie? (1) |  |
| 3. Processing problems (7; 15.6%, n=45): | 9. What you think about your mom this movie? (1) |
| 4. Can you speak Spanish? (1) |  |
| 6. What did your mom think about this film? (1) |  |
| 13. Where you was yeastery evening? (1) |  |
| 15. Where you wer last evening? (1) |  |
| 17. Where did you spent your last evening? (1) |  |
| 19. Where were you last night? (1) |  |
| 21. Do this you last film? (1) |  |
| 4. Communication strategies (3; 6.6%, n=45): |  |
| 14. Where you ___ last night? (1) |  |
| 25. Was that your last movie? (2) |  |
| **Total: 45 errors (100%)** |  |
Table 9b: The errors produced by the CLIL students in translation.

Table 9b: The errors produced by the CLIL students in translation.

| Error types (number and % within class 5C, total n of errors =13, n of responses =14): |
| 1b. Transfer of L1 structure (3; 23.1%, n=13): |
| 26. Where you were last night? (3) transfer of Finnish word order |
| 2a. Overgeneralizations (4, 30.7%, n=13): |
| 27. Where did you spent last night? (2) using ‘did’+‘spent’ in the same phrase |
| 28. Where did you spent you last night? (1) using ‘did’+‘spent’ in the same phrase |
| 29. Where did you were last night? (1) using ‘did’+‘were’ in the same phrase |
| 2c. Omissions (3, 23.1%, n=13): |
| 30. What your mother think about this film? (3) omission of ‘does’ |
| 3. Processing problems (3, 23.1%, n=13): |
| 28. Where did you spent you last night? (1) ‘you’ is a spelling mistake |
| 31. What did your mother think about this film? (1) accidental use of ‘did’ instead of ‘does’ |
| 32. Where this your last movie? (1) ‘where’ instead of ‘were; stemming most likely from phonology, ‘where’ and ‘were’ are pronounced similarly |
| Total: 13 errors (100%) |

**Figure 3:** The translation of questions: the number of errors produced per category.

Tables 9a—9b and Figure 3 show that the CLIL students produced significantly fewer errors (13 occurrences) than the mainstream students (45 occurrences) in translation from Finnish into English. The errors produced by the CLIL students were fairly evenly distributed into four categories: *transfer of L1 structures*
(23.1%), overgeneralizations (30.7%), omissions (23.1%) and processing problems (23.1%), the most common type being the overgeneralization of two past tenses within the same phrase (30.7% of all errors within the CLIL group). Likely explanations for these errors include a lack of explicit L2 knowledge, fossilized forms in the interlanguage, little explicit instruction about the past tense, or just plain inattentiveness. The CLIL students do, after all, encounter a great variety of past tenses during the fifth grade in subjects like history, so the lack of input cannot account for these errors.

Interestingly, overgeneralizations were not as common in the mainstream students’ translations, with only four occurrences (8.9%) out of all the 45 erroneous sentences produced by the mainstream students. Instead, the majority of the mainstream students’ errors involved omissions (26.7%) of either auxiliary verbs or other grammatical or semantic constituents. Transfer of L1 structures was also rather frequent (15.6%), mostly concerning the Finnish word order where the subject comes right after the interrogative pronoun. False concepts hypothesized (13.3%) and processing problems (15.6%) were also fairly common. These mainly included incorrect auxiliary verbs or spelling mistakes. Other error types produced by the mainstream students were overgeneralizations (8.9%), structural failures (8.9%), errors originating in communication strategies (6.6%), and ignorance of rule restrictions (4.4%). The overgeneralization errors mainly concerned verb forms such as present tense third person singular –s (#2 “Can you speaks Spanish well?”) or past tense in questions (#12 “Where did you spent your last night?”), whereas the structural failures refer to incorrect word order.

If we examine the CLIL and mainstream groups’ results separately and calculate the frequency (%) of each error type in relation to the total number of erroneous translations produced within the group (6B n=45; 5C n=13), we can see that three out of the four error types produced by the CLIL students were also the most common types among the mainstream students: the transfer of L1 structures, omissions, and processing problems (in this case, spelling mistakes). The distribution of errors is illustrated in Figure 4. This implies that in spite of

---

8 It should be noted that the percentages given demonstrate the prevalence of the error types within the two distinct groups, not within the whole data set.
having learned English in different ways, there are similarities in the students’
interlanguage, which further suggests that the acquisition and learning of at least
some grammatical features of English might indeed follow a developmental
pattern despite the type of instruction received (Krashen 1984; Selinker 1972).
Albeit intriguing, no reliable conclusions can be drawn from the translation data
alone because the CLIL students produced such a small number of errors in total.
Interestingly, though, the written composition errors exhibit a similar
distribution. This brings us to the next subsection.

![The translation of questions: the frequency of error types produced within the group](image)

**Figure 4:** The translation of questions: the frequency of error types within the
group.

### 6.3.3. The Types of Errors Produced: Written Composition

As anticipated, the written composition brought out a somewhat different set of
errors than the translation exercise because it allowed for more freedom of
expression. A total of 29 compositions were analysed: 14 by CLIL and 15 by
mainstream students. One composition was excluded from the analysis because a
mainstream student had misunderstood the instructions and written in Finnish.
All other compositions were included in the analysis because the CLIL students’
compositions were notably longer. This levelled out the small difference in
number.
Following the same principles as for the translation exercise (see section 6.3.2. and Figure 2), the error categorization was clear and straightforward for the most part. Nevertheless, a few difficult cases were encountered: not only did some sentences contain multiple errors, some fell easily into multiple categories or did not fit any given category. Eventually, some errors were placed into more than one category – for example, all errors from #49 through to #53 could be regarded as both overgeneralizations and ignorance of rule restrictions. This should be taken into account when examining the results.

The CLIL students again outperform the mainstream students in accuracy: the mainstream students produced a total of 113 errors while the CLIL students only 52 in their written compositions (see Figure 5). This is a remarkable difference, considering the fact that apart from the required ten words or expressions that had to be found in the composition (see Appendix 1), the students could more or less freely decide what to write, hence avoiding expressions that were too difficult for them. While this type of avoidance can, indeed, affect the results to some degree, a clear difference in productive accuracy can still be seen.

Figures 5 and 6 illustrate that the most common types of errors in the students’ compositions were omissions (mainstream students: 43 out of 113, 37.2%; CLIL students: 19 out of 52, 36.5%), processing problems (mainstream: 22 out of 113, 19.5%; CLIL: 20 out of 52, 38.5%), overgeneralizations (mainstream: 19 out of 113, 16.8%; CLIL: 7 out of 52, 13.5%) and false concepts hypothesized (mainstream: 13 out of 113, 11.5%; CLIL: 3 out of 52, 5.8%). Interestingly, the first two of these categories also stood out in the translation exercise. It should be noted though that a few errors are found in two of these categories (#89, #43, #70, #123, #129, #135), so not all the occurrences are individual cases.

The complete list of errors can be found in Appendix 2.
Figure 5: Written composition: the number of errors produced per category.

Figure 6: Written composition: the frequency (%) of error types produced within the group.
The most common errors were omissions (mainstream students: 37.2%; CLIL: 36.5% out of all errors produced). Both CLIL and mainstream students had omitted mainly articles, verb forms and prepositions. The omitted verb forms included mostly third person –s, past tense ending –ed, or irregular past tenses that had been replaced by the same verb in present tense. Even though the CLIL students’ compositions involved fewer omissions, the fact that both students groups produced similar errors suggests there is a general pattern of interlanguage development like Krashen (1984) and Selinker (1972) have proposed decades ago.

Processing problems represented the second-most common type of errors (mainstream students: 19.5%; CLIL: 38.5% out of all errors produced). Again, I decided to place all spelling mistakes in this category. Most of them seemed to stem from phonology: the students had guessed the written form of the word on the basis of its pronunciation. The students had, for example, used ‘wierd’ for ‘weird’ (#109, #154), ‘nowing’ for ‘knowing’ (#146), ‘fortunetly’ for ‘fortunately’ (#118), ‘allredy’ for ‘already’ (#157) and ‘scarryest’ for ‘scariest’ (#153). The fact that these types of errors were most common in the CLIL students’ written production (20 occurrences, 38.5% out of all the errors produced by the CLIL students) might actually reflect the differences between the two L2 learning contexts: the CLIL students acquire a substantial proportion of the L2 in (spoken) interaction, with less emphasis on accurate spelling. Consequently, they have to guess the correct spelling of words from the way they are pronounced. In fact, many of the CLIL students’ spelling mistakes resembled those produced by native speakers of English.

The third-most common errors were overgeneralizations (mainstream students: 16.8%; CLIL: 13.5% out of all errors produced). The mainstream students mostly added the past tense ending –ed to irregular verbs: out of the 19 overgeneralizations in total, 11 involved past tense endings. In terms of fluency, the overgeneralization of the past tense ending –ed is actually an effective communication strategy, because the student manages to deliver the meaning in spite of not knowing the accurate irregular verb form. There were also unnecessary prepositions for expressions of time, such as ‘twice on a week’ (#50) or ‘on this morning’ (#51). Other cases included a double negative in ‘I didn’t
want never’ (#48) and an incorrect superlative ending –est in ‘beautifullest’ (#49). The overgeneralizations produced by the CLIL students were more various and also fewer in number, including regular past tense endings for irregular verbs (#123 ‘hearted’ for ‘hurt’), a double past tense (#124 ‘didn’t die’ for ‘didn’t die’), a comparative ending –er for a long adjective (#127 ‘horribler’), a plural –s for an irregular plural (#128 ‘peoples’), and a third person singular –s (#129 ‘he haves’), which was also placed in the false concepts hypothesized category.

The fourth-most common category of errors was false concepts hypothesized (mainstream students 11.5%; CLIL 5.8% out of all errors produced), or the inability to understand a distinction in the target language. These included mainly incorrect objective and possessive pronouns as well as subject-verb disagreement like #94 “My parents doesn’t know”, #96 “Rosa are beautiful”, and #135 “There were lot of fire”. Errors in this category were much more common in the mainstream students’ compositions, once again an indication that the CLIL students have a better command of the structures of English when it comes to subject-verb agreement or the usage of prepositions, for example.

An error type not encountered in the translation of questions was category 1c. Interlingual errors, which referred to the direct translation of an idiomatic structure from Finnish into English. Examples include #33 “I see nightmares”, #36 “my friend said me crazy”, #121 “a dream what I have sean”, and #122 “the slowest scream what I’ve ever heard”. Both mainstream (2.6%) and CLIL students (3.8%) produced these errors, which offers an insight into their interlanguage: negative transfer is more common at the early stages of L2 acquisition.

The errors linked to communication strategies were particularly different from those produced in translation. On two occasions (#119, #120), a mainstream student had used a closely related word from the same semantic category as the targeted word, possibly due to not knowing or remembering the correct word. The CLIL students produced no errors belonging to this category. Neither did they produce any transfer of L1 structures nor structural failures.
It should be noted that all mainstream students’ errors (#49—53) under the *ignorance of rule restrictions* category were also classified as *overgeneralizations* due to the ambiguity of categorization described earlier. Thus, the aforementioned six errors in this category (5.3% out of all errors by the mainstream students) represent duplicate instead of individual cases. The CLIL students made only one error (1.9% out of all errors) in this category.

All in all, the CLIL students produced more accurate, colourful and idiomatic written compositions in English than the mainstream students. They made significantly fewer errors than the mainstream students. Moreover, a major part of their errors derived from phonology and were categorized as processing problems, which actually represent occasional mistakes that do not have a marked effect on the grammaticality or fluency of output.
7. Reliability and Validity

The present study relied on quantitative methods for the most part. However, there was also a qualitative ingredient as concerns the descriptive error analysis, in the sense that the classification depended primarily on my own interpretation of each error. Therefore, I chose to assess the reliability and validity of the study not only according to criteria presented for quantitative research but also with a brief focus on the qualitative nature of the error analysis.

7.1. The Quantitative Content

The quantitative content of the study included the statistical analysis of the students’ (n = 30) answers to the exercises in the English grammar test. In quantitative research, reliability refers to the extent to which the data collection method, such as the grammar test, would produce similar results if repeated (Metsämuuronen 2005: 64—65). Validity refers to the extent to which the conclusions made from the data are meaningful and appropriate – that is, the instrument and its content corresponds to what is being measured (Metsämuuronen 2005: 64—65).

The time and place of data collection might have affected the reliability of the test answers to some extent. The test was administered during normal school days and took about 60 minutes to complete. The CLIL class took the test on a morning lesson right before lunch while the mainstream class took the test in the afternoon, starting at 13:30. These conditions considered, some students might have been either hungry or tired by the time they reached the written composition (part 4). Nevertheless, the answers contained no signs of hastiness apart from one misinterpretation of instructions, where the student had written the composition in Finnish instead of English. The lack of incomplete and hasty responses implies that the students took the test seriously and generally tried their best.

To increase the validity of the test and to ensure the test corresponded to the students’ expected knowledge of English, I formulated the exercises on the basis of curricular contents (Töölö Primary School Curriculum, 2005 and 2012) and
past National English Examinations (valtakunnallinen englannin koe) for sixth graders. I also paid special attention to make the content of the exercises reflect the interests and experiences of students aged between 11 and 13.

A statistically significant positive correlation (\( r = .511—.631, p < 0.01 \); see Appendix 3) was found between the students’ past English marks and the test results, an indication that the test was successfully constructed to measure the students’ grammatical proficiency. The test also displayed high internal consistency: the intercorrelations for the test results from parts 1 (grammaticality judgments), 2 (cloze passage) and 3 (translation of questions) ranged between \( r = .531—.635 \) \((p < 0.01) \) (see Appendix 3, Table 10).

As for part 1 of the test, grammaticality judgments have been deemed a valid means of gaining information about the learners’ implicit knowledge of the target language (cf. Järvinen 1999). However, with the methods used it is impossible to completely distinguish whether the students resort to implicit or explicit knowledge or both when judging the grammaticality of sentences. DeKeyser (2003: 320, quoted in De Graaff & Housen 2009: 743) has, likewise, pointed out that no perfect tests exist for distinguishing the results of implicit and explicit learning or the availability of implicit or explicit knowledge. Therefore, it is worth noting that the results from the grammaticality judgments exercise (part 1) do not necessarily represent the students’ implicit knowledge alone but are a reflection of their overall English proficiency.

### 7.2. The Qualitative Content: Descriptive Error Analysis

One of the fundamental concerns of qualitative research is that of credibility: how well do the researcher’s interpretations of the data correspond with reality (Tuomi & Sarajärvi 2003: 136–137). Another fundamental concern is researcher bias: the analysis filters through the researcher’s perceptual lens, theories and beliefs (Maxwell 2005: 108). Both of these concerns need to be taken into account when assessing the reliability and validity of the descriptive error analysis I conducted on the translation of questions and the written composition.
The error analysis relied primarily on the description of the types and sources of errors by Ellis (1994: 58), Lott (1983) and Richards (1971), with some modifications and additions that stemmed from the present data (see Figure 2 in section 6.3.2.). However, the actual categorization process was based on my best interpretations about the source and nature of the errors encountered. Another researcher might have had different conclusions. The reliability of the error analysis would have increased by interviewing the students about the errors they had made, but this was impossible within the scope of this study. Regardless of these problems, however, the error analysis suited the diagnostic purposes of this study really well.

As for the validity of the data, researchers have observed that clinical elicitation in the form of a grammar test, for example, prompts different language use than naturalistic situations, and the results obtained might not reflect the students’ actual competence (Lococo 1976; Ellis 1995: 50). However, data collection in naturalistic situations would have been ill-suited for the purposes of this study: the goal was, indeed, to measure the students’ grammatical proficiency in written English. Also, language learners and native speakers alike are more likely to aim at correct language forms in situations that call for formal language use, such as written tests, so should they have any incorrect hypotheses, a formal grammar test would most presumably bring them out (Ellis 1995: 22). To obtain as comprehensive data as possible, the students’ grammatical proficiency was measured by four different exercises: grammaticality judgments, cloze passage (‘fill in the gaps’), translation, and written composition.

As for the generalizability of the findings, this is a case study in its purest essence. With only 30 respondents from two different classes, the purpose of the study was not to obtain generalizable data but to evaluate the language learning outcomes of the CLIL programme in Töölö Primary School. The results cannot be generalized to all CLIL contexts because of the differences in the starting age of CLIL, the linguistic aptitude and repertoire required from the students before starting the programme, the amount of foreign language content instruction, and the amount of formal and explicit English instruction given. The teachers’ methods and students’ abilities also vary, meaning that another study on different student groups might produce contradictory results even within the
context of Töölö Primary School. Therefore, the results are only approximate and provide grounds for further research. To obtain more generalizable data, longitudinal and more process-oriented research should be conducted. Unfortunately, this type of approach was not possible within the scope and aims of this study.
8. Discussion and Conclusion

The purpose of this study was to evaluate the learning of English grammar in a CLIL (Content and Language Integrated Learning) programme in Töölö Primary School, Helsinki, that combines content-based, implicit foreign language acquisition in content lessons with explicit foreign language instruction in formal English as an A1 lessons. CLIL was defined as “a dual-focused educational context in which an additional language, thus not usually the first language of the learners involved, is used as a medium in the teaching and learning of non-language content” for at least 25% of the time (Marsh & Hartiala 2001: 17). The research built on a comparison of the grammatical proficiency of English between 14 fifth-grade CLIL students and 16 sixth-grade mainstream students who receive only formal English instruction. The main aim was to analyse, describe and interpret what differences could be observed in the students’ grammatical proficiency and what types of grammatical errors the students produced in written English, both in translation and in a written composition. The data was be collected by means of an English language test and analysed mainly quantitatively, drawing on statistical methods and descriptive errors analysis.

8.1. The Main Findings

The findings on the learning outcomes of CLIL in Töölö Primary School were very promising. The CLIL students outperformed the mainstream students in all areas of grammatical proficiency that this study focused on: the implicit knowledge of English measured by grammaticality judgments, the structural knowledge of English measured by a cloze passage or ‘fill-in-the-gaps’ exercise, and productive skills in English measured by a translation exercise and a written composition.

Firstly, with a mean total of 8.07 out of 10.0 for the grammaticality judgments exercise, the CLIL students were found to possess a somewhat better implicit knowledge of English than the mainstream students who got a mean total of 7.44 out of 10.0. This finding is in line with Järvinen’s (1999: 136—137) study where she found that the implicit knowledge of the target language develops faster in CLIL than mainstream L2 instruction. Not only did the CLIL students discern
grammatically correct sentences more accurately but also more consistently: while the mainstream students’ results displayed greater variance and standard deviation, the CLIL students’ skills were more homogeneous. The most difficult sentence for both the CLIL and mainstream students was, rather surprisingly, one that tested how the students would place the word ‘enough’ in the sentence: most of them would have used a Finnish word order, most likely due to negative transfer. Other difficult structures were non-finite clauses and indirect speech. According to the Töölö Primary School curriculum for English as an A1 language (2005 and 2012), neither of these structures had been explicitly taught to the students before the data collection. Instead, the students could more easily judge the grammaticality of structures that were included in the curriculum contents and thus had been explicitly taught to them in formal EFL lessons. This finding supports the output hypothesis (Swain 1985; see also e.g. Meriläinen 2008: 29; Nikula & Marsh 1997: 90, 1999d: 80) and contrasts with the view that implicit and explicit knowledge of a foreign language are two distinct systems (Eubank & Beck 1995: 40; Krashen 1984; Truscott 1998: 119; VanPatten & Cadierno 1993; Wong 2004). The present finding suggests that mere input is not enough for the development of grammatical proficiency, but explicit instruction and opportunities for producing output are needed as well. Consequently, teachers need to keep on creating these opportunities in both content-based and instructed foreign language lessons.

The same tendency continued in the cloze passage exercise that measured the knowledge of some common structures of English in five broad categories: verbs, nouns, pronouns, adjectives and prepositions. The CLIL students possessed a better knowledge of all grammatical categories except for pronouns. However, the difference between the student groups’ mean totals for the pronoun part was statistically insignificant. The comparison of adjectives caused most problems for both the CLIL and mainstream students. Again, the standard deviation and variance of total means were lower for the CLIL group, an indication of more homogeneous grammatical proficiency.

The CLIL students also outperformed the mainstream students in terms of productive accuracy in both translation and written composition. As for the former, the CLIL students got higher mean totals for all five translations of
questions, indicating that they translate more accurately from Finnish into English than the mainstream students. In general, the mainstream students were more likely to omit constituents, use incorrect auxiliary verbs, or transfer Finnish structures into English. There are many possible explanations to these findings. First of all, and rather unsurprisingly, translation tasks have been found to induce more L1 transfer errors than other types of exercises like creative writing (Lococo 1976). The reason as to why only the mainstream students made more transfer errors in this translation exercise could be a tendency to first think in Finnish and then translate the phrase into English. Likewise, the explanation might lie in the difference between naturalistic and instructed ways of foreign language learning and the fact that the mainstream students in this study had simply received less English input and had fewer opportunities for producing English output at school than the CLIL group. The CLIL students might also find it easier to write idiomatic English due to having attended naturalistic, content-based English lessons. Besides, they are generally advised against thinking first in Finnish and then translating into English.

The groups’ mean totals in the translation of questions differed surprisingly little, considering the fact that the CLIL students had been exposed to a much greater amount of English throughout their school years. This could have been a sign of task familiarity as well as simplicity: translation from Finnish into English is a very common type of exercise in the instructed EFL lessons attended by the mainstream students. The small range of total scores for both of the groups also suggests the exercise might have been too easy. Then again, the CLIL students did produce significantly fewer errors (13 occurrences) in translation than the mainstream students (45 occurrences). The CLIL students’ errors fell into four categories: transfer of L1 structures (23.1%) i.e. Finnish word order, overgeneralizations (30.7%) of past tense, omissions (23.1%) of auxiliary verbs, and processing problems (23.1%), meaning mostly spelling mistakes. The last three of these categories also stood out in the written composition. The mainstream students’ errors mostly included omissions (26.7%) of either auxiliary verbs or other structural or semantic constituents, transfer of L1 structures (15.6%) i.e. Finnish word order, processing problems (15.6%) in spelling, and false concepts hypothesized (13.3%) i.e. mostly incorrect auxiliary verbs. Other error types produced by the mainstream students were
overgeneralizations (8.9%) of e.g. verb forms like the third person singular –s or past tense, structural failures (8.9%) i.e. incorrect word order, errors originating in communication strategies (6.6%), and ignorance of rule restrictions (4.4%).

The error analysis conducted for the written compositions, likewise, indicated that the CLIL students produce more accurate, colourful and idiomatic English than the mainstream students. These findings agree with Järvinen’s (1999) study where CLIL students outperformed mainstream students in both productive complexity and accuracy in the target language. The present data indicated that the CLIL students made significantly fewer errors than the mainstream students, who produced a total of 113 errors while the CLIL students only 52. The four most common error categories were the same for both student groups: omissions (mainstream 37.2%; CLIL 36.5%) of mostly articles, verb forms and prepositions, processing problems (mainstream 19.5%; CLIL 38.5%) in spelling, overgeneralizations (mainstream 16.8%; CLIL 13.5%) of e.g. prepositions and the past tense marker –ed added to irregular verbs, and false concepts hypothesized (mainstream 11.5%; CLIL 5.8%) in terms of e.g. incorrect objective and possessive pronouns and subject-verb disagreement. A particularly noteworthy error category not encountered in the translation data was interlingual errors, the direct translation of Finnish idiomatic structures into English and a clear indication of L1 transfer or interference in the students’ interlanguage. After all, the L1 plays a major role in L2 acquisition (Ellis 1995: 28—29).

A major part of the CLIL students’ processing problems i.e. spelling mistakes in the written composition derived from phonology. The students had tried to guess the written form of the word on the basis of pronunciation. These errors were occasional mistakes and did not have a notable effect on the grammaticality or comprehensibility of output. However, they actually manifested the most essential differences between CLIL and mainstream EFL learning contexts in Töölö Primary School. While the mainstream students focus more on accurate spelling and learn a substantial proportion of English from written texts in formal EFL lessons, the CLIL students also pick up a great number of words from daily spoken interaction in class and never see them written. Accordingly, some of the spelling mistakes resembled those made by native speakers of English who also learn the language in naturalistic situations.
Interestingly, three out of the four most common error categories the CLIL students produced in the translation of questions – transfer of L1 structures, omissions, and processing problems – were also the most common among the mainstream students. The same kind of tendency was seen in the written composition: the student groups had all of the four most frequent error categories in common, involving omissions, processing problems, overgeneralizations and false concepts hypothesized. This implies that in spite of having learned English in different ways, there are similarities in the students’ interlanguage, which further suggests that the acquisition and learning of at least some grammatical features of English might indeed follow a developmental pattern – or ‘Natural Order’ in Krashen’s terms (1984) – despite the type of instruction received (Krashen 1984; Selinker 1972).

In earlier research, instructed, mainstream foreign language learners have more or less systematically outperformed CLIL students in grammatical accuracy (e.g. Genesee 1987; Kokkila 1996; Nikula & Marsh 1997; Stanutz 1992; Swain 1996), albeit there have been a few contradictory findings (e.g. Järvinen 1999). Researchers have suggested that the explanation lies in the lack of formal and explicit target language instruction in CLIL programmes. The present data disagree with these earlier findings. The reason could well be that in Töölö Primary School, the CLIL students attend formal EFL lessons from the first through to sixth grade and receive explicit instruction on the forms of English. A multitude of findings on the benefits of explicit and formal L2 instruction support this view (DeKeyser 2003; Doughty 1991; Ellis 1990, 2002; Genesee 1987; Lightbown & Spada 1990; Long 1983; Norris and Ortega 2000; Pica 1983; White et al. 1991), along with studies focusing specifically on integrating explicit grammar instruction to content-based, meaning-focused approaches (Doughty & Varela 1998; Lyster 2004; Sheen 2005; Spada & Lightbown 1993, 2008). According to Spada & Lightbown (2008: 184), today there is a growing consensus that L2 instruction is most effective when attention is directed to both form and meaning (see also Nikula & Marsh 1997: 90—91, 113). The present findings indicate, likewise, that explicit instruction combined with natural foreign language acquisition in the CLIL setting leads to successful learning outcomes and improves grammatical accuracy. The CLIL students also have more
opportunities to produce meaningful L2 output at school than the mainstream students – yet another path to successful L2 learning (Meriläinen 2008: 29; Nikula & Marsh 1997: 90, 1999d: 80; Swain 1985).

There are alternative explanations to the findings as well. Fact is that the CLIL students’ language aptitude is good to start with: they need to pass a language aptitude test to be admitted to the CLIL programme in Töölö Primary School (see section 3.6.). Due to an increasing number of applicants, today the children need to obtain relatively high test scores to fall in the accepted range. Therefore, an intriguing question is to what extent the present findings represent the innate language aptitude of the CLIL students, and what proportion of the successful learning outcomes of the CLIL students owes specifically to the CLIL programme followed in Töölö Primary School.

The lower variation and standard deviation in the CLIL students’ test results were, indeed, a clear indication of more homogenous English skills. But whether these skills are the outcome of a successful CLIL programme or the manifestation of better language aptitude is, unfortunately, a question beyond the scope of this study, and further research is needed. However, a more important question is if the CLIL and mainstream groups can even be compared to each other: did they stand on the same line or did the CLIL students have a clear advantage?

First of all, the CLIL students’ previous English marks were good to start with – there were no marks below 8, denoting ‘good knowledge’ overall (see Table 4, section 5.3.). Nevertheless, the mainstream group did have more students with top marks, and their mean total was a bit better (8.88) than the CLIL group’s mean total of previous English marks (8.79). I know personally that the student assessment these marks represent was very much based on the same English exams, because the groups had used the same book series and studied the same grammatical contents prior to data collection. In this respect, the groups were not too far apart for a reliable comparison to be made, except for the fact that the CLIL students receive more English input in the content lessons on a daily basis, of course.
8.2. Implications for Teaching and Further Research

Despite the fact that this was a case study and the results are generalizable only if similar conditions apply, some suggestions can still be made. The data demonstrated that a CLIL programme that combines content-based foreign language acquisition and formal, instructed foreign language learning can have very successful learning outcomes. These findings find wide support in earlier research. On this account, this type of model can be utilized when it is time to improve current CLIL programmes or develop completely new ones. CLIL teachers can also include formal, explicit foreign language instruction in the content lessons as much as they feel is necessary for successful learning, especially if separate foreign language lessons cannot be implemented in the curriculum or in the students’ schedule.

Further research to examine successful CLIL practises is needed. To reliably evaluate the effects of innate language aptitude on not only the present but also past findings, further research should be conducted in CLIL settings that, firstly, involve both instructed and content-based target language lessons, and secondly, offer the CLIL programme for all students regardless of language aptitude. In addition, comparative research on the target language learning outcomes in CLIL settings with and without formal instruction in the language would also be salient and very much needed for the development of better CLIL programmes. Further studies should, naturally, have more respondents. All in all, longitudinal research is needed that focuses on foreign language learning as a process and not merely as an outcome, so as to gain more in-depth knowledge about the language acquisition process in CLIL settings.
References


University of Turku. Faculty of Education. Research Reports A: 168.
Appendix 1: The Grammar Test

Hyvä vastaaja,


Erittäin suuri kiitos avustasi! 😊

Laura

Vastaa ensin muutamiin taustakysymyksiin.

Syntymävuosi: ______
Luokka: ______

Sukupuoli (merkitse rastilla): [ ] tytö [ ] poika

Milloin aloitit englannin opiskelun? (merkitse rastilla)

[ ] päiväkodissa/esikoulussa
[ ] 1. luokalla
[ ] 2. luokalla
[ ] 3. luokalla
[ ] 4. luokalla
[ ] 5. luokalla

Englannin arvosana viime syysyn väliarvioinnissa: ______

Mikä on äidinkielesi? __________________

Opiskelen suomea toisena kielenä. [ ] kyllä [ ] ei

Englanti on toinen äidinkieleeni/kotikieleni. [ ] kyllä [ ] ei
Tehtävä 1. Virheellisten lauseiden tunnistaminen.


1.
   a. I watched **a film** that you told me about last Saturday.
   b. I watched **the film** that you told me about last Saturday.
   c. I watched **films** that you told me about last Saturday.

2.
   a. I think the film was awesome but my sister **don't like** it.
   b. I think the film was awesome but my sister **didn't like** it.
   c. I think the film was awesome but my sister **didn't likes** it.

3.
   a. **Her** taste in films is just too different from mine.
   b. **His** taste in films is just too different from mine.
   c. **Our** taste in films is just too different from mine.

4.
   a. In the film **were many frightening scenes**.
   b. **There were many frightening scenes** in the film.
   c. **Many frightening scenes were** in the film.

5.
   a. Yesterday I **came** to school by bike.
   b. Yesterday I **come** to school by bike.
   c. Yesterday I **am coming** to school by bike.

6.
   a. My bike is broken but my **dad said** that **he won't buy me** a new one before summer.
   b. My bike is broken but my **dad says** that **he won't buy myself** a new one before summer.
   c. My bike is broken but my **dad say** that **he won't buy me** a new one before summer.

7.
   a. I don't know **why did he say** that.
   b. I don't know **why he saying** that.
   c. I don't know **why he said** that.
8.
a. Maybe because **he hasn't got enough money**.
b. Maybe because **he doesn't have money enough**.
c. Maybe because **he haven't got enough money**.

9.
a. On my way to school I saw a **dog walking** in the street and **jumping on people**.
b. On my way to school I saw a **dog who walking** in the street and **jumping on people**.
c. On my way to school I saw a **dog who walked** in the street and **jumping on people**.

10.
a. I was **late from school** because I stopped to watch **him**.
b. I was **late from school** because I stopped to watch **his**.
c. I was **late to school** because I stopped to watch **himself**.

**Total: ___ / 10p**
Man’s best friend

________________(1) people love animals and many families have a pet. One of the most popular pets is a dog. We can find dogs all over the world ______(2) Alaska ______(2) Australia.

Wild dogs __________ (3) first like wolves but _________________(4) years ago man and dog __________ (5) friends. The close friendship _______________ (6) man and dog is long.

Though dogs like meat __________ (7) they can eat almost anything, for example fruit and ________________(8). They can taste, hear and smell in much the same way humans can, but we humans are __________ (9) than dogs in these things. It is no wonder that humans __________(10) to take a dog into __________ (11) homes.

Dogs can help us in many ways. They can help ____________ (12) find lost __________ (13) through snow and darkness. They can also help the blind walk safely in traffic. Dogs are great company for __________(14) and old people. They even visit ____ (15) old people’s homes today.

A dog can be ____________ (16) family member.

Vastausvaihtoehdot:

1. a) more b) much c) most
2. a) up Alaska to Australia b) from Alaska to Australia c) in Alaska down Australia
3. a) where b) was c) were
4. a) thousands of b) thousands in c) thousands on
5. a) come b) become c) became
6. a) beside b) between c) below
7. a) better b) good c) best
8. a) berries b) berrys c) berry’s
9. a) worst b) worse c) bad
10. a) wants b) wanting c) wanted
11. a) them b) their c) there
12. a) policemans b) policemen’s c) policemen
13. a) people b) peoples c) people’s
14. a) childs b) children c) childrens
15. a) to b) for c) –
16. a) nicer b) nicest c) the nicest
It ______________(17) to be alone. It likes to belong to a group and loves its family a lot.

After reading the following story ______(18) can realise why there is a saying: "the dog is man’s best friend”.

A Japanese dog died in March, 1935. It was then the most famous dog in Japan. Almost _____________(19) knew it because the dog was waiting for ______(20) master coming back from work at the railway station for 10 years. The master never came. The dog didn’t know that the master was dead.

| 17. | a) don’t like  b) doesn’t like  c) didn’t like |
| 18. | a) us  b) you  c) man |
| 19. | a) nobody  b) anybody  c) everybody |
| 20. | a) its  b) it’s  c) its’ |

For the teacher:
Nouns: ___/4p
Verbs: ___/4p
Adjectives: ___/4p
Pronouns: ___/4p
Prepositions: ___/4p
Total: ___ /20p

Olet toimittajana juorulehdessä ja sinun täytyy haastatella Suomessa vieraillevaa elokuvatähteä. Kirjoita kysymykset valmiiksi.

1. Järjestä annetut sanat kysymykseksi.
   Missä sinä asut nyt? (do – live – you – now – where)

   ____________________________________________________________
   ___ / 1p

2. Kirjoita kysymys englanniksi. Suluissa on apusanoja, mutta sinun pitää lisätä kysymykseen myös muita sanoja.
   Osaatko puhua espanjaa hyvin? (can – Spanish – well)

   ____________________________________________________________
   ___ / 2p

   Mitä mieltä äitisi on tästä elokuvasta? (think about)

   ____________________________________________________________
   ___ / 2p

4. Kirjoita kysymys englanniksi.
   Missä sinä vietit eilisillän?

   ____________________________________________________________
   ___ / 2p

5. Kirjoita kysymys englanniksi.
   Oliko tämä sinun viimeinen elokuvasi?

   ____________________________________________________________
   ___ / 3p

Total: ___ / 10p

89

Kirjoita tarina jostakin tapahtumasta englanniksi. Tarinan ei tarvitse olla totta. Kirjoituksen pitää olla vähintään 40 sanaa pitkä.

Käytä tarinassasi kahdeksaa alla olevaa sanaa/ilmausta siten, että käytät kahta sanaa/ilmausta kustakin laatikosta:

<table>
<thead>
<tr>
<th>hirviö</th>
<th>kauheampi</th>
<th>juoksi</th>
<th>tänä aamuna</th>
</tr>
</thead>
<tbody>
<tr>
<td>käärme</td>
<td>nopein</td>
<td>kuuli</td>
<td>eilen</td>
</tr>
<tr>
<td>opettaja</td>
<td>jännittävin</td>
<td>näki</td>
<td>kolme päivää sitten</td>
</tr>
<tr>
<td>koira</td>
<td>kauniimpi</td>
<td>söi</td>
<td>viime keskiviikkona</td>
</tr>
<tr>
<td>paras ystäväni</td>
<td>hitain</td>
<td>heräsi</td>
<td>kaksi vuotta sitten</td>
</tr>
</tbody>
</table>

Esimerkkilauseet auttavat, jos et keksi, kuinka aloittaisit tarinan:

- *Roosa heräsi kuudelta, koska hän kuuli jotakin outoa. Roosa käveli ikkunan luo ja katsoi ulos. Hän näki kolme isoaa...*
- *Eilinen oli elämäni jännittävin päivä. Sain kirjeen...*
Valmista on! Tarkista vielä vastauksesi ja palauta sitten paperi Lauralle 😊
Appendix 2: The errors produced in the written composition

Error types 6B (number and % within class 6B; total n=113)

1b. Transfer of L1 structure (3; 2.6%, n=113):
33. I see nightmares (2) from Finnish ‘näen painajaisia’
36. My friend said me crazy (1) from Finnish ‘sanoi minua hulluksi’

1c. Interlingual errors (2; 1.8%, n=113):
34. my mother – – he (1) using ‘he’ where ‘she’ is required; Finnish makes no distinction between genders, ‘hän’ is used for both males and females
35. She – – his dog (1) mixing ‘her’ and ‘his’; Finnish ‘hän’ makes no distinction between genders

2a. Overgeneralizations (19; 16.8%, n=113):
37. drived (1) overgeneralization of past tense –ed
38. thinked (1) overgeneralization of past tense –ed
39. caught (1) overgeneralization of past tense –ed
40. I waked up (1) overgeneralization of past tense –ed
41. Dad saied (1) overgeneralization of past tense –ed
42. He goed (1) overgeneralization of past tense –ed
43. I telled her (1) overgeneralization of past tense –ed
44. falled (1) overgeneralization of past tense –ed
45. blewed up (1) overgeneralization of past tense –ed
46. founded (from find, found) (2) overgeneralization of past tense –ed
47. he went played (1) overgeneralization of past tense –ed
48. I didn’t want never to clean it (1) overgeneralization of negative form (double neg.)
49. beautifullest (1) overgeneralization of superlat. ending –est
50. twice on a week (1) overgeneralization of preposition ‘on’
51. on this morning (2) overgeneralization of preposition ‘on’
52. She came back on yeasterday (1) overgeneralization of preposition ‘on’
53. at this morning (1) overgeneralization of preposition ‘at’

2b. Ignorance of rule restrictions (6; 5.3%, n=113):
49. Beautifullest (1) superlative ending –est not used here
50. twice on a week (1) preposition ‘on’ not needed
51. on this morning (2) preposition ‘on’ not needed
52. She came back on yesterday (1) preposition ‘on’ not needed
53. at this morning (1) preposition ‘at’ not needed

2c. Omissions (42; 37.2%, n=113):
54. slowest (1) omission of definite article ‘the’
55. fastest (2) omission of definite article ‘the’
56. wizard was old (1) omission of definite article ‘the’
57. boys saw the unexpected quest (1) omission of definite article ‘the’
58. Are they go to dog show? (1) omission of definite article ‘the’
59. boys were tired (1)  omission of definite article ‘the’
60. She go school (1)  omission of definite article ‘the’
61. Pekka watched movie (1)  omission of indefinite article ‘a’
62. I have postcard in my mailbox (1)  omission of indefinite article ‘a’
63. it’s from snake (1)  omission of indefinite article ‘a’
64. true monster (1)  omission of indefinite article ‘a’
65. It was like snake (1)  omission of indefinite article ‘a’
66. logo of Finnish Kennel club (1)  omission of indefinite article ‘a’
67. Pekka kill spiders (1)  omission of 3rd person singular –s
68. Pekka go to sleep (1)  omission of 3rd person singular –s
69. Rosa wake up (1)  omission of 3rd person singular –s
70. She go school (1)  omission of 3rd person singular –s
71. She eat yoghurt (1)  omission of 3rd person singular –s
72. After school she go (1)  omission of 3rd person singular –s
73. Rosa go home (1)  omission of 3rd person singular –s
74. she go she best fried (1)  omission of 3rd person singular –s
75. Roosa walked ---- and watch (1)  omission of past tense –ed
76. Roosa seem scared (1)  omission of past tense –ed
77. She hug her dog (1)  omission of past tense –ed
78. Roosa watch the monster (1)  omission of past tense –ed
79. Roosa smile (1)  omission of past tense –ed
80. I miss you! (1)  omission of past tense –ed
81. Roosa take their cat (1)  omission of irregular past tense (‘took’)
82. I go to school and saw (1)  omission of irregular past tense (‘went’)
83. I feel very sad (1)  omission of irregular past tense (‘felt’)
84. he ate banana and drink coffee (1)  omission of irregular past tense (‘drank’)
85. What they do here? (1)  omission of auxiliary verb ‘do’
86. I get up six (1)  omission of preposition ‘at’
70. She go school (1)  omission of preposition ‘to’
87. He woke up five o’clock (1)  omission of preposition ‘at’
88. he started run away (1)  omission of preposition ‘to’
43. I telled her my monster (1)  omission of preposition ‘about’
89. she go she best fried (1)  omission of preposition+verb structure ‘to+see’
90. She saw three big monster (1)  omission of plural –s
91. he ate pasta and lots of fruit (1)  omission of plural –s
92. Are they go to dog show? (1)  omission of present continuous –ing

2d. False concepts hypothesized (13; 11.5%, n=113)
93. on the garden (2)  using ‘on’ where ‘in’ is required
94. My parents doesn’t now (1)  using ‘doesn’t’ where ‘don’t’ is required
95. They names is.. (1)  using ‘is’ where ‘are’ is required
96. Rosa are beautiful (1)  using ‘are’ where ‘is’ is required
97. He teacher are he dad (1)  using ‘are’ where ‘is’ is required
97. He teacher are he dad (1)  using ‘he’ where ‘her’ is required
98. I saw they (1)  using ‘they’ where ‘them’ is required
95. They names is... (1)  using ‘they’ where ‘their’ is required
89. she best fried (1)  using ‘she’ where ‘her’ is required
99. His slower than (1)  using ‘his’ where ‘he’s’ is required
100. His so ugly (1)  using ‘his’ where ‘he’s’ is required
101. snakes are worst than dogs (1)  using ‘worst’ where ‘worse’ is required
2e. Structural failures (4; 3.5%, n=113)
47. he went played (1) incorrect past continuous structure
48. I didn’t want never to clean it (1) double negative
89. After school she go she best fried (1) failed to develop a structure fully, constituents missing
102. Last day is my life interesting day (1) wrong word order, constituents missing

3. Processing problems: phonology-based mistakes (22; 19.5%, n=113)
103. cloc is eight (1) misspelled ‘clock’
104. Mikes neighbor (1) misspelled genitive form, apostrophe missing
105. pageage (1) misspelled ‘baggage’
106. exiting (1) misspelled ‘exciting’
107. laught (2) misspelled laughed
52. She came back on yeasterday (1) misspelled yesterday
108. beatiful (1) misspelled ‘beautiful’
109. wierd (1) misspelled ‘weird’
110. didn’t (3) misspelled ‘didn’t’
111. supraised (1) misspelled ‘surprised’
112. togheter (1) misspelled ‘together’
113. now (2) misspelled ‘know’
114. whit (2) misspelled ‘with’
115. along (1) misspelled ‘alone’
116. caime (1) misspelled ‘came’
117. chicken (1) misspelled ‘chicken’
118. fortunately (1) misspelled ‘fortunately’

4. Communication strategies (2; 1.8%, n=113)
119. She sounded something. (1) using ‘sounded’ for ‘heard’
120. Let’s seem your monster! (1) using ‘seem’ for ‘see’

(in both occasions, using a closely related word from the same semantic category as the targeted word)

Error types 5C (number and % within class 5C; total n=52)

1c. Interlingual errors (2; 3.8%, n=52)
121. a dream what I have sean (1) using ‘what’ for ‘that’; no distinction in Finnish
122. the slowest scream what I've ever seen (1) using ‘what’ for ‘that’; no distinction in Finnish

2a. Overgeneralizations (7; 13.5%, n=52)
123. my head hearted (1) overgeneralization of past tense –ed (for ‘hurt’)
124. he didn’t died (1) overgeneralization of past tense
125. in this morning (1) overgeneralization of preposition ‘in’
126. his a dog (1) overgeneralization indefinite article ‘a’
127. *horribler* (1) overgeneralization of comparative ending –er
128. *peoples* (1) overgeneralization of plural –s
129. he *haves* really blue eyes (1) overgeneralization of 3rd person sg. –s

2b. Ignorance of rule restrictions (1; 1.9%, n=52)
130. teacher tell to her friends (1) no need for ‘to’ here

2c. Omissions (19; 36.5%, n=52)
131. I’m *orphan* (1) omission of indefinite article ‘a’
132. Nella had *nice life* (1) omission of indefinite article ‘a’
133. *teatcher* (3) omission of indefinite article ‘a’
134. *dog is faster* (1) omission of indefinite article ‘a’
135. there were *lot of fire* (1) omission of indefinite article ‘a’
136. *fastest* (3) omission of definite article ‘the’
137. *slowest* (1) omission of definite article ‘the’
138. *snake* (1) omission of definite article ‘the’
139. *start screaming* (2) omission of past tense ending –ed
140. snake *eat that* (1) omission of irregular past tense (‘ate’)
141. it *bite him* (1) omission of irregular past tense (‘bit’)
142. he *get medicine* (1) omission of irregular past tense (‘got’)
143. my best friend *feed the dog* (1) omission of irregular past tense (‘fed’)
144. my *neighbor dog* (1) omission of genitive marker ‘s

2d. False concepts hypothesized (3; 5.8%, n=52)
129. he *haves* really blue eyes (1) using ‘haves’ where ‘has’ is required
135. there *were lot of fire* (1) using ‘were’ where ‘was’ is required
145. coming *in a door* (1) using ‘in’ where ‘through’ is required

3. Processing problems: phonology-based mistakes (20; 38.5%, n=52)
146. without *nowing* (1) misspelled ‘knowing’
148. trying to *come me down* (1) misspelled ‘calm me down’
149. *rememberd* (1) misspelled ‘remembered’
150. Nella *past away* (1) misspelled ‘passed’
151. *teatcher* (3) misspelled ‘teacher’
152. *herd* (1) misspelled ‘heard’
153. *scarryest* (1) misspelled ‘scariest’
154. *wierd* (1) misspelled ‘weird’
155. *yeasterday* (1) misspelled ‘yesterday’
156. *exiting* (1) misspelled ‘exciting’
157. *already* (1) misspelled ‘already’
158. *prettyer* (1) misspelled ‘prettier’
159. *frend* (4) misspelled ‘friend’
121. a dream what I have *sean* (1) misspelled ‘seen’
123. my head *hearted* (1) misspelled ‘hurt’
## Appendix 3: Tables

### Table 10: Correlations between the students’ previous English marks and the test results.

Correlations

<table>
<thead>
<tr>
<th></th>
<th>English mark in previous certificate</th>
<th>TOTAL SCORE: GRAMMATICALITY JUDGMENTS</th>
<th>TOTAL SCORE: CLOZE PASSAGE</th>
<th>TOTAL SCORE: TRANSLATION OF QUESTIONS</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>English mark in previous certificate</strong></td>
<td>Pearson</td>
<td>1</td>
<td>.631**</td>
<td>.537**</td>
</tr>
<tr>
<td></td>
<td>Correlation</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Sig. (2-tailed)</td>
<td>.000</td>
<td>.002</td>
<td>.004</td>
</tr>
<tr>
<td><strong>TOTAL SCORE: GRAMMATICALITY JUDGMENTS</strong></td>
<td>Pearson</td>
<td>.631**</td>
<td>1</td>
<td>.531**</td>
</tr>
<tr>
<td></td>
<td>Correlation</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Sig. (2-tailed)</td>
<td>.000</td>
<td>.003</td>
<td>.000</td>
</tr>
<tr>
<td><strong>TOTAL SCORE: CLOZE PASSAGE</strong></td>
<td>Pearson</td>
<td>.537**</td>
<td>.531**</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>Correlation</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Sig. (2-tailed)</td>
<td>.002</td>
<td>.003</td>
<td>.000</td>
</tr>
<tr>
<td><strong>TOTAL SCORE: TRANSLATION OF QUESTIONS</strong></td>
<td>Pearson</td>
<td>.511**</td>
<td>.635**</td>
<td>.606**</td>
</tr>
<tr>
<td></td>
<td>Correlation</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Sig. (2-tailed)</td>
<td>.004</td>
<td>.000</td>
<td>.000</td>
</tr>
<tr>
<td><strong>N</strong></td>
<td></td>
<td>30</td>
<td>30</td>
<td>30</td>
</tr>
</tbody>
</table>

**. Correlation is significant at the 0.01 level (2-tailed).}