

Vocabulary Testing in CLIL

The Effect of Incidental Vocabulary Learning in CLIL on the Vocabulary of Learners

Nicholas Kristian Austad



Mastergradsavhandling ved Institutt for Lærerutdanning
og Skoleutvikling, Engelsk Fagdidaktikk

UNIVERSITETET I OSLO

Vår 2013

Vocabulary Testing in CLIL

The Effect of Incidental Vocabulary Learning in CLIL on the Vocabulary of Learners

© Nicholas Kristian Austad

2013

Vocabulary Testing in CLIL: The Effect of Incidental Vocabulary Learning on the
Vocabulary of Learners

Nicholas Kristian Austad

<http://www.duo.uio.no/>

Trykk: Reprosentralen, Universitetet i Oslo

IV

Abstract

The present study examines whether students in Content and Language Integrated Learning (CLIL) classes will score better than students receiving only normal EFL classes on vocabulary tests. The study uses a mixed methods design to inspect the vocabulary scores of the students. Included in this mixed-methods design are interviews with the CLIL teachers and CLIL students, a survey and two vocabulary tests, one testing vocabulary breadth, and the other testing productive vocabulary. The sample consisted of four classes from two schools, two CLIL classes, and two EFL classes. The classes were all from the tenth grade, in lower secondary school.

The breadth test used was an X-lex test, which uses vocabulary from the 5000 most frequent English words and an additional group of false words to avoid overconfidence in the test-taker. The other vocabulary test consisted of an analysis of 200 word samples from the students looking at type/token ratios. The survey and interviews were used to supplement the vocabulary tests. They were designed to help explain the results of the vocabulary tests, and give a better picture of why the scores came out the way they did.

The findings of this study showed that the EFL students scored better on the vocabulary tests than the CLIL students. The survey and interviews managed to reveal some possible explanations to why the CLIL students had not scored better than the EFL students. The small amount of English that was used in the CLIL classes coupled with using Norwegian textbooks helped explain why the CLIL students had done worse than the EFL students. In addition to this, the grades of the EFL students in English were higher, which could also help explain the higher scores of the EFL students.

In the discussion I argue that the language portion of CLIL must become more integrated into the subject. Currently the language is seen as a barrier rather than a goal in CLIL. Most importantly I argue that if CLIL is to have any effect on the vocabulary of students both the quality and the quantity of vocabulary related tasks must go up. Also, I strongly suggest the use of good English texts and textbooks, as reading is one of the best and simplest forms of vocabulary learning in CLIL.

Acknowledgements

I would like to start with thanking my supervisor, Glenn Ole Hellekjær. His tireless work to help me was much appreciated, always having good advice when things seemed impossible. I would also like to thank my friends, for bothering to put up with me and being wonderfully distracting. A special mention has to go to Mari for your delightful company. Most of all I would like to thank my family, this goes out to all of you, for being such wonderful people, and making me push for more each day. The biggest thanks goes out to my immediate family, my brother, my father, and my mother. You have always made me want to do my very best, I can't thank you all enough for having faith in me, and encouraging me when things are difficult. This has been a challenging year, and without you I would never have made it this far. Most of all, I would like to thank my mother, you are with me every day, and you are the reason I am where I am today.

Index

1) Introduction	1
Personal Anecdote.....	1
Introduction to the Chapter	2
Goal of the Study.....	2
The Research Question.....	3
Overview of the Thesis	4
2) Content and Language Integrated Learning.....	5
Introduction	5
Content and Language Integrated Learning	5
The Roots of Content and Language Integrated Learning	7
Content and Language Integrated Learning in Norway and Scandinavia.....	8
2.1.1 CLIL in Norway	9
Research on Vocubular in CLIL.....	11
Summary of the Chapter	12
3) Vocabulary Knowledge and Vocabulary in the LK06.....	13
Introduction	13
The Nature of a Word.....	13
3.1.1 Multi-Word Items.....	14
Word Knowledge	15
The Most Important Words in the English Language	16
Summary of Work Knowledge and Word Learning	17
The English Subject and the LK06	19
3.1.2 The Basic Skills and Vocabulary	20
The English Subject Curriculum and Vocabulary.....	20
3.1.3 Competency Aims in the Subject Curriculum Relating to Vocabulary	22
Discussion on the Place of Vocabulary in the Subject Curriculum	23
4) Vocabulary Learning.....	25
Introduction	25
Incidental Vocabulary Learning.....	25
Paribakht and Wesche’s Study on Incidental Vocabulary Through Reading	29
The Involvement Load Hypothesis	30

4.1.1	Laufer and Hulstijns Research on the Involvement Load Hypothesis	32
	Hill and Laufer’s Study on the Involvement Load.....	33
	Folse’s Study on Involvement Load and Repetition	34
	Lee and Hirsh’s Study on Involvement Load.....	35
	Summary of Chapter	37
5)	Method	38
	Introduction	38
	Defining the Research Question.....	38
	Selecting the Sample	39
5.1.1	Details Regarding the Sample	39
	The Research Design.....	41
5.1.2	Strengths and Weaknesses of a Mixed-Methods Approach.....	41
	Testing Vocabulary	43
5.1.3	Breadth Testing	43
5.1.4	Depth Testing	44
5.1.5	Receptive and Productive Vocabulary	45
	The Vocabulary Tests.....	45
5.1.6	The X-lex Test.....	46
5.1.7	The Text Analysis	48
	The Interviews.....	49
5.1.8	The Teacher Interviews	50
5.1.9	The Student Interviews.....	51
	The Survey	52
	Collecting and Processing the Data.....	53
5.1.10	Processing and Interpreting the data	54
	External Validity	54
5.1.11	Representativeness of the Sample	55
5.1.12	Non-respondents.....	56
5.1.13	Conclusion of External Validity Discussion	56
6)	Results and Analysis	58
	Introduction	58
	The Sample.....	58
	X-lex Results	59

Difference Between Classes.....	61
Text Results.....	62
The Survey Questions	64
6.1.1 CLIL Students	66
6.1.2 EFL Students	66
6.1.3 Comparison of Answers to Unknown Words	67
6.1.4 Summary of Survey Questions about Unknown Words	69
Analysis of Quantitative Data	69
6.1.5 Analysis of Strategies to Deal with Unknown Words.....	70
Summary of Data Analysis	74
Introduction to Interview Section.....	75
Interview with CLIL Teacher A, General Section	75
6.1.6 Interview with CLIL Teacher A, Vocabulary Section.....	77
6.1.7 Interview with CLIL Teacher A, Difficulties and Thoughts.....	78
Interview With CLIL Teacher B, General Section	78
6.1.8 Interview with CLIL Teacher B, Vocabulary Section	80
6.1.9 Interview with CLIL Teacher B, Difficulties and Thoughts.....	80
Comparison of the Teacher Interviews	81
Interviews with CLIL Students	82
6.1.10 Interview with CLIL Students from School A.....	82
6.1.11 Interview with CLIL Students from School B	83
Summary of Interviews	85
Overall Summary of Results	85
7) Discussion	87
Introduction	87
Relevant Findings from the Study.....	87
The Vocabulary Test Findings	88
Findings from the Survey and Interviews	89
Theoretical Discussion of Findings.....	92
7.1.1 Overview of Theory	92
7.1.2 The Quantity Aspect.....	93
7.1.3 The Quality Aspect.....	94
7.1.4 Conclusion to the Theoretical Discussion.....	97

Studies on Incidental Vocabulary Learning	98
7.1.5 Studies on CLIL	98
7.1.6 Comparison with the Present Study	99
Validity of the Study	100
Summary of the Chapter	102
8) Conclusion.....	103
Implications of the Study	103
Further Research	105
Concluding Remarks	106
Litteraturliste	107
Appendices	110
Figure 3.1 The Vicious Circle of Reading	17
Figure 4.1 The road from ambient speech to automated output.....	27
Fig 4.2 VKS Elicitation Scale	29
Fig 4.3 Laufer & Hulstijn's Three components of Involvement	31
Fig 4.4 Comparison of exercises in Lee & Hirsch's study	35
Table 5.1 Total Numbers of Students who have Completed Tests	40
Table 5.2 Comparison of breadth and depth tests	46
Table 6.1 Total Numbers of students who have completed tests	58
Table 6.2 X-lex scores for CLIL and EFL students	59
<i>Table 6.3 X-lex Scores without Students with 5 or more False Words</i>	60
Table 6.4 X-Lex Scores Sorted by Individual Classes.....	61
Table 6.5 200 Word Text Analysis using RANGE taken from written texts of the CLIL and EFL students.....	63
Table 6.6 200 word text analysis using Range on the students from school B only.....	63
Table 6.7 Comparison of Answers to Questions Regarding Unknown Words Between EFL and CLIL Students N=73	65
Table 6.8 Correlation of How often do you meet unknown words with the other questions in the survey for the students in the sample who completed both the X-lex vocabulary test and Survey, N=71	71
Table 6.11 Correlation of Grades with the answers to questions regarding unknown words for students in the sample who completed both the X-lex vocabulary test and survey N=71	73
Table 7.1 Important Findings from the Survey and Interviews	90

1) Introduction

Personal Anecdote

My relationship with the English language has always been a close one, with my mother being Canadian I grew up in a house that was filled with both Norwegian and English. I learned both languages equally, and learned to read and write in both English and Norwegian. If words failed me in one language I could easily shift to the other, often mid-sentence. Throughout my childhood we would travel to Canada to visit family. It was here that I first encountered immersion programs. My cousins all went to immersion schools, learning French from a very young age. Indeed, all of them had learned an impressive amount of French, especially compared to my own paltry vocabulary in the language after a few years in lower secondary school in Norway. I was impressed, but it would be many years before I heard of anything like it in Norway. It was during my studies at the University of Oslo that I first heard of Content and Language Integrated Learning (CLIL), the teaching of a foreign language in other subjects such as Mathematics or Social Sciences. It was a relatively new experiment, and there was little information about it, but from the start it excited me. Having seen the rapid development of language in my cousins due to immersion programs I thought that CLIL had an incredible potential to help improve the English of Norwegian students. Also the possibility of learning subject specific vocabulary was something I had always felt that I could have used in my own education, since I felt I lacked the specific terminology needed to discuss certain topics in English.

Introduction to the Chapter

This study was therefore a golden opportunity for me to further explore the prospects of bilingual teaching, as I got the opportunity to work with some of the leading people in Norway on CLIL, including my supervisor Glenn Ole Hellekjær. In the study I was able to work with two lower secondary schools in Norway who were already involved in a larger study on CLIL as part of a Comenius project. Their enthusiasm and engagement to the project helped immensely and allowed me to investigate the subject. CLIL is still in its infancy in Norway, and there are few organized projects at the present time, which meant that it was hard to come by schools who were willing to give up valuable classroom time for my tests. Therefore the schools participating in the study were a godsend, not only did they allow ample opportunity to test, but both schools were from the same area and had the same grade levels doing CLIL instruction. The schools had CLIL classes in the final year of lower secondary school, meaning the students were in grade 10, with the age of the students being approximately 15.

This study will be testing the vocabulary of students in Norway who are in Content and Language Integrated Learning classes. I will be investigating if these CLIL students have obtained a better vocabulary than that of their peers who received EFL instruction only. First I will be presenting the goals of this study and then the research question. After this I will be giving a quick overview of the thesis.

Goal of the Study

The purpose of this study is to look at how CLIL can influence the English of Norwegian learners. I firmly believe that one of the most important factors when learning a language is vocabulary. To obtain a vocabulary that allows for widespread use of English concerning a great deal of subjects will help students further enhance their language. This is something that I consider extremely important, and it is the main area of interest for me regarding CLIL instruction. The potential increase in exposure to language, and the ability to teach a type of

vocabulary that is often lacking in ordinary EFL teaching are things that could greatly improve upon the language level of Norwegian learners of English. Learning how best to teach vocabulary, especially in the challenging environment of CLIL, could also be extremely valuable. Indeed this is a problem that every CLIL teacher will come into contact with, whether they like it or not. Vocabulary will create challenges for any CLIL instructor, not to mention EFL teachers, and having good knowledge and tools available to help alleviate these issues could be not only beneficial for learning, but also time-saving.

The Research Question

Wanting to look at the improvements in vocabulary that CLIL could yield I settled on the following research question for this thesis: “Do classes with EFL & CLIL score better on vocabulary tests than classes with EFL only?” Originally I had hoped to have both pre- and post tests which would allow for an analysis looking at the improvement of students in CLIL classes compared to EFL only classes. This proved to be time-consuming and I settled for a more manageable alternative. My hypothesis is that successful CLIL instruction should lead to better vocabulary scores in the students than the students who are not receiving CLIL. The wording implies that classes who have received CLIL instruction should be, on average, scoring higher on vocabulary tests than those who receive only EFL instruction. The question focuses on the tests and the test scores themselves, not any extenuating circumstances around them. Although these extenuating circumstances will be documented and looked at in the thesis, they are simply there to provide clarity to the results of the tests, and an attempt at explaining the results.

Overview of the Thesis

In this thesis I will attempt to answer the research question stated above “Do classes with EFL & CLIL score better on vocabulary tests than classes with EFL only?”. I will do so by first presenting an overview of CLIL, its history, its place in Norway, and look at a study relating to vocabulary development in CLIL. In chapter 3 I will be defining what vocabulary really is, what it means to know vocabulary and discuss the place of vocabulary in the current curriculum in Norway, the LK06. Then in chapter 4 I will be presenting some theory and studies on the implicit learning of vocabulary. This theoretical perspective will be used as the basis for my arguments later in the thesis. Chapter 5 will contain the methods being applied to this study, how and why I decided to use them and I will also answer questions regarding the validity of the thesis. In chapter 6 the results of the study will be presented and analyzed. Then in chapter 7 I will be discussing these results, bringing together the separate pieces of data, comparing them and discussing them in light of the theory presented in chapter 4. I will also be discussing the validity further in this section. Finally in chapter 8 I will make some concluding remarks, including suggestions for further research and the implications of this study.

2) Content and Language Integrated Learning

Introduction

In this chapter I will provide an overview of CLIL, its history, and its definition. CLIL is still a relatively new subject in Norway, and I will be discussing the current status of CLIL and how it has been implemented so far. I will also be looking more closely at a study by Sylven (2010) who investigated the effects of CLIL on the vocabulary of students in Sweden.

Content and Language Integrated Learning

CLIL is considered as a catch-all definition that encompasses all forms of instruction in a non-language subject taught in a foreign language (Svenhard 2010, p.5). Christiane Dalton-Puffer (2007) explains that one of the main pro-CLIL arguments is that:

...the curricula of the so-called content subject (eg. Geography, history, business studies etc.) constitute a reservoir of concepts, topics, and meanings which can become the object of 'real communication' where natural use of the target language is possible.

(Dalton-Puffer 2007, p. 3).

One of the main concerns when looking at CLIL instruction is the relationship between the content and the language. According to Dalton-Puffer there is a good deal of tension between

these two despite the presence of the connecting word integrated between the content and language in CLIL.

The concern of many subject teachers is that the presence of a foreign language will negatively affect the student's knowledge of the subject. These teachers are concerned with both the amount and the quality of learning. Some believe that a lower language proficiency in the foreign language will negatively impact both the complexity of the subject matter and the amount of subject matter that is to be taught. As it stands today, CLIL instruction seems to put Content ahead of language without any clear indication as to why this is (Dalton-Puffer 2007, p.5-6). The language is considered secondary, however it must be assumed that there are language related goals to the instruction as well as content specific goals. If there are no language related goals to teaching CLIL then there is little purpose to it in the first place. Most teachers are eager to try CLIL because it could improve language. However, they become wary when they feel that the language is getting in the way of the subject teaching. The CLIL compendium, which is an EU funded project that was concluded in 2002, lists several language related goals for CLIL instruction:

- A. Improve overall target language competence
- B. Develop oral communication skills
- C. Deepen awareness of both mother tongue and target language
- D. Develop plurilingual interests and attitudes
- E. Introduce a target language

(<http://www.clilcompendium.com/clilcompendium.htm>)

This list of language related goals is quite general as we can see, with no clear- cut goals relating to the language. Lacking these more specific goals relating to the language is a problem according to Dalton-Puffer (2007), and needs to be looked at more closely. In the

next section I will detail the history and background of CLIL, looking at the Immersion programs in Canada first of all, before shifting to Europe and then to Scandinavia and Norway before discussing the current situation for CLIL in Norway.

The Roots of Content and Language Integrated Learning

Content and Language Integrated Learning has its roots in the immersion programs implemented in Canada. The immersion programs were made to help improve the French of the English-speaking majority in a country that considers itself bilingual, with both French and English as official languages. First implemented in the 1970's, the immersion programs focused on developing communicative competence in French, with over half of the instruction being in French. Several forms of immersion programs have sprung up over the years, the earliest form was coined the total immersion program, where children would be taught exclusively in French throughout kindergarten and English would be gradually implemented into the teaching as the children grew. At approximately year 6 the students would receive a 50/50 split of English and French instruction. Other forms of immersion programs have also been implemented in different forms, but with varying amounts of French instruction. The late immersion program implements French as late as grade 8, with the amount of instruction varying somewhat between schools (Sylvén 2010, p.14-16). The best results have come from the early total immersion programs with regards to the children's abilities in French. In these immersion programs near-native like listening and reading comprehension is achieved in French, although the same cannot be said for speaking and writing. This means that their receptive skills are more improved than their productive skills in the language. The late immersion program also has students who lag behind the skills of students who receive early immersion. In addition to this, it has been shown that the academic achievements of the students has not been negatively affected by immersion programs. In fact, early immersion has shown signs of enhancing the academic abilities of students rather than harming them (Sylvén 2010, p.17).

Canada has not been alone in experimenting with the effects of bilingual teaching, with several European countries testing the waters of bilingual instruction starting as early as the 1960's in Germany. We can say that the interest in bilingual teaching has increased considerably over the last 20 years, starting in the 1990's (Sylvén 2010, p.18-19).

Content and Language Integrated Learning in Norway and Scandinavia

The use of Content and Language Integrated Learning in the Scandinavian countries has seen a dramatic increase since the 1990's. In Finland the use of CLIL instruction to teach Swedish has been in use since 1987. As Finland is a bilingual country with both Finnish and Swedish languages, much like Canada, the use was at first restricted to teaching Swedish. Since the 1990's however, the teaching of English in CLIL instruction has become more and more popular in the country. In Sweden the teaching of foreign languages in other subjects has been present for a long time, with some of the oldest schools in Sweden, such as the Deutsche Schule in Stockholm, teaching, unsurprisingly, in German. These schools are not organized in any way, work independently from each other, and function as international schools rather than normal schools teaching CLIL. The teaching of CLIL in Sweden has been used mostly by individual experimenters, starting up in the late 1970's (Sylvén 2010 p.19-24).

The situation in Norway is somewhat different from Sweden and Finland. In 1993 the Norwegian Ministry of Research and Education created the first CLIL classes. In doing so the ministry also set down the requirements demanded of CLIL instruction, the requirements being that at least 30 percent of the teaching is performed through the target language and that the participation of the students is voluntary (Sylvén 2010, p.19). Since these first CLIL classes the use of CLIL instruction has grown slowly, and it is not implemented to any large degree in Norway at the present time. In the next section I will be looking more closely at the current status of CLIL instruction in Norway.

2.1.1 CLIL in Norway

The teaching of CLIL in Norway, since its beginning in 1993, has grown slowly but surely towards what it is today. Still on the fringes of education, it is becoming a more important and interesting part of the foreign language teaching in Norway. CLIL has been used to teach languages such as English, German, and French in Norway (Kunnskapsdepartementet 2007, p.24). CLIL is also mentioned in the Ministry of Education's strategy plan *Språk Åpner Dører* for language teaching. The plan of action was to implement CLIL instruction in the elementary school level for students starting in 2005 (Kunnskapsdepartementet 2007, p.39). In 2009 the Norwegian Centre for Foreign Languages in Education, which was established in 2005, began a second implementation of CLIL in Norway, this time at the upper secondary school level. In this project a total of nine different upper secondary schools participated. In their work they discovered that although CLIL instruction had existed in Norway since 1993, the widespread implementation of this was hampered by external factors limiting the ability to teach CLIL. Looking at how to best implement CLIL instruction therefore became an important part of the project, as well as identifying the limiting factors that hindered it. Just as with the earlier project from 2005, the main focus of the project was to look at the effects of CLIL instruction on the basic skills of reading and writing, comparing their work with earlier works such as Hellekjær (2005), who looked at reading (Svenhard 2012, p.3).

The project initiated by the Norwegian Centre for Foreign Languages in Education did not put any restraints on the teachers participating in the study when it came to their approach to CLIL instruction. Svenhard (2012) argued that every school should have an individual approach to CLIL if it was to thrive and grow at the school. Therefore the approach of the teachers has been different, although Svenhard again notes that many of the same topics were brought up when the teachers discussed their experience. (Svenhard 2012, p.4).

Hellekjær (1996) explains some of his experience from teaching a Norwegian history class in English. His focus was on teaching the content of the subject, limiting the teaching of language to especially important points. In this history class the textbook being used was in English. The students struggled with using this textbook, lacking the proper reading strategies to read effectively. Many students were attempting to read word for word, looking up each

word in the dictionary as they went. They were simply reading the way they would in an English class (Hellekjær 1996), which means they would miss the forest for the trees so to speak. The solution to this, according to Hellekjær, was not to abandon English textbooks, but instead to get rid of the students' old habits. With extensive reading, and the teaching of good study skills and reading strategies the students should become more adept at managing the English textbook (Hellekjær 1996). Hellekjær also recommended a high degree of fluency in the target language for the teacher, preferably with a degree in it as well as a degree in the subject. This was noted as being somewhat challenging to achieve, however (Hellekjær 1996). The teaching of CLIL is still new, with a lacking infrastructure around it. Hellekjær's points regarding it's teaching shows how CLIL is still a very new concept. Lacking proper materials and textbooks for the subjects, and the difficulty in finding teachers with suitable qualifications makes CLIL very much a work in progress still in Norway.

Hellekjær & Hopfenbeck (2012) looked at the effects of CLIL on the reading skills of students, discovering a marked improvement in the reading skills of Norwegian students since 2002. They also argue that the use of CLIL instruction can be effective in teaching reading skills and comprehension as long as the instruction is of sufficient quality and quantity. They mention that CLIL instruction should ideally consist of at least 50 percent of the total instruction in the subject being taught in the target language (Hellekjær and Hopfenbeck 2012, p.117-118). Their critical view of the current CLIL instruction shows that if the language, as Dalton-Puffer has claimed, takes a back seat to the content, then the language teaching in CLIL will have a lesser effect than that which is desired. While there has been a considerable amount of research looking at CLIL instruction in Norway, most of it has been focused on reading, or the implementation of CLIL. In the next section I will outline a study by Liss Kerstin Sylven, looking at the incidental vocabulary acquisition of Swedish learners of English in CLIL.

Research on Vocabular in CLIL

The research of vocabulary acquisition in CLIL in Norway has, as was mentioned above, been more focused on the effects on reading. Sylven (2010) has, however, looked at the effect of CLIL instruction on the vocabulary of learners in Sweden. More specifically she looked at the incidental vocabulary acquisition enjoyed by CLIL students as opposed to the vocabulary of students in normal programs. The study was based on the idea that a greater exposure to the language and more input should lead to a larger and richer vocabulary (Sylven 2010, p.6-7). Her study consisted of four different Swedish schools teaching CLIL, in addition to this she also included control groups of non-CLIL students in her study so as to compare their results to the CLIL results (Sylven 2010, p.8). The study lasted for a full two years and consisted of three rounds of tests, which contained several different types of tests. This meant that the students had a full school year between each of the test periods (Sylven 2010, p.215-216). Her results seemed to indicate at first that CLIL instruction had helped the students to acquire a greater vocabulary than that of the control groups. There were problems however, as the CLIL students seemed to enjoy a larger amount of exposure to English outside of school than that did the control groups. This meant that it was difficult to know what could be attributed to CLIL and what could be attributed other factors. Sylven did note that the amount of input seemed to have a large effect on the vocabulary of the learner. One interesting thing that came out of the study was that one of the CLIL classes who received the least amount of English instruction in CLIL scored the highest on the tests. Sylven argues that this indicates that it is the quality and not the quantity of the instruction that is the most important (Sylven 2010, p.218-219).

Looking at the habits of the students outside of school Sylven discovered that this might have been the greatest impact on the results of the study, more so than the CLIL instruction. Those students who would read English texts outside of school consistently scored better than those who did not. She noted that most of the CLIL students did exactly this, while most of the control students did not. In addition to this, gender habits seemed to impact the vocabulary of the students as well. The boys would often play computer games outside of school containing English language. The boys scored better than the girls on the vocabulary tests. Interestingly

Sylven found that the CLIL instruction seems to have been most beneficial to the girls, as the girls in the CLIL classes start catching up to the vocabulary of the boys towards the end of the study (Sylven 2010, p.220-221). The end result of the study indicates that the CLIL instruction has had an effect on the vocabulary of the students, with the teachers noting how the subject-specific terminology of the students has been improved. Sylven however, argues that this is not what should be of primary concern, as the students lack a general broad vocabulary that needs to be acquired.

Sylven also touches on the issue of content vs language as Dalton-Puffer mentioned above. Findings in Sweden, contrary to that of others, seem to indicate that the content has suffered from the CLIL instruction. Indeed, one student even claimed it was a waste of her time, due to the poor English of her teacher (Sylven 2010, p.223-224). This claim is interesting because it highlights the importance of not only content knowledge from the teacher, but also language knowledge, lest both content and language should suffer as a result.

Summary of the Chapter

In this chapter I have given a brief overview of the roots of CLIL, following them up to the present day in Norway, detailing some of the work that has gone into it's study here. However, CLIL has a sort of grass-roots element to it. With many teachers starting their own smaller projects, getting their feet wet if you will, experimenting with the CLIL approach. Therefore, while CLIL has been initiated by the department of education in Norway, there are many more teachers attempting CLIL in the Norwegian school system on their own. The focus in Norway has mostly been on the basic skills, and primarily on reading, looking at the effects of CLIL on the reading skills of students. Vocabulary has in many ways taken a back seat, always there, but never being the centre of attention. This study hopes to show how important vocabulary is to successfully teaching CLIL, and how vocabulary is a lynchpin ingredient in any successful English program, whether the focus is on reading or writing.

3) Vocabulary Knowledge and Vocabulary in the LK06

Introduction

In this chapter I will be looking at the nature of vocabulary, what it takes to know a word, and what vocabulary really is. I will also be looking at what the most important vocabulary to learn in the English language and why this is. I will then follow with a section on the role of vocabulary in the Norwegian school system and the LK06 syllabus, about what the focus on vocabulary is in the current curriculum and discussing if the role of vocabulary in the syllabus is defined well enough.

The Nature of a Word

To discuss vocabulary we will first need to define what a vocabulary truly is. The most common definition of a word is the lemma, which constitutes the base and inflected forms of a word, e.g run, running, ran. These would all count as a single lemma and knowledge of one implies knowledge of the others (Read 2000, p.18). However, words can also take on different meanings and might not be forms of the same lemma, an example of which would be the words social and society. These two words have a connection in that they both relate to people, sharing the root soci- form. Their meaning is different from each other placing them into different word families, which is a way of classifying words that share a common meaning (Read 2000, p.19). The reason this is so important to vocabulary is the way we measure vocabulary size. A person who knows the meaning of the word social could not necessarily be expected to know the meaning of the word society. Different researchers measure vocabulary according to different criteria, a study focusing on word families might

estimate a larger vocabulary size than one focusing on lemmas (Read 2000, p.19) In the next section we will look further at the complicated issue of what vocabulary truly is by looking at multi-word items.

3.1.1 Multi-Word Items

While the common understanding of a word is closely related to individual units of words such as what you might encounter in a dictionary, corpus research has revealed that language is not made up of individual units of words that are completely separate from each other. Corpus and text research has shown that words are linked together, and that there are patterns linking the different words together (Moon 1997, p.40). The definition of a multi-word item according to Rosamund Moon (1997) is:

A multi-word item is a vocabulary item which consists of a sequence of two or more words (a word being simply an orthographic unit). This sequence of words semantically and/or syntactically forms a meaningful and inseparable unit. Multi-word items are the results of lexical (and semantic) processes of fossilisation and word-formation, rather than the results of the operation of grammatical rules.

(Moon 1997, p.43)

Multi-word items that fit with this description include: Idioms, compound words, phrasal verbs, and fixed expressions. Finally there is another category of multi-word items, namely prefabricated sentences, or sentence stems. A sentence stem consists of a chunk of words or phrase that has become institutionalized, an example of such a sentence stem could be that reminds me (Moon 1997, p.45-47). These sentence stems are not as strictly joined together as idioms or fixed expressions, they are instead somewhat more flexible, only that they are commonly used in conjunction with each other. This kind of extreme collocation can be seen throughout the English language. An example of how collocation has become ingrained in the language we can look at an example. When describing that it is no longer raining outside we would typically say that the rain has stopped, not the rain has ended or finished. Continuing this we can look at the collocation between torrential and rain and see from corpus research

that torrential is almost always followed by the word rain (Moon 1997, 40-41). The idea that we do not learn individual words outside of any context but that we learn, adapt and use sentence stems to create language, is an interesting field of vocabulary research that is still in it's infancy.

Word Knowledge

When testing vocabulary it is also important to consider what it means to know a word. Knowing a word is not the same as recognizing it or even being able to give a meaning. In 1976 Richards wrote an article regarding lexical competence, and in this article he also outlined what is required to know a word:

- Knowing a word means knowing the degree of probability of encountering that word in speech or print. For many words we also know the sort of words most likely to be found associated with the word.
- Knowing a word implies knowing the limitations on the use of the word according to variations of function and situation.
- Knowing a word means knowing the syntactic behaviour associated with the word.
- Knowing a word entails knowledge of the underlying form of a word and the derivations that can be made from it.
- Knowing a word entails knowledge of the network of associations between that word and other words in the language.
- Knowing a word means knowing the semantic value of a word.
- Knowing a word means knowing many of the different meanings associated with a word.

(as cited in Sylven 2010, p.36)

This list shows how complex the nature of knowing a word can be, and how difficult it can be to measure vocabulary. Other researchers have in later years attempted to better define and qualify what it means to know a word. Nation adapts Richards list further, looking at the requirements of fully knowing a word:

1. The spoken form of a word.
2. The written form of a word.
3. The grammatical behavior of the word.
4. The collocational behavior of the word.
5. How frequent the word is.
6. The stylistic register constraints of a word.
7. The conceptual meaning of a word.
8. The associations a word has with other related words.

(Nation 1990, p.31)

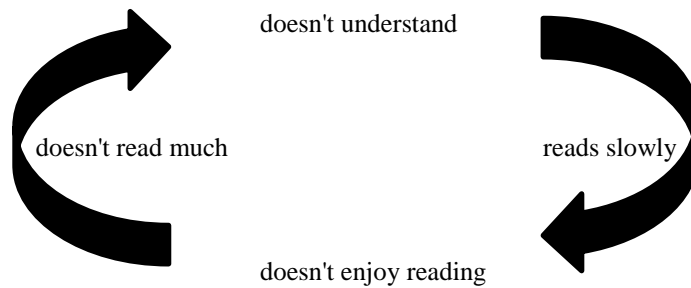
This list shows the difference between identifying and understanding meaning compared to using and applying a word in context. This list, and other similar ones, are fairly impractical in use. It is difficult to get any real sense of vocabulary, as in practice only a few words can be tested in this way. (Read 2000, p.26-27) In the next section I will be looking at the importance of the most frequent words in the English language, and at their importance when learning vocabulary.

The Most Important Words in the English Language

According to Nation the most frequent words in the English language account for approximately 80% of most texts. These headwords consist of approximately 2000 words (Milton 2009, p.45-46) and these should be focused on when teaching vocabulary. Without these words it will be impossible to use the English language properly (Nation and Newton 1997, p.239). After having learned these words it is important to distinguish what use the learner will have of English. If the learner is supposed to do academic study then Nation recommends that the 800 headwords in the academic vocabulary list made by Nation (1990) should also be learned. The academic vocabulary that Nation recommends learning consists of words taken from texts used in upper secondary school, university and newspapers. Some examples of words are abandon, comply, denote, evident (Nation and Newton 1997, p.239). As Nation and Newton (1997) claim, guessing a word from context occurs successfully when approximately 98% of the lexical items in the text are known. This is problematic according to Laufer as to know 98% of the words in a text a learner will need to know a total of 5000

word families, or 8000 lexical items (as cited in Coady 1997: 229). This is a daunting number and creates a problem when attempting to teach, especially when reading. In other words, too many unknown words will create a barrier for the learner, quickly becoming unintelligible. This creates a vicious circle according to Nuttall (as cited in Coady 1997, p.233).

Figure 3.1 The Vicious Circle of Reading



(Coady 1997, p.233)

This circle shows that a learner who does not read very much will be trapped in a circle where reading becomes so difficult that it is not enjoyable and the learner therefore reads less, thereby continuing the cycle. The reason this creates a problem for language teaching is that written English contains far more word types than in spoken English. Therefore a great deal of the language can only be encountered and thereby learned through reading (Coady 1997, p.230). In dealing with this vicious circle of reading, language teachers have given learners texts where the vocabulary is severely limited to be easier to read for learners without a large enough vocabulary. This has been widely criticized as these texts are not seen as “authentic” changing the syntactic and semantic usage of the vocabulary. The critics claim that the texts do not prepare foreign language learners for real texts. Coady (1997) notes that native speakers also read simplified texts when beginning to read, and many proponents for this approach argue that it is usable but learners should as quickly as possible switch to more authentic texts (Coady 1997, p.230-231).

Summary of Work Knowledge and Word Learning

To sum up the discussion on the nature of a word and vocabulary knowledge we have seen that vocabulary, and words, are complex language items, not easily classified and often difficult to fully understand. Knowing a word is a demanding task, and complete knowledge of a word needs far more than simply recognizing the word or being able to give the meaning of the word. I have also looked at the words that are most important to learn in the English language and why they are so important.

In the next few sections I will be looking at the curriculum in Norway and the English language curriculum specifically. In doing so I will be discussing the place of vocabulary in the curriculum and what the focus of the vocabulary learning is in the English subject curriculum.

The English Subject and the LK06

In 2006 a new school reform was initiated in Norway called the National Curriculum for Knowledge promotion in Primary and Secondary Education, commonly referred to as LK06.

The LK06 comprises five different areas, which are as follows:

- The Core Curriculum
- The Quality Framework
- The Subject Curriculum
- Distribution of Teaching Hours per Subject
- Individual Assessment

The core curriculum deals with overarching goals and values that are deemed important in all subjects while the subject curriculum deals with specific goals relating to the individual subjects. I will in this chapter focus on the subject curriculum in English. English is taught in the Norwegian school system from year 1 in primary school up until year 1 in upper secondary, with an option to select English in years 2 and 3 of upper secondary as well, which we will not look at in this chapter. Vocational students also have English in both the first and second year of upper secondary. The following list shows the amount of teaching hours given to the subject (a teaching hour is considered 60 minutes):

- Primary School (Years 1 to 7): 328 teaching hours
- Lower Secondary (Years 8 to 10): 227 teaching hours
- Program for General Studies; Upper Secondary Year 11 (vg1): 140 teaching hours
- Vocational Education Programs; Upper Secondary Year 11 and 12 (vg1 and vg2): 84 (vg1) and 56 (vg2) teaching hours

(Kunnskapsdepartementet 2010)

3.1.2 The Basic Skills and Vocabulary

In the LK06 there are four basic skills that have been deemed necessary in all subjects, these are reading, oral and written production, arithmetic and use of digital tools. In the English subject curriculum they are understood as:

- Being able to express oneself in writing and orally
- Being able to read
- Numeracy
- Being able to use digital tools

(Kunnskapsdepartementet 2010)

All of these basic skills are important to master in English according to the LK06. While not being a basic skill, vocabulary can be seen as an integral part of mastering these basic skills. For example under the basic skill of numeracy the LK06 states “...being able to supplement mathematical competence in one's native language with the necessary terms in English.” (Kunnskapsdepartementet 2010). Vocabulary becomes an integral part of the basic skills as it is needed in all of them.

The English Subject Curriculum and Vocabulary

The English subject curriculum states the goals and objectives that students are required to learn throughout their English education through competency aims. These competency aims are after the second, fourth, seventh and tenth years in primary and lower secondary school, and then after the first year of upper secondary for general studies, while for vocational studies they are after the second year. These competency aims describe what is expected of the student at these intervals in their education. The competency aims are divided amongst three main subject areas that the English curriculum has, these are as follows:

- Language Learning
- Communication
- Culture, Society and Literature

Language learning focuses on how we use the language and the strategies we use in acquiring it. Communication focuses on being able to convey meaning using oral and written interaction. Finally culture, society and literature is aimed at teaching the student about the English speaking culture and world. With regards to vocabulary we find this is mostly under the area of communication. The area of communication mentions vocabulary amongst other things stating “Good communication requires knowledge and skills in using vocabulary and idiomatic structures...” (Kunnskapsdepartementet 2010, p.3)

The English subject curriculum mentions vocabulary in discussing the curriculum's main purposes.

To succeed in a world where English is used for international interpersonal communication, it is necessary to master the English language. Thus we need to develop our vocabulary and our skills in using the systems of the English language; its phonology, grammar and text structuring. We need these skills to listen, speak, read and write, and to adapt our language to an ever -increasing number of topics, areas of interest and communication situations.

(Kunnskapsdepartementet 2010, p.1)

Here the subject curriculum specifically mentions vocabulary as being necessary to communicate with the world around us. Vocabulary is necessary in order to use many of the basic skills mentioned above as well as mastering the language itself. This prominent place in the curriculum implies that the authors of the curriculum feel that vocabulary is a very important part of English language learning.

3.1.3 Competency Aims in the Subject Curriculum Relating to Vocabulary

In the subject curriculum there are many competency aims that involve a need for vocabulary, I will in this section highlight those that directly relate to vocabulary teaching in English. I will be looking at the competency aims from the end of the tenth year of lower secondary school, as this is the group I will also be investigating in my study. While I have selected to look at this year in particular, the competency aims are mostly similar throughout, the only difference being the degree of mastery.

The word vocabulary is only mentioned once in the subject curriculum for the end of the tenth grade, it is mentioned in the following competency aim:

- Master vocabulary that covers a range of topics

(Kunnskapsdepartementet 2010)

This is the only time that vocabulary is specifically mentioned in the subject curriculum, while there are many more competency aims that allude to the topic of vocabulary, or demand vocabulary to be achieved. However, this is the only competency aim that specifically states that the learners has vocabulary knowledge and only vocabulary knowledge. As we can see from the competency aim, it is extremely vague. Mastering vocabulary that covers a range of topics is not very specific, however the word master implies a degree of knowledge at least. To master vocabulary, or to master a word, we can assume complete knowledge about that word. This is problematic as we have shown in this chapter so far. Looking at Nation's list in section 3.3 we can see that to fully know a word the learner is tasked with a great deal. In fact this competency aim spills over into other competency aims as we can see in the following list:

- understand spoken and written texts on a variety of topics
- write texts that narrate, describe, argue or give messages, with the appropriate basic structure and adequate paragraphing
- use basic terminology to describe grammar and text structure

(Kunnskapsdepartementet 2010)

While this list is by no means complete it gives an idea of the scope that the single competency aim specifically regarding vocabulary demands of the learner, and how vital vocabulary is to the completion of the other competency aims in the subject curriculum.

Discussion on the Place of Vocabulary in the Subject Curriculum

The place of vocabulary in the current subject curriculum in the Norwegian school system is one that seems to be lacking, not only for the tenth grade. The entire subject curriculum is in my opinion so vague when discussing vocabulary that this is problematic. The only place, in the tenth grade curriculum, where vocabulary is mentioned as a specific competency aim is in the communication section of the curriculum as shown in section 3.7.1 above: master vocabulary that covers a wide range of topics. As we have discussed previously in this chapter knowing the 2000 most frequent words, as defined by Nation, in the English language are absolutely necessary to be able to use English at an acceptable level. The lack of focus on the specific vocabulary that is to be learned by the tenth grade means that teachers are left to their own devices in discovering what vocabulary is most important for their students. A student who lacks part of those 2000 words will continue to struggle with English and as shown by Nuttall (as cited in Coady 1997) this will lead to difficulties in reading. While vocabulary is a basic requirement for many of the competency aims in the English subject curriculum more focus is needed with regard to vocabulary that should be learned, and greater importance needs to be placed on the students actually learning these words. For such a vital and important building block in the language of the learner the curriculum is surprisingly vague

and seems to put little importance on the teaching of vocabulary specifically. It seems that one is almost expecting vocabulary to come as a by-product of language teaching, and that it does not need to be focused on specifically.

In the next chapter I will be looking more closely at incidental vocabulary learning, this is due to its natural affinity with Content and Language Integrated Learning, and this will help shed further light on the challenges that teachers face with when teaching vocabulary.

4) Vocabulary Learning

Introduction

In this chapter I will be discussing the idea that vocabulary can be obtained by being exposed to the language, this is called incidental or implicit vocabulary learning, which will be referred to as incidental vocabulary learning in this thesis. This is one of the basic ideas behind the CLIL method. This chapter will be used to discuss the findings from the study in the discussion part of the thesis, chapter 7. Obtaining vocabulary from exposure to the language is something that CLIL gives ample opportunity to. The question remains if this is sufficient for vocabulary acquisition. The following theory and studies look at what effects incidental vocabulary learning has, and how and if additional tasks or repetitions can improve the amount of vocabulary being learned. I will be first discussing the idea of incidental vocabulary learning and some of the theory regarding this, and then I will be presenting a number of studies that look into the effects of incidental vocabulary learning, and whether incidental contact alone is enough or if the learner needs something more to improve their vocabulary effectively.

Incidental Vocabulary Learning

When teaching vocabulary, there are two approaches to choose between, the incidental and explicit approaches. The incidental approach is based on the natural language learning of our first language, and stems from the Natural Method that was based on how a child would learn a language from their family and environment (Sylvén 2010, p.28-29). This is also one of the most important aspects of the CLIL method. When immersion programs first appeared in Canada the belief was that the vocabulary would be learned naturally through exposure to the language, by placing the learners in an environment that would give them this incidental

contact with the language (Sylvén 2010, p.29). The other approach, the explicit approach, focuses on teaching vocabulary directly. This can be done through a number of ways, such as direct translation, description, or associated words for example. The focus is on teaching very specific words.

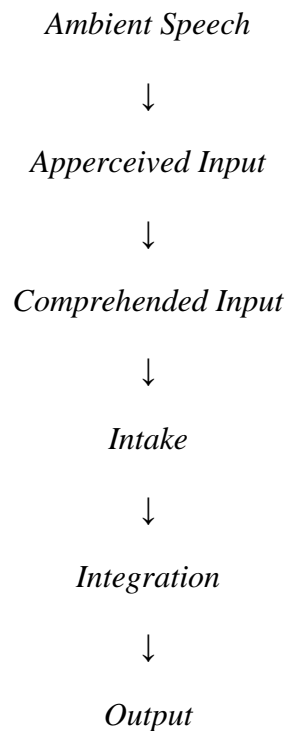
Nation talks about a cost/benefit, with more frequent words being worthwhile to teach explicitly while other less frequent words can be left to incidental contact (Schmitt 2000, p.120-121). This is because these words will occur so often in the English language that to comprehend a text the learner will struggle without knowledge of these words. The frequency of words in a language tend to follow Zipf's law, which states that the most common word in a language is likely to occur twice as often as the second most common word in a language. This shows that knowing approximately 2000 words should allow the learner to understand 80% of all text. It is these 2000 words that Nation recommends could be taught explicitly as they are so useful in the English language (as cited in Milton 2009, p.45-47).

The focus in the CLIL method, as stated above, is on the incidental approach, exposure to the language should naturally lead to an expanded vocabulary. Stephen Krashen (1989) argued for the importance of input in vocabulary learning, with his Input Hypothesis. This stated that reading texts where approximately 95% of the text was known would allow for the new vocabulary to be learned simply through input. Krashen (1989) based his model for second language acquisition on Chomsky's theory of Universal Grammar, which states that children have an innate ability to learn and develop language. The Universal Grammar theory is primarily constructed to explain the acquisition of a first-language, not a second-language. While Krashen's input hypothesis has been heavily criticized it has still been a major factor in the development of foreign language teaching programs such as CLIL (Lightbown and Spada 2006, p.34-38).

A theoretical framework for how vocabulary is acquired and learned was presented by Gass (1988). Her theory was that vocabulary learning goes through five stages before input can

become output, complicating Krashen's view on the nature of vocabulary learning. The five stages are as follows:

Figure 4.1 The road from ambient speech to automated output



(Gass 1988, p.200)

In this framework Ambient Speech is the exposure to the language, whether in the form of a book or a classroom discussion. The apperceived input is the vocabulary that is noticed by the learner, some of which becomes comprehended input through analysis, attempting to understand the vocabulary. Some of this comprehended input is then understood and analyzed to such a degree that it in turn becomes intake, which is integrated with the internal grammar and speech of the learner to finally become output (Sylvén 2010: 31-33). There are also studies that have looked at the effects of the form of input on the output. What kind of input is most effective for certain kinds of output such as written production to give an example.

A study performed by Saragi, Nation and Meister tested learners' understanding of the Russian words being used in the novel *A Clockwork Orange* (as cited in Read 2000). The goal of the study was to look at the amount of words that were learned incidentally through contact with them in the book. Testing the Russian words they could ensure themselves that few if any of the learners had any other contact with the target words. The result of the study showed that 76% of the 90 Russian target words were known on average by the learners (as cited in Read 2000, p.45). To learn a word they suggested that the minimum number of repetitions that was required needed to be approximately ten (as cited in Paribakht and Wesche 1997, p.176). This implies that incidental contact with the target word is required to be fairly frequent to result in learning. Other studies testing the incidental acquisition of new vocabulary in texts have shown modest gains of only a few words per text (Lee and Hirsh 2012, p.81). Schmidt has argued that the learner needs to be consciously aware of the words to be able to learn them and are unable to learn new vocabulary without noticing them in some way (as cited in Read 2000: 44).

It is important to note that most of the focus on incidental learning of vocabulary has been conducted through reading, as the amount of vocabulary in written English is far greater than that of oral English. Speech also tends to be less formal and less academic than what we see in writing, resulting in a different vocabulary being used (Milton 2009, p.55). In addition to this the language spoken by the teacher is contains a different vocabulary than that found in the written material learners are subject to, there is however little evidence or research conducted on the incidental vocabulary acquisition of spoken language (Milton 2009, p.197). In the next section we will be looking at a number of studies testing for the effects of incidental vocabulary learning on learners.

Paribakht and Wesche's Study on Incidental Vocabulary Through Reading

Based on the research on incidental vocabulary acquisition Paribakht and Wesche (1997) wanted to look at vocabulary acquisition through reading. They wanted to find out if reading plus vocabulary exercises would lead to a more effective vocabulary acquisition than reading plus reading additional texts if the amount of time spent was equal between the two approaches. In addition to this, they wished to see which vocabulary exercises were perceived as the most useful by teachers and students (Paribakht and Wesche 1997, p.177-178).

The subjects were taken from two ESL classes of an intermediate level at the University of Ottawa, a total of 38 students participated, having a variety of L1 backgrounds. The group was split into two parts, one part was to receive reading exercises plus vocabulary exercises relating to the text, called Reading Plus. The other group received reading exercises as well as additional reading material on the subject, this group was called Reading Only. Both groups were to receive the same amount of time to study the materials (Paribakht and Wesche 1997, p.182-187). To analyze the results of the study Paribakht and Wesche used the Vocabulary Knowledge Scale to help distinguish the stages of word knowledge that the students had. The VKS scale has five self-report categories:

Fig 4.2 VKS Elicitation Scale

Self-report Categories

- i. I don't remember having seen this word before*
- ii. I have seen this word before, but I don't know what it means*
- iii. I have seen this word before, and I think it means _____. (synonym or translation)*
- iv. I know this word. It means _____. (synonym or translation)*
- v. I can use this word in a sentence: _____. (Write a sentence.) (If you do this section, please also do Section IV.)*

(Paribakht and Wesche 1997: 180)

The scale allows for the measurement of word knowledge from total unfamiliarity to correct use, both grammatically and semantically (Paribakht and Wesche 1997: 179). The results of the study showed that both groups showed significant gains in vocabulary learning. The Reading Plus group did however have greater gains than the Reading Only group, obtaining more vocabulary. Paribakht and Wesche (1997) commented that it was interesting to see the results of the Reading Only test score as high as they did, however, it is important to note that the texts supplied to the learners were thematically related and the target words were repeated several times throughout the texts . Using the Vocabulary Knowledge Scale shown above they also found that the Reading Only group tended more to simply recognize words, while the Reading Plus group reached a higher level of knowledge regarding the target words (Paribakht and Wesche 1997, p.195-196). Their conclusion from the study was that while reading that utilized repetition of words would lead to vocabulary acquisition, reading supplemented by vocabulary exercises allowed for greater gains, and a higher level of knowledge regarding the words (Paribakht and Wesche 1997: 197). Having also looked at the teacher and students' opinions regarding the usefulness of the different vocabulary exercises it was revealed that the students found the text content to be more important than the vocabulary exercises. They commented that having texts that they liked mattered more than the type of vocabulary exercise utilized (Paribakht and Wesche 1997, p.195). In the next section we will be taking a look at what this involvement from the student can have as an effect on the vocabulary acquisition process.

The Involvement Load Hypothesis

Laufer and Hulstijn (2001) claimed that learning of a word was facilitated not by what kind of exposure the learner had to it, but instead to what need and involvement there was on the part of the learner. Laufer and Hulstijn (2001) presented three key points in word learning, namely Need, Search, and Evaluation. The Need dimension consists simply of whether or not the learner needs the word to complete a task, while the Search dimension is if the learner attempts to understand the meaning of the word or not. This Search can be from any source, it

only demands that the learner is consciously attempting to understand the meaning of the word without having this presented immediately. The Evaluation dimension is the comparison of the new word with others and deciding on its contextual use. The learner makes a decision on when this word is appropriate to be used (Laufer and Hulstijn 2001).

Fig 4.3 Laufer & Hulstijn's Three components of Involvement

Components	Feature	Operationalisation	Prominence
<i>Need</i>	<i>Motivational</i>	<i>Whether knowledge of new words is required to complete a given task</i>	<i>Absent (0) Moderate (1) Strong (2)</i>
<i>Search</i>	<i>Cognitive</i>	<i>The attempt learners make to ascertain the meaning of unknown words in a task</i>	<i>Absent (0) Present (1)</i>
<i>Evaluation</i>	<i>Cognitive</i>	<i>Comparing a new word with other words and making a decision as to its suitability in a given context</i>	<i>Absent (0) Moderate (1) Strong (2)</i>

(Laufer and Hulstijn 2001)

Each dimension is given a score, Need and Evaluation have from 0 (absent) to 3 (Strong) while Search has only 0 (Absent) and 1 (Present). The higher the score, the greater the involvement and therefore the greater retention of the word. This is in line with the thinking that vocabulary cannot be, at least not effectively, learned without the learner being consciously aware of it. Input is simply not enough by itself and more is needed if the word is to be learned. This is especially interesting with regards to CLIL, as the potential for vocabulary learning seems to need more than simply being subject to the language.

4.1.1 Laufer and Hulstijns Research on the Involvement Load Hypothesis

Laufer & Hulstijn's (2001) research with Dutch learners of English seemed to confirm their hypothesis, with those being given tasks with a higher involvement load being more successful in vocabulary acquisition. Laufer and Hulstijn (2001) tested advanced Hebrew and Dutch speaking learners of English and were randomly assigned to one of three different tasks. The first task was to read a passage of text and then answer multiple-choice questions that required the knowledge of ten target words. The target words were highlighted and also in a glossary on the side of the text. Task 2 consisted of the same passage of text and questions as task 1, the target words were removed however and replaced with blank spaces. Laufer and Hulstijn (2001) then gave them a list of the target words to fill out these blank spaces with. In task 3 the learners were only supplied with the target words and were required to write an original text using these target words in the form of a letter to a newspaper editor. The two involvement load components of need and search were the same in all three different tasks, with the need component being moderate (1) while the search component was absent (0). The varying component in the three tasks, according to Laufer and Hulstijn (2001), was the evaluation component. Task one had a level of absent (0), task 2 had a level of moderate (1), and finally in task three the level was strong (2). This meant that the final involvement score of the three different tasks was 1 for task 1, 2 for task 2, and 3 for task 3. To test the learning of the target words Laufer and Hulstijn (2001) asked the students to give an L1 translation or English explanation for all ten of the target words immediately after the tasks had been concluded. They would then test the Dutch learners one week later, and the Hebrew learners two weeks later in the same fashion (Laufer and Hulstijn 2001).

The conclusion of this study was that the learners who did task 3 did better than those who performed task 1 and 2. This seemed to be in line with their idea of the Involvement Load Hypothesis, with more involvement on the part of the learner the better the vocabulary retention would be. Other studies have also been done testing the validity of this theory by Kim as well as Keating, both coming to somewhat the same conclusions as those in the original study (as cited in Lee and Hirsh 2012, p.86-87). We will now look at another study that investigates the effects of other tasks in addition to incidental vocabulary learning.

Hill and Laufer's Study on the Involvement Load

Monica Hill and Batia Laufer (2003) wanted to test the effects of involvement on the incidental vocabulary learning of learners by testing to see if the number of operations had any effect on how well a learner would retain a word. They also wanted to look at the amount of time these activities took, as it was noted in Laufer and Hulstijn's study (2001) that the students doing task 3 spent more time finishing the task than the other students (Hill and Laufer 2003, p.90-91). The sample consisted of 96 learners, who were all English learners going to the University of Hong Kong. All of the students in the sample were second language learners of English. Each of these learners were then randomly assigned to one of three different tasks relating to a text that all learners had read. The entire test was conducted by computer. The text consisted of approximately 93% words that were known by the learners, and the learners also had the ability to get a dictionary meaning of any word in the text should they choose to by selecting the word on the screen (Hill and Laufer 2003, p.91-92). In task 1 the learner was required to answer twelve yes/no comprehension questions regarding the text. To answer these questions the learner would need an understanding of the different target words. Task 2 required the learner to select one of four meanings of the different target words. The meanings consisted of high frequency words that should be known by the learners. Task 3 required the learner to select the correct target word from a list of four different target words when presented with a synonym or paraphrase of the target word. (Hill and Laufer 2003, p.93-94). The results of the testing showed that the learners who scored highest in both an immediate and delayed test were the learners who had done task 3, which was followed by the learners having done task 2, while task 1 scored the lowest of the three tests. Testing for significance they found that task 1 was significantly less effective than both the other two tests (Hill and Laufer 2003, p.99). Also wanting to see the amount of time it took to conduct the different tests they found that there were no significant differences between the tests when it came to the length of time it took to finish them, with all of the tasks taking a little over five minutes to complete (Hill and Laufer 2003, p.101). Finally Hill and Laufer (2003) looked at the amount of time the test-takers had clicked on dictionary definitions of words in the text. Again task 1 scored the lowest in this category, showing a significant difference between task 1 and 2, and task 1 and 3. There was no difference between task 2 and 3 however (Hill and Laufer 2003, p.100-101). To explain the efficiency of

the different tasks, and why tasks 2 and 3 scored better than task 1, Hill and Laufer (2003) used the amount of dictionary uses the different tasks encouraged (Hill and Laufer 2003, p.102). As task 1 saw fewer dictionary uses than the other two we can therefore conclude that the involvement load was higher in these tasks than in task 1. What was also noted in the study was that the learners conducting tasks 2 and 3 did not use a longer amount of time on the task than the learners in task 1. This was despite the fact that task 2 and 3 elicited more dictionary uses than in task 1. They speculated that the reason for this was that task 1 required the learner to not only focus on the target word but also on the sentence or sentences around it (Hill and Laufer 2003, p.102). Since the test with the highest scores was also the test with the highest amount of dictionary uses and the lowest scores with the lowest dictionary uses, this seemed to show "...that an important factor determining task effectiveness for vocabulary learning is the amount of word-related activity the task induces" (Hill and Laufer 2003, p.104). This seems to confirm the Involvement Load Hypothesis, that a more involving task will yield better results when it comes to vocabulary learning and retention.

Folse's Study on Involvement Load and Repetition

Another study looking at the Involvement Load Hypothesis was Folse's study in 2006 (as cited in Lee and Hirsh 2012). The results of his study differed from those presented above however. Folse tested 154 university students on vocabulary exercises with three different levels of involvement. One group was given one fill in the blank exercise, another three fill in the blank exercises, and finally one group was given one original sentence writing task. Finally a post-test was administered unannounced testing for the meaning of the target words and the usage of the target words in learner made sentences. The results of the test did not fit perfectly with the Involvement Load Hypothesis set forth by Laufer and Hulstijn (2001), as the students who performed three fill in the blank exercises scored higher than the students who were tasked with writing an original sentence. On the scale of the Involvement Load the original sentence writing scored a five for involvement while the three fill in the blank exercises scored a four. Folse claimed that this was due to the greater importance of repetition than the involvement the exercise took. Meeting a word more often would lead to greater

vocabulary retention than a more involved task without repetition (as cited in Lee and Hirsh 2012, p.87).

Folse's claim that repetition can lead to improved vocabulary retention is supported by others. Nation has claimed that repetition is useful not only to learn vocabulary but to know it well enough that it can be accessed and used by the learner. This repetition can be gained through different means. Long texts, for example, have shown to contain a great deal of repetition of vocabulary. The use of texts can be effective especially if the texts given to the learner are about the same topic area, providing much better repetition of vocabulary. (as cited in Matsuoka 2012, p.157-158). Some research has also been done on whether spaced or massed repetition is more effective. It seems that an initial massed repetition followed by a more spaced approach might be most effective according to Nation (as cited in Matsuoka 2012, p.158).

Lee and Hirsh's Study on Involvement Load

Wanting to further test the ideas that Folse claimed, Lee & Hirsh (2012) began a study looking at “Whether task type (i.e. the quality of exposure) or the number of tasks (i.e. the quantity of exposure; frequency) more effectively promotes vocabulary learning...” (Lee and Hirsh 2012, p.88). The study tested 131 year 8 students at a junior high school in Taiwan that were learning English as a foreign language (Lee and Hirsh 2012, p.89). Doing much the same as Folse they would give the students three different tasks, also calculating the involvement index according to Laufer & Hulstijn's (2001) Involvement Load Hypothesis which is shown in the following table.

Fig 4.4 Comparison of exercises in Lee & Hirsch's study

Condition	Conceptualisation	Involvement Load	Involvement Index
One multiple-choice exercise	Low involvement Single Exposure	Moderate need (1) No search (0) Moderate evaluation (1)	1+0+1=2
Three different multiple-choice exercises	Low involvement Multiple exposures	Moderate need (1) No search (0) Moderate evaluation (1)	1+0+1=2
One original sentence writing	High involvement Single exposure	Moderate need (1) No search (0) Strong evaluation (2)	1+0+2=3

(Laufer and Hulstijn 2001)

Twelve target words were selected in the study and the researchers made sure to avoid words that would be known already by the students. These words were then separated into three different groups A, B, and C. Each student was required to practice all three groups of words, with each group being randomly assigned to one of the three tasks above. Finally two unannounced post-tests were used to test the students' knowledge of the target words. These post-tests were adapted from the Vocabulary Knowledge Scale mentioned in section 3.3. One of these post-tests was administered immediately and the other was administered two weeks later (Lee and Hirsh 2012, p.93-95). The results showed that all three approaches helped with vocabulary retention, having significant effects on the vocabulary retention of the students. The approach that scored the highest was approach number two, using three different multiple-choice exercises. This result coincided with Folse's results from 2006 that showed that repetition had a greater effect on vocabulary retention than involvement load. However, on the second post-test taken two weeks later approach number three, writing an original sentence scored better than the other two approaches. In this test the amount of target words retained was far smaller however, and all three approaches did not significantly effect the retention of vocabulary (Lee and Hirsh 2012, p.104-105). While not significant, it is interesting to note that the more involving task scored highest on the second post-test. The involvement of the task might prove more important over time than the number of repetitions. The most important aspect of this study is that all three vocabulary tasks gave results, showing that both the involvement and the number of repetitions are effective ways of improving on incidental vocabulary learning.

Summary of Chapter

The studies shown in this chapter have all looked at the effects of incidental contact of vocabulary, and how this incidental contact effects the vocabulary acquisition and knowledge of the learner. The results of these studies have indicated that incidental vocabulary acquisition can be greatly improved through two methods. One by supplementing the incidental contact with vocabulary exercises, as shown in Paribakht and Wesche's (1997) study in section 4.3 in which the level of word knowledge can be greatly increased through vocabulary exercises. However, as it has also been shown in this chapter there is a second option that has proved to be effective. Using texts with repeating words will also allow for a greater vocabulary acquisition. As Folse concluded that the repetition of words can be more important than the involvement load of the exercises (as cited in Lee and Hirsh 2012 p.87). There is however an indication that both measures should be possible if given the necessary amount of time, which should in theory increase the vocabulary acquisition further. The fact that both of these approaches help improve on the incidental vocabulary learning is something that should be considered very important. Having the ability to improve upon the slow progress of incidental vocabulary acquisition with vocabulary tasks can be very valuable for CLIL.

In my own study I will be looking at the incidental vocabulary acquisition that should happen through CLIL. In theory CLIL should be able to satisfy many aspects of the Involvement Load Hypothesis and the potential for texts with a repeating vocabulary is also there. The conclusion from these studies indicates that the potential for vocabulary growth could be greatly improved by focused vocabulary teaching. Incidental contact with vocabulary will be far more effective if there is a dedicated focus on vocabulary learning and acquisition throughout the learning process.

5) Method

Introduction

In this chapter I present an overview of the design and process of the current study. I will begin by explaining how the study was planned and designed, explaining how I selected the sample, and the reasoning behind the research design I elected to use. I will then explain in detail the different elements to the study and the reasoning behind them. Finally I will explain the vocabulary tests and tools that I have used to analyze the vocabulary of the students. It is important to note that the present study was conducted alongside that by another master student. Some parts of the study were done jointly due to time limitations. This will be stated where this is the case.

Defining the Research Question

Prior to starting my master thesis I wrote a pilot study regarding the teaching of vocabulary in CLIL. With this study I started to wonder what effects CLIL instruction actually has on the vocabulary of the learner. My initial goal was to devise both pre and post-tests to examine the improvement that CLIL instruction had on the vocabulary of the students, this proved, unfortunately, to be too time consuming. With only a single year to write my thesis this would not give sufficient time between the required pre and post-tests to be of much use. I therefore decided to devise the research question presented in Chapter 1 instead. I was eager to test the students' vocabulary and discover what improvements, if any, the CLIL instruction had on their vocabulary. I therefore decided that I would need to find suitable ways to measure the vocabulary of the students, and settled on the tests that will be presented later in this chapter. I wanted to use a quantitative approach, but was eager to supplement this with qualitative data as well. I therefore went with a mixed-methods approach, which I will present below.

Selecting the Sample

The first step that is needed when selecting a sample is to identify what the target population is according to Ary, Jacobs, and Sorenson (2010). The target population for this particular study is the students currently in lower secondary school in Norway who are being taught through CLIL, as well as those students in EFL instruction who are not receiving CLIL, to be used as a control group. Since the goal of the study was to see if those students who had CLIL would score better on vocabulary tests than students who had EFL only, the ideal situation would be to test the students after they had been having CLIL instruction for some time. Unfortunately the number of classes who are taught CLIL in Norway is not very large. In addition to this, most of the research would need to be finished before January, this to give enough time to properly analyze and assess it before the master was to be finished. Most of the CLIL programs were only starting up then, so none of the students would have had CLIL instruction in previous years which meant that the students would only have CLIL for a limited time before they could be tested. I reached out to several schools asking if they were teaching CLIL and were interested in taking part in a research study. Since so few schools were teaching CLIL and the fact that the study would eat up time that could be spent teaching I could not afford to be too selective and therefore reached out to both lower and upper secondary schools in Norway. Finally I was informed of two lower secondary schools participating in a larger research project that were interested in letting me perform my research at their schools. The two schools were involved in a larger CORE project, looking at CLIL throughout Europe. The CORE project was to examine, how CLIL was implemented in various countries around Europe. These schools were both lower secondary schools and the classes were all from the tenth grade, the final grade of lower secondary school. This was very beneficial as this meant the entire group would be as homogenous as possible.

5.1.1 Details Regarding the Sample

The sample that was selected in this study schools in Norway. Two classes came from each school, one with CLIL instruction and one class that was not receiving CLIL instruction. All the classes had EFL instruction as well. One class had CLIL instruction in religion and the

other received CLIL instruction in social studies. The total CLIL instruction reported by both of these classes was approximately 30% of the hours the subject had, although neither teacher was very sure of the exact number. So to summarize:consisted of four classes from two lower secondary schools. All four classes were in the tenth grade, which is the final year of lower secondary school. Both of the schools are state run schools, as are the large majority of lower secondary

- The sample was comprised of four classes from two lower secondary schools
- There was one CLIL class and one EFL class from each school
- The total CLIL instruction in both CLIL classes was reported to be about 30% of the total hours of the subject in each class.

The following table shows the number of students that took part in each of the four classes, each of the two schools taking part in the study supplied two classes, one CLIL class and one EFL class. This shows the number of students who completed the X-lex, 200 word text, and survey, as well as showing how many students completed both the X-lex test and the survey.

Table 5.1 Total Numbers of Students who have Completed Tests

	X-lex Tests	200 Word Texts	Survey	Both X-Lex and Survey Completed
CLIL Students	36	17	37	35
EFL Students	38	20	36	36
Total Number	74	47	73	71

As can be seen 73 students completed the survey and 74 students completed the X-Lex vocabulary test. In addition to this, 47 students had usable texts that were analyzed. All of these students completed the X-Lex test while 46 students completed all of the different tests. The number of students who completed both the X-lex and survey numbered 71. The number of students who completed at least one part of the study was 78. The number of students in a CLIL class was 38 while for EFL it was 40. In addition to this, School A supplied 49 students while school B supplied a total of 29 students.

The Research Design

When considering the research question I kept in mind that the main focus was on the vocabulary tests that I would use during the study. However, lacking the necessary time to conduct pre- and post-tests I wanted to supplement the data collected from the vocabulary tests with other data. Being only able to test the students' vocabulary knowledge at a single point in time, I knew I would need additional data to shed some light on the results of these tests. These additional data I wanted could be gathered primarily in the form of a survey. This survey would be used to gather some general information about the students and will be discussed more in detail later in the chapter. In addition, the teachers of the two CLIL classes were interviewed along with a small random sample of their students, which will be discussed later in this chapter. I also wanted to know more about the teachers approach to teaching CLIL and therefore settled on a mixed methods approach to the study. I will in the following section discuss the merits and challenges of using a mixed methods approach.

5.1.2 Strengths and Weaknesses of a Mixed-Methods Approach

Conventionally, the two approaches, quantitative and qualitative have been separated into two distinct fields. A mixed methods approach seeks to combine these two. The quantitative and qualitative approaches have been split in part because they, traditionally, have been seen as different ways to understand and interpret the truth (Calfee and Sperling 2010, p.8). A mixed-methods approach rejects the belief that quantitative and qualitative approaches to research are incompatible with each other. I knew that I would need a quantitative approach in my study to test the students' vocabulary, as I was doubtful that the students would be able to identify and reflect sufficiently on their own vocabulary if I used a qualitative approach. However, using only a quantitative approach might leave many questions unanswered. Therefore, using a qualitative approach as well would allow for a better analysis of the quantitative results of the tests. Mixing the two methods together can allow the two forms to interact with each other, strengthening both (Calfee and Sperling 2010, p.9). I therefore settled on using an embedded design for my study. An embedded design is when one data set is used to support another set of data. The reason for doing this is that some aspects or

questions might be unanswerable using only one form of data (Ary et al 2010, p.564). In this study the interviews with the teachers and students as well as the survey are used to support the data collected in the vocabulary tests. It was therefore of interest to use a mixed methods approach. This, however, is not without its problems.

The two main problems facing a researcher wanting to use a mixed methods approach are:

- Time commitment; conducting a mixed methods approach is a time consuming endeavour (Ary et al 2010, p.568), the larger amount of different types of data that is to be collected will necessitate a greater time commitment, and the research will have to be conducted in an efficient manner.
- Skills in both Qualitative and Quantitative research; Being able to analyze both quantitative and qualitative research is a demanding task, especially when you are supposed to integrate the two approaches to reach a single goal (Ary et al 2010, p.568). Creating, conducting and analyzing a mixed methods approach can therefore be a challenging task.

While it is demanding to conduct a study using a mixed methods approach, it does have several strengths that can be very valuable, it:

- Removes the weaknesses of the two approaches; if conducted correctly a mixed methods approach can negate the weaknesses of a quantitative or qualitative approach by means of using the other to compensate (Ary et al 2010, p.567).
- A more complete picture; using a mixed methods approach the researcher can spot aspects that might have been missed if using only a single approach. This combination can provide a more accurate image as a result (Ary et al 2010, p.567).
- A mixed methods approach can give a more valid interpretation than what might otherwise be possible (Calfee and Sperling 2010, p.11).

As mentioned above, in the present study I saw it as being necessary to supplement the data collected from the tests with interviews, especially of the teachers, as the CLIL instruction might vary and these interviews could clarify the results of the tests that have been conducted. Settling on an embedded design for my mixed-methods approach I felt that this would ease the difficulty in analyzing the different sets of data, and allow the use of the interviews to supplement and explain what is seen in the test data.

Testing Vocabulary

When attempting to measure vocabulary knowledge, there are two different approaches that have seen frequent use in testing. In this next section I will look at the differences and issues with these two approaches, as well as looking at some additional problems and issues with vocabulary testing. Finally, I will discuss in depth the tests that I have elected to use, and the reasons why I have settled on using these tests for my study.

5.1.3 Breadth Testing

There are two main methods for testing vocabulary, breadth and depth tests, in this section I will look at breadth tests. These tests are intended to test the quantity of vocabulary that the learner knows. These tests will often take the form of simple checklists, where the learner checks if they know a word or not. The advantages to such a test are the speed at which they can be produced and executed, and the large amount of words that can be tested. The greatest advantage of this is that the test should be more reliable than those testing fewer words, and less prone to randomness as a result (Milton 2009, p.72). These tests are usually designed from a corpus such as the General Service List or BNC. The words are then selected from different frequency bands within depending on what it is meant to test (Read 2000, p.87). The major disadvantage with the checklist test are that it does not measure how well the learner knows the word, only that the student identifies it as a word, and this in turn creates a problem with validity. There is no way of knowing if the learner guessed or not. This problem can partially be avoided by including false words within the test such as in Meara's X-Lex test (Milton 2009, p.73). Another widely used test to measure vocabulary breadth is Nation's

Levels Test (Milton 2009, p.74). In this test learners are given a list of words and must match them to a description. This test does test a somewhat deeper knowledge of the word, however, it requires more than an understanding of the target word if a learner does not understand the description. In addition to this some words can be guessed at through a process of elimination (Milton 2009, p.74-75). Overall the measurement of vocabulary breadth allows for a quick and simple overview of the learner's vocabulary knowledge and has also been shown to correlate with examination grades and CEFR levels in language proficiency in English. Meara & Milton have compared results from X-Lex tests with the CEFR suggesting that tests of vocabulary size can be compared favourably to CEFR levels, giving accurate predictions (as cited in Milton 2009, p.191). Furthermore, Nation's Vocabulary Levels Test has been used in New Zealand as well as other countries to test immigrant learners when they arrive in the country (Read 2000, p.118).

5.1.4 Depth Testing

The measurement of vocabulary depth is, in contrast with breadth measurement, still in its infancy. Depth tests how well the learner knows certain words, testing them on several different concepts of word knowledge. While Breadth tests have often been criticized for only giving a basic or superficial level of vocabulary, depth tests are meant to remedy this by going deeper, asking for more knowledge from the learner. This depth is often measured along the lines of Nation and Richard's definition of word knowledge (Read 2000, p.93). We will now take a look at an example of a depth test, namely Paribakht and Wesche's (1997) Vocabulary Knowledge Scale, or VKS, mentioned in section 4.3. The VKS was designed to test the knowledge of specific words as a result of several reading activities. The learner is given a list of words together with five self-report categories. The VKS elicitation scale tests the learner in several aspects of word knowledge, in addition to requiring them at the highest category, category V, to be able to produce a sentence containing the word. There are however problems with the scale, firstly it does not take into account multiple meanings of words, although this could be avoided by selecting specific words not subject to multiple meanings. The other issue is that category V might not give a good indication of use in practice, as the learner can produce a sentence that is neutral, giving little indication as to whether or not the learner has a good grasp of the words usage (Read 2007, p.136-137).

5.1.5 Receptive and Productive Vocabulary

As mentioned in section 3.3, the degree that you know a word can differ greatly, there is a difference between what a learner can understand and what they can produce themselves in reading and writing. This is referred to as a learner's receptive and productive vocabulary. This is a difficult thing to measure in vocabulary tests, as many will fall short when attempting to measure the productive vocabulary of the learner. The ideal productive test would allow for the learner to produce a word by retrieving it themselves, instead of being given a word and asked for a meaning (Schmitt 2001: 98). If we look at Nation's table for knowing a word in section 3.3, we can see that there is a distinction between what a learner can understand and what a learner can actively use. The Vocabulary Knowledge Scale of Paribakht and Wesche (1997) attempts to test for this to some extent, although not perfect it perhaps comes closer to measuring productive vocabulary than other tests (Melka 1997, p.99). The use of checklist tests, however, does not properly measure productive vocabulary. There are however estimates of the relationship between receptive and productive vocabulary, with receptive vocabulary being approximately double that of the productive vocabulary of the learner, though there are differing reports on the subject (Melka 2001, p.92).

The Vocabulary Tests

In the following table (table 4.1) I list up the pros and cons of the breadth and the depth test to better examine them. Summing up the strengths and weaknesses of the two approaches we find the following:

Table 5.2 Comparison of breadth and depth tests

Breadth Test	Depth Test
Pros	Pros
Gives an overview of the total vocabulary knowledge of a learner.	Gives in-depth information regarding the degrees of knowledge a learner has.
Is quick and takes little time.	Is a better measure for productive vocabulary.
Allows for the testing of many items becoming more reliable as a result.	Can test for the learning of specific terminology or words.
Cons	Cons
Only tests for a basic understanding of a word.	Is unable to test for a large number of words, thereby becoming more affected by chance.
Can suffer from overconfidence in the test taker	Is quite time consuming

Looking at these pros and cons of the two test types, and considering my own study, I came to the conclusion that I would need to utilize a breadth test over a depth test. The reasoning behind this was that, as mentioned above, I was interested in seeing if CLIL students would score better on vocabulary tests than EFL students. Using a depth test would make the study vulnerable to chance, where some students would know a word better than others, a breadth test would limit this issue. In addition to this, time was, again, a factor in my decision. Breadth tests are much quicker to conduct, while a depth test is more complex demanding more time and energy from the respondent. With this decision made I settled on using the X-lex test by Meara, and I will be discussing the reasoning why in the following section.

5.1.6 The X-lex Test

The X-lex test is a vocabulary test that is freely available from the website <http://www.lognostics.co.uk>. Originally designed by Paul Meara and colleagues and called the eurocentres' vocabulary size test (Read 1997, p.312-313) it has been further developed into the computerized X-lex test that is available online. The computerized version of the test presents the learner with 120 words, 20 words each from five different vocabulary lists. These lists detail the 5000 most common words in the English language taken from West's General Service list. The other 20 words are pseudo-words designed to test for overconfidence in the

learner when taking the test. For each correctly known word the learner is scored 50 points, for each pseudo-word, or false word that is claimed to be known, the learner is subtracted 250 points. The test has been used successfully in several places, showing promising results. The X-lex test was used in a study examining the vocabulary of Greek learners of English, tracking the development of their vocabulary through several years of study with promising results (Milton 2009, p.79-80). Meara has noted that the two main issues with the test have been that it tends not to work well with low-level learners of English, nor with French learners of English. A final issue is that the scores of some students can become too low as a result of claiming to know too many false words (Read 1997, p.313). In my study only one of these problems was seen as problematic. As the students were in the tenth grade I could assume that their level of English would be adequate for the test. While there has been no use of the X-lex test to my knowledge on Norwegian learners, I could safely assume that the issues plaguing French learners would not affect Norwegian learners due to the difference between the languages. The final problem, students claiming too many false words, was however something that needed to be considered. In my results section I solved this by simply removing the students who claimed to know five or more false words. This number was rather low as well, so this particular problem did not seem to affect this study too much. The score is not the only important aspect of the X-lex test, as the test tracks the knowledge of each 1000 word frequency band in the five thousand most frequent words it is important that learners score more at each band than they will score on the following frequency band. The stronger the student the more complete each frequency level becomes. It is no guarantee that every test-taker should have a profile that looks like this, instead it is useful as a baseline to make sure that the learner's profile is not too skewed in favour of higher frequency words.

To be able to use the X-lex test I was required to alter it slightly as the students would not be able to complete the test on the computer as is normal. Instead I selected a sample test from the X-lex program and printed this out in a table with where the students could check off in one of two columns for each word. The columns were 'I know this word' and 'I do not know this word or am uncertain of this word'. All the test results were then calculated after the students had handed them in. All students answered on all of the words and the test took a maximum of ten minutes to complete, with some students completing it in just five minutes.

An example of the test can be seen in the appendix (appendix 1) showing the complete list of words as presented to the students.

5.1.7 The Text Analysis

The secondary part of the vocabulary testing I wished to include in my study was to look at the written production of the learners and compare this to the texts the classes had read.

Unfortunately as the study was being conducted it became clear that there was little English written material available to the CLIL classes, and this part of the study was scrapped as a result. I did however, still obtain essays that were written by the students, and decided to include these in the study as they might give some interesting information. The text analysis was done using 200 word text samples taken from written material that the students wrote, and looking at the type/token ratio as well as the amount of different frequency words that were being used. The written material from school A was taken from an examination they had, while the written material from school B was taken from a test the CLIL teacher held for both the EFL and CLIL students. Therefore the texts for CLIL and EFL classes were the same, although the texts between schools were different. It is important to note however that the final tally of students that supplied these essays was very small. With such a small sample little weight could be given to it when analyzing the results, I did however elect to keep the results as they are still an interesting part of the study. The texts were analyzed using Paul Nation's Range program which is freely available to download from <http://www.victoria.ac.nz/lals/about/staff/paul-nation>. Nation's Range program is primarily used to test the suitability of texts for use in English classes. The initial reasoning behind using this program for analyzing the texts the students had written was to be able to compare them to the texts they have read. All the texts were prepared as instructed by the range program, double-checking to avoid any mistakes. The range program would then be run, and would output the type/token ratio as well as the type and tokens of three different word lists making up the most common words of the English language in the two first lists, while the third list includes words that are common in upper secondary school and university. Any words that are not included in these three lists will be listed as other in the results. Laufer and Nation argue that this Lexical Frequency Profile (LFP) can be used as a measure of vocabulary knowledge and size. (as cited in Read 1997, p.314)

The reasoning as to why I wanted to include an analysis of the students' written texts was originally that I wanted to look at their productive vocabulary in contrast to the written material they were being given. This would give an indication of the relationship between implicit learning and productive output, which could have proven interesting had it worked. As stated earlier this did not come to fruition, and given the small amount of texts produced that were comprised of at least 200 words, it could be stated that this part of the study was unsuccessful. I do however believe that the analysis of written materials could be very useful, the issue is that this is a very complex and difficult subject, as grammar issues can be present that are not counted by a program such as Range. The testing of vocabulary in it's "natural environment" is something that I think would be extremely beneficial and would provide for a level of analysis that is not available in either breadth nor depth tests. In the next section I will discuss the interviews that were included in the study, which were used to supplement the data collected from the vocabulary tests.

The Interviews

All of the interviews were conducted during the span of two days. The interviews with the CLIL teachers were done separately, while the interviews with the students were done in groups of three students from each CLIL class. I had initially decided to interview the EFL teachers and students as well. In the end I elected to not do this, as I was interested in the CLIL classes, and the EFL classes were only there as a control group. The interest of the interviews was to shed some more light on the teaching and possible effects of CLIL. I was not looking at the effects of vocabulary teaching in EFL. Another important point as to why these interviews were in the end not included, was that the EFL teachers of the CLIL students were not interviewed, and it would not be possible to state one way or another if the EFL control classes in the study had received better or worse teaching in EFL than the CLIL students. The student interviews were conducted while the other researcher, who was mentioned at the beginning of the chapter, was present as well. This researcher did not interject or ask any questions during my interview, nor have I used any part of her interview

in my own study. All of the interviews were taken with an audio recorder and later transcribed. Also, all of the interviews were conducted in Norwegian, as this was preferred by the people being interviewed.

5.1.8 The Teacher Interviews

Before I began writing my master thesis I, as mentioned above, did a pilot study looking at the teaching of vocabulary in CLIL. In this pilot study I interviewed teachers who taught CLIL about their vocabulary teaching in their class. In these interviews I was especially interested in how much explicit and implicit vocabulary teaching they conducted, and how aware they were of vocabulary learning in CLIL. I discovered through these interviews that finding the right materials was a great challenge for them, especially with regard to vocabulary, something I had not properly considered beforehand. Therefore I chose to alter my interviews somewhat to give more focus on the selection of materials in CLIL. I still focused much of the interview on how they taught vocabulary, implicitly and explicitly, and how aware they were of it, and how important they felt vocabulary was in CLIL. Wanting to supplement the data that would be present in the vocabulary tests I felt it was important to allow the teachers to describe their process and thoughts when they considered vocabulary in their class. It was important to identify their views on this and how they felt they could best include vocabulary teaching in the class. The interviews were also meant to give the teachers a chance at giving their thoughts on CLIL in general as well. I therefore settled on using a semi-structured interview guide. A semi-structured interview would allow the teachers to give insight into what they felt was important with CLIL and with vocabulary, with questions that could not simply be answered by a yes or no. A semi-structured interview would also give me a chance to expand on topics that could come up (Ary et al 2010, p.438), allowing for a depth that would be important to the study. The interview guide (see appendix 2) was divided into four main sections, the first section included background questions regarding their teaching experience. The second section asked general questions about CLIL. The third section of the interview guide contained questions regarding the use and selection of materials in the CLIL class. The final section investigated the vocabulary teaching in the class, and their views on vocabulary in CLIL. Finally they were asked if they had anything else that they wanted to add

to the interview. It was important to give them a chance to explain their thought process rather than getting canned responses, which is why the loose interview form was used.

5.1.9 The Student Interviews

I had already done a pilot study, as I had with the teacher interviews, where I had interviewed students about vocabulary in CLIL. In this pilot study I interviewed the students in groups of three asking them questions relating to their CLIL instruction and vocabulary teaching. I had in this interview asked questions regarding their thoughts about CLIL and if they felt their vocabulary had improved much during their CLIL instruction. This proved not to work too well, the students would give answers, that often contradicted those of their teacher, one of the problems when interviewing students is that it is difficult to know how aware they are of the teaching they are receiving. How trustworthy are they really? Asking them too complicated questions could result in incorrect answers. The pilot study was very helpful in giving me an idea of the type of questions they could and could not answer effectively. I therefore altered the interview from the pilot study to better facilitate the type of responses I would get. Asking questions mostly about their use of English outside of school, and their thoughts on their own English education. In these interviews I also wanted to use a semi-structured interview guide, while it might be simpler with a more structured form of interview the purpose of the interviews was to give thoughts and explanations that would be impossible to detect with a survey or through other data that I collected. Selecting the students was done by picking three students at random, also making sure to include at least one boy and one girl, from the students that had volunteered to take part in an interview. The interview guide (see appendix 3) consisted of three main topics. The first section contained some very basic information about the students, asking if they had any English-speaking family and what class they were in. The second section contained questions asking about their use of English outside of school. The third section asked what they would do when encountering an unfamiliar word in English and general questions regarding vocabulary in CLIL. Finally they were, just like the teachers, given the opportunity to add anything that they thought important. The students were asked the questions as a group and I left it mostly up to them if they wanted to answer or not. Some questions were answered by all three, others by only one or two. Only when I felt that it was important would I push for an answer.

This method of interview was perhaps a mistake, interviewing three at a time meant that the answers became much shorter and less in depth than ideal. It was also more difficult to follow up on interesting things the students would say, while I attempted to do this as best I could there were a few times where it would have been nice to have the students expand on what they said. While perhaps more daunting for the students it would probably have been a better choice to interview them one at a time. However, time constraints were again a factor, and with limited time it was difficult to interview three students separately. Only interviewing one student more in depth would be problematic as well, as the answers given would vary depending on the student being interviewed. In the end I feel that interviewing three students at once was perhaps not ideal, but the most practical approach to take.

The Survey

The survey was done jointly with the other master student mentioned earlier in this chapter. This was done mostly due to time constraints, since it was more efficient to conduct a joint survey of the classes rather than have separate surveys, especially considering the amount of questions that were the same. The parts of the survey (appendix 4) used by me were the general section, and the questions regarding unknown words, the questions included in these were questions from 24-42. The most important aspect of the survey was to gather general information about the students. Therefore questions were asked about their background in English and what grades they had received. All of the general questions were used by myself and the other master student. The grades were of vital importance to the study, as without them it would not be possible to say anything meaningful regarding the vocabulary scores whatsoever. The students were also asked some questions regarding unknown words. These questions were designed to question the students on what strategies they employed to deal with unknown words when they encountered them, and perhaps most importantly how often they encountered them. The survey was directly administered to the students and the questions regarding vocabulary were closed-ended questions, which is the ideal form of questions if at all possible. This would allow for easier tabulating of the results as well as

allowing for faster more accurate answers (Ary et al 2010, p.391-392). The questions regarding vocabulary consisted of scaling items, with five different responses ranging from never to always. As these questions related to how often they employed specific strategies in dealing with unknown words it was important that they were given enough range to give an accurate response. My part of the survey was not very long, which was important because of the time it would take to complete the entire survey, which also included a reading test the other researcher was conducting. The total length of the survey and reading test was approximately 30 minutes. The main goal of the survey was to give some insight into the students work habits regarding vocabulary. The survey data would be useful in explaining the results of the vocabulary tests that were also conducted.

Collecting and Processing the Data

The ideal situation in a study such as this is that the researcher or an assistant is at hand when the surveys and tests are being conducted. Thankfully, for the most part this was feasible in this study. The surveys and X-lex tests were all supervised by myself or the other master student. Since there were several classes and limited time in which we could conduct our research we needed to split up between the classes at certain times. This was only done at school B however, in School A we were able to both take part in all parts of the surveys and X-lex tests. The written texts of the students were done without myself or an assistant being present, as this was not possible in the time we had available. The texts from school A were finished prior to my arrival at the school and were collected by hand, while the texts from school B were sent over by e-mail one week later. While it was less than ideal that the students wrote these at different times, it was the only way of collecting the data. All the tests and surveys contained brief but clear instructions and the students were walked through them prior to them answering them. I or the other researcher was also at hand to answer any questions they might have had.

5.1.10 Processing and Interpreting the data

The data was coded into SPSS 20, this included the data from the survey, as well as the results of the X-lex test and Range analysis of written texts by the students. All of the data was cross-checked several times to ensure that it was plotted in correctly. While there was a lot of different data in the study, I was able to keep the analysis fairly simple. Looking mostly at the frequency of results I could garner a fair amount of information from just this. I also used bivariate correlations using Spearmans rho, which proved extremely useful in comparing the different sets of data. I correlated the questions regarding unknown words with each other, finding the most interesting results from this, and most importantly I could correlate the results of the vocabulary tests with the survey data. This was done by looking at the correlation between the test scores and grades for instance. The X-lex test was calculated using the method described in section 5.6.1 and the final score put into SPSS, the texts were also analyzed using Range as described in section 5.6.2 and also put into SPSS. The main part of the study, seeing if CLIL students scored better than those with EFL only on vocabulary tests was fairly straightforward, and could be done by examining the mean scores of the classes on these tests. The ancillary part of the study was in analyzing why they scored the way they did which included the results of the survey and the interviews. These two parts were conducted separately and then brought together to shed light on the results of the vocabulary tests. The interviews were also related to separately at first before including them in the analysis of the quantitative data.

External Validity

It is necessary to look at the external validity of any study, such as this one, which incorporates quantitative data. The external validity refers to the extent that the findings of the study can be generalized to apply to the rest of the population (Ary et al 2010, p.292). Any such discussion will need to look at how applicable the sample is to the rest of the population that you are comparing it to. The population for this study is considered to be the students in upper and lower secondary school in Norway who are taking both EFL and CLIL, and those

students who are taking only EFL. The main threats to the external validity of this study are the following:

- The representativeness of the sample
- Non-respondents

I will in the following two sections discuss and address these problems, starting with the representativeness of the sample.

5.1.11 Representativeness of the Sample

In this study the selection of participating schools was done on a take what you can get basis. As there are few schools in Norway who have CLIL programs currently it was necessary to take what was available. How representative a sample is of the general population that you are comparing it to is important for the significance of the study. Ary et al (2010) points out that:

The results of a reading study that used first-graders enrolled in an affluent suburban school district as subjects might not be the same if first-graders in a rural school district had been the subjects.

(Ary et al 2010, 292-293)

This matters to the present study as both schools are from the same smaller, rural district in Norway and therefore might have other characteristics than that of other schools. Another more important issue is that the students in this study are all in lower secondary school, which means that the sample does not really reflect students in upper secondary schools. The results might have been different had the study been performed on upper secondary school students rather than lower secondary school students. In addition to these concerns, the sample size of the study is worth mentioning. The current sample is not a very large one, and it is therefore necessary to take the results with a great caution, and they are probably not representative for other students than those in the sample. There are however things that strengthen the representativeness of the sample. The students are not volunteers, as is the case in many other

CLIL programs. This strengthens the validity of the study, as the sample is not made up only of students who, for example, are interested in improving their English, or are already very skilled in English.

5.1.12 Non-respondents

This study included three separate components that the students were required to complete. These three components consisted of the survey, the vocabulary test, and the written text. The number of students who completed all three tests was not very large. Especially large was the number of non-respondents to the written text. This has meant that the actual sample size is made smaller and also that the non-respondents can have affected the greater generalization of the sample. The non-respondents could consist of very weak students. As the cut-off point for my text analysis was 200 words, this meant that any texts below this word count could not be used in the study, which removed a decent portion of texts from the sample. We can therefore assume that some of the weaker students when it comes to productive vocabulary did not produce texts. This means that the sample taken from the text analysis might be less applicable to the general population. This has been taken into consideration when analyzing the data. The text analysis has been included, but given little weight in the final analysis and discussion. Both the survey and vocabulary test included a few non-respondents, but this number was far smaller, making the total sample of the survey and X-lex vocabulary test sufficiently large.

5.1.13 Conclusion of External Validity Discussion

The main weaknesses of this study are the fact that both schools are from the same district and that they are lower secondary schools. This threatens the external validity of the study somewhat, and makes it more applicable to lower secondary schools than both lower and upper secondary schools teaching CLIL. The amount of different components in the study has been problematic, since some components have not been completed by all students. This is most applicable to the text analysis, and this has not been considered as heavily in the final analysis as a result.

The final sample of the study is therefore not very large, nor is it applicable to all of the population, which means that there are some very real threats to the external validity of this study. To a certain extent however, it can still be used as a discussion point on the effects of CLIL on vocabulary.

6) Results and Analysis

Introduction

In this chapter I will be presenting the test results from the two schools that participated in the study, as well as looking at some answers from the survey that was conducted. Finally I will present the data from the interviews of the teachers and students, using them to supplement the quantitative findings. I will start by looking at the sample in greater detail.

The Sample

As mentioned in section 5.3.1 the sample was selected from two schools, each supplying one CLIL class and one EFL class. The table 5.1 is presented again below showing the amount of students who completed each part of the study. This includes the X-lex test, the students' texts and the survey.

Table 6.1 Total Numbers of students who have completed tests

	X-lex Tests	200 Word Texts	Survey	Both X-Lex and Survey Completed
CLIL Students	36	17	37	35
EFL Students	38	20	36	36
Total Number	74	47	73	71

The most important number here is the number of students who completed both the X-lex test and the survey. Out of a total of 78 students who completed at least one part of the study, there were 71 who completed both the X-lex test and the survey.

X-lex Results

As stated in section 6.1 a total of 74 students completed the X-lex test. This test shows the general breadth of vocabulary among the students. In this section I will be looking more closely at these results and what they can tell us about the students' vocabulary. We will be seeing if the CLIL lessons have had a positive effect on the students' vocabulary knowledge compared to the EFL classes.

The X-lex test is calculated by adding up 50 points for each known word, and subtracting 250 points for each false word claimed to be known by the participant. In the following table 6.2 we can see the calculated scores, having subtracted from the score the number of false words claimed to be known, by students. The table shows the scores for the CLIL classes, EFL classes and the CLIL and EFL classes combined. This table also contains the mean grade that the students answered about their grades in written English. The scores can be thought of as a representation of the total vocabulary of the student. For example, a student with a score of 3500 words can roughly be thought to have a vocabulary consisting of 3500 words. The X-lex test measures words at each 1000 word frequency band, and it is important that students have a frequency profile that approximately fits the standard profile. Meaning that students know more words at from the first 1000 words than they know words from the 2000 word frequency band and so on (see section 5.6.1 for more detail).

Table 6.2 X-lex scores for CLIL and EFL students

X-Lex Calculated Scores	Mean	Median	Upper Value	Lower Value	Std. Deviation	Total Number	Mean Grade
CLIL	3402,7	3550	4950	1450	900,23	36	3.76
EFL	3607,69	3750	4800	1300	803,33	38	4.28
CLIL+EFL	3507,89	3675	4950	1300	852,41	74	4.01

As we can see from the table above, the mean score for the CLIL students is actually lower than that of the EFL students. This is somewhat disappointing, as I hoped that CLIL would

yield superior results to EFL. This can probably be explained by the EFL students having a higher grade average than the CLIL students. I will be return to these results later in the chapter.

As mentioned above the X-Lex test uses a series of false words that subtract 250 points from the score of the participant if they claim to know them. Those students who have claimed to know a large amount of these words will have somewhat unreliable results as they might have been guessing about many of the words. Removing the students who claimed to know 5 or more false words gives us the following results:

Table 6.3 X-lex Scores without Students with 5 or more False Words

X-Lex Calc. False Positives Removed	Mean	Median	Upper Value	Lower Value	Std. Deviation	Total Number	Mean Grade
CLIL	3455,88	3650	4950	1450	914,76	34	3.79
EFL	3755,71	3800	4800	2300	655,16	35	4.45
CLIL+EFL	3607,97	3750	4950	1450	802,19	69	4.11

The results are now skewed even further in favour of the EFL students with regards to scores. This means that most of the students who claimed to know 5 or more false words are to be found amongst the EFL students. We can also see how the grade average has increased when we removed the students claiming to know too many of the false words. This suggests that some weaker students have guessed more when selecting words than the stronger students. Furthermore, when checking the correlation between the grade average and the number of false words we find a weak negative correlation $r=-.25$, $p<05$, $N=69$. This seems to indicate that the stronger students are less likely to guess when selecting words, or that they have better strategies for recognizing these false words and separating them from the real ones.

Difference Between Classes

One important distinction to consider is the differences between the classes. A total of four classes participated in the study. The two schools have not cooperated, and there could be differences in the teaching of the two sets of CLIL classes. The level of the English in the different classes might be different as well. In table 6.3 below I look to see if there are individual differences between the CLIL classes, and also including the EFL classes to see if there is one class there that is skewing the results of the X-Lex test. As in table 6.3 the students claiming to know 5 or more false words have been removed from this sample as well.

Table 6.4 X-Lex Scores Sorted by Individual Classes

X-Lex Scores for Classes	Mean	Median	Upper Value	Lower Value	Std. Deviation	Total Number	Mean Grade
School A CLIL	3490,48	3650	4950	1550	872,16	21	4.05
School A EFL	3719,57	3800	4800	2300	720,12	23	4.57
School B CLIL	3400	3550	4700	1450	1013,86	13	3.38
School B EFL	3825	2800	4800	2900	530,65	12	4.2

As we can see the table indicates the same result as in table 6.3, with both EFL classes scoring better than the two CLIL classes, both EFL classes have better grades as well. Interestingly, the CLIL class from School A has a slightly higher grade average than the CLIL class from school B without having much higher X-Lex scores. The EFL class of school B is also greatly affected by removing the students who claim to know too many false words, with these students included in their sample the class has a mean score of 3446,88, which is almost 400 less than without these students. The mean grade of the class also goes up when removing these students as has been seen before when removing the students claiming to know too many false words. With them the class has a mean grade of 3,93, which is somewhat lower

than in table 6.3 above. However, this sample is very small as there are only 16 students that are included in the sample for the EFL class from school B. Removing the students who claim to know too many false words brings this down to 13, which could explain why the numbers change so significantly. It is interesting to note however that the EFL classes selected more false words than the students in the CLIL classes did.

To summarize the EFL students scored better by a fairly large margin, the simplest explanation for why this is would be the grade average being better for the EFL students. It seems as if the CLIL instruction has not had any noticeable impact on the scores of the CLIL students compared to the EFL students. However, we have no pre-test to compare our results to so we cannot say for certain that the CLIL instruction has had no impact on the vocabulary scores of the students in the CLIL classes.

Text Results

In this section I will be taking a look at the other vocabulary test that I employed to see if the results will show anything different from the X-lex test above. Analyzing texts written by the students I wanted to look at the productive vocabulary of the students. The texts were 200 word samples from material that the students had written analyzed using Nation's RANGE program. This gives a type/token ratio as well as separating the words used into separate lists. These lists are comprised of different frequency words, with list 1 containing the most frequent, list 2 second most, and list 3 less frequent. The final list is of words that do not belong to any of the lists. This could be names or very specific terminology, for example. The texts that were analyzed were different for the two schools, they were not however, different for the two sets of classes. The EFL and CLIL class of each school wrote texts on the same topic. In the following table I have therefore chosen to group the two CLIL and the two EFL classes together as the sample size becomes rather small when looking at the individual classes as to be too unreliable. In the table we can see the scores for the CLIL students and the EFL students as well as their combined scores. The table shows the percentage of different

words being used according as well as the type/token ratio of the texts. All the texts consist of approximately 200 words.

Table 6.5 200 Word Text Analysis using RANGE taken from written texts of the CLIL and EFL students

200 Word Text Analysis	Type/Token Ratio Mean	List 1 Words Mean	List 2 Words Mean	List 3 Words Mean	Other Words Mean
CLIL	49.18%	82.09%	6.28%	3.82%	7.82%
EFL	52.20%	82.86%	6.27%	3.65%	7.25%
CLIL+EFL	5.11%	82.58%	6.27%	3.71%	7.45%

Looking at these results we can yet again see favourable results for the EFL classes, with a mean score of 52,2% type/token ratio compared to the CLIL classes with a mean score of 49,2% type/token ratio. This 3% increase is not very great however, but it might indicate that the CLIL students do not have a stronger productive vocabulary. In addition to this, the number of words taken from the different lists is almost exactly alike in the two groups. There is a slight increase in the use of list 2, 3 and other words for the CLIL students, but this increase is too small to be significant. This is especially disappointing considering that the students from School B wrote texts relating to the subject that the CLIL students had been working on. With all the students from school B writing texts relating to the subject the CLIL students had received CLIL instruction in we could expect better results from the CLIL students at this school compared to the EFL students at the school. If we separate the students from School B out to look at only the results from them we find the following.

Table 6.6 200 word text analysis using Range on the students from school B only

	Type/Token Ratio Mean	List 1 Words Mean	List 2 Words Mean	List 3 Words Mean	Other Words Mean
School B CLIL	47.31	75.2	7.2	5.51	12.11
School B EFL	48.66	78.26	5.72	5.22	10.81

There are two interesting things to look at in this table. The first is that even though the CLIL students are able to write about a subject they should have a better vocabulary in, they still have a lower type/token ratio than do the students in the EFL class. The second thing to note

is that while they have a lower type/token ratio they also have a lower amount of words taken from word list 1. Furthermore, while the CLIL students have used a larger amount of words that are outside the three word lists. The mean scores from the Other Words list are a positive sign that the CLIL students have gained, if not a greater vocabulary, perhaps a vocabulary that is more suited to the task at hand. However, this sample is too small and the difference too insignificant to be of much value. Perhaps of greater interest is that these students have a somewhat lower type/token ratio than the average. They have however used a greater number of list 3 words in their texts than the mean score for this in table 6.5. This might be an indication that the assignment demanded a more challenging vocabulary to be used – which then resulted in less variety in the text as a result. This seems to suggest that the type of assignment could affect the vocabulary used in it.

The Survey Questions

In this section we will look more closely at the answers given in the survey regarding how the students dealt with vocabulary issues, and which strategies they would employ to deal with unknown words. In table 6.7 the answers of the CLIL and EFL students are presented. I will start by briefly talking about the results of each of the classes separately, starting with the CLIL students, before I compare and contrast the two groups. All of the questions had five different answers ranging from never (1) to always (5).

Table 6.7 Comparison of Answers to Questions Regarding Unknown Words Between EFL and CLIL Students
N=73

Questions regarding unfamiliar words	<i>Never</i>		<i>Sometimes</i>		<i>Often</i>		<i>Almost Always</i>		<i>Always</i>		<i>Missing</i>	
	EFL	CLIL	EFL	CLIL	EFL	CLIL	EFL	CLIL	EFL	CLIL	EFL	CLIL
V24 How often do you encounter new and unfamiliar words	3%	3%	58%	47%	15%	34%	15%	5%	3%	8%	8%	3%
V25 How often do you stop reading in English because of words you cannot understand?	23%	18%	53%	55%	15%	11%	3%	13%	0%	0%	8%	3%
V26 How often do you look up words in a dictionary?	23%	34%	53%	42%	10%	16%	5%	5%	3%	0%	8%	3%
V27 How often do you guess the meaning of a word from context?	3%	8%	10%	29%	35%	40%	33%	18%	13%	3%	8%	3%
V28 How often do you ask the teacher?	13%	3%	60%	55%	15%	24%	3%	16%	3%	0%	8%	3%
V29 How often do you ask your parents or others at home?	38%	34%	43%	34%	8%	21%	5%	8%	0%	0%	8%	3%
V30 How often do you ask friends or fellow students?	13%	5%	38%	45%	30%	34%	13%	11%	0%	3%	8%	3%
V31 How often do you just ignore the word and continue reading?	18%	21%	48%	45%	13%	16%	15%	11%	0%	5%	8%	3%
V32 How often do you give up reading completely?	60%	55%	23%	32%	8%	3%	3%	3%	0%	5%	8%	3%

6.1.1 CLIL Students

The most important statistic in this table is the answers to the how many new words the CLIL students encounter. Studying CLIL one would expect that the number of unfamiliar words would be rather large. A total of 47% of the students claim to encounter words often or more, with most of these, 34% saying that they encounter them only often. We would expect that those students studying CLIL should encounter new words more often than those in EFL. We will now take a look at some of the strategies that these students most often employ, and which are rarely used.

Looking at the strategies for dealing with, and learning, vocabulary that the students have claimed to use often or more we can get a better picture of what tools the students have available to them when encountering new words. The most common strategy that was used often or more was guessing the meaning from context, with a total of 61% of the students saying that they used this often or more. Following this strategy were asking fellow students or friends with 47% of students utilizing this often or more, and asking the teacher with 40% of students saying they would do this often or more. 32% of students would ignore the words they could not understand often or more, while 29% of students would ask their parents or others at home often or more. The least used strategy amongst the CLIL students was looking up words in a dictionary, with only 21% of students saying that they would do this often or more. Using a dictionary was also the strategy that most students claimed to never use, tied with asking parents or others at home at 34%. The strategies that the fewest would never use were asking the teacher at only 3% and asking other students or friends at 5%. In addition to these results a total of 55% of students said they would never give up reading an English text because of words they did not understand. 11% of students would give up often or more because of words they did not understand.

6.1.2 EFL Students

The EFL students said that they encountered new words often or more 33% of the time, with only 15% of these stating that they encountered new words only often. The most commonly applied strategy for the EFL students is by a clear margin, understanding words from context.

With a total of 80% of the students asked saying that they would use this strategy often or more. The second most common strategy coming in some way behind contextual understanding was asking friends or fellow students for help, with a total of 43% of the students using this strategy often or more. We then see another drop before we reach the next strategy, which was asking the teacher for help, being utilized often or more by only 20% of the students. The least used strategy for the EFL students was asking their parents or others at home, with only 13% of students saying they used this often or more. Using a dictionary was also only used, often or more by 18% of the students. As with the CLIL students the two strategies that were most often never applied were asking their parents or others at home with 38% of students saying they never did this and using a dictionary to look up words being never used by 23% of the EFL students. Interestingly 28% of students said they would often or more ignore unfamiliar words and continue reading, while there might be some that consider this the same as understanding contextually it might imply that some of those saying they are reading contextually are actually only ignoring the unfamiliar words and pushing on without actively thinking about their meaning.

6.1.3 Comparison of Answers to Unknown Words

In this section the answers of the CLIL and EFL students will be compared and discussed in greater detail. In table 6.7, that was shown above, the CLIL and EFL students are both compiled into a single table. This can show if the CLIL students have different strategies from the EFL students when they encounter new words and what strategies are more common among the CLIL students compared to the EFL students.

For the most part the two groups are fairly homogenous, there are however a few things that separate the two groups. Firstly we can see that 47% of the CLIL students have crossed off saying they encounter new words often, almost always, or always, while for the EFL students there are only 33% of the students that have crossed off that they encounter new words often or more. This fits in nicely with the hypothesis that CLIL instruction should expose students to more new vocabulary than EFL instruction. This is at least a positive sign that the CLIL lessons are introducing the students to more new vocabulary than they would encounter

otherwise. Whether this difference of 14% is enough of a difference will not be discussed here, it is enough to note that the CLIL students feel that they encounter more new vocabulary than the EFL students.

When it comes to the strategies that are employed by the two groups when encountering new or unfamiliar words the answers are fairly similar, there are, however, some differences between the groups. The most common strategy for both groups was understanding from context, in the EFL class however 80% of students would claim to use this often or more, while 61% of the CLIL students said the same. This difference between the groups is rather large, and it seems like the EFL students rely more heavily on this strategy when encountering new words than the CLIL students do. There could of course be explanations for this, such as the difficulty of words encountered in CLIL being greater than in EFL, and that they are more difficult to understand contextually from the text. This might explain why the CLIL students rely more heavily on asking the teacher or their parents at home about new words than do the EFL students. Since 40% use the teacher often or more and 29% use their parents often or more in the CLIL class compared to only 20% and 13% of the EFL class. In addition to this the EFL students will more often wholly reject strategies other than contextual understanding, choosing not to use them at all. The one strategy that is rejected more often by the CLIL students is, other than contextual understanding, is looking up words in the dictionary, 23% of the EFL students never use this, while 34% of the CLIL students elect to never use this. The CLIL students seem more reliant on using others to help them understand while the EFL students use strategies they can employ more independently, although both groups will ask their fellow students and classmates the same amount when dealing with new words. This could be attributed to the nature of CLIL instruction, where the words are of a more unfamiliar nature to the students, and so the CLIL students might need more aid from the teacher than those in an EFL class. If we check to see the correlation between CLIL instruction and asking the teacher for help we find that there is a correlation of $r=.233$, $p<05$ $N=73$. This indicates that there is a weak correlation between students who have CLIL instruction and asking the teacher for help when they encounter an unfamiliar word. Meanwhile there is somewhat higher correlation of $r=.35$, $p<05$ when it comes to the EFL class and the use of contextual understanding, which confirms our suspicions.

6.1.4 Summary of Survey Questions about Unknown Words

There are a few differences between the CLIL and EFL students, most importantly, the CLIL students encounter more new words than the EFL students. They also use their teacher more than the EFL students. This greater emphasis on using the teacher is interesting, this might suggest that the teacher's vocabulary knowledge will be even more valuable in a CLIL class than in an EFL class. In section 2.4.1 I mentioned how difficult it was to find teachers with the proper qualifications in both English and the subject being taught for CLIL instruction. This result shows the value of having such skilled individuals teaching CLIL. Other than these two differences there are few important distinctions between the groups. The EFL students understand more words contextually than the CLIL students, but this could be attributed to the fact that they also encounter fewer unknown words than the CLIL students. In the next section we will be looking at some of the significance of these results and answers, analyzing them further.

Analysis of Quantitative Data

In this section I will be investigating the significance of these results and look into the effects of CLIL and other factors on the students vocabulary. As seen in table 6.1 the EFL students scored higher on the X-Lex test than the CLIL students. This result shows that the CLIL students do not have a better vocabulary than do the students in the EFL group. However, it is highly doubtful that a student's English would become worse from CLIL instruction. Therefore there must be other factors that can explain why the CLIL students have not managed to do better than the EFL students, and to some degree have done even worse. If we look at the grades of the students, we can see that the grades are somewhat higher in the EFL classes than in the CLIL classes, especially when we remove those students who have claimed to know too many false words. Indeed, when looking at the correlation between X-Lex scores and grades for all of the students, we find a strong correlation of .652 with a significance of .000 indicating that there is a strong relationship between grades and X-Lex scores. This is a positive sign that the X-Lex test is able to assess learners, in a valid manner, correctly, and also is an indication of the importance of vocabulary in the English subject. In that a good

vocabulary should correlate with a higher grade. This can also explain the reason why the EFL students scored higher on the X-Lex test, as their grade averages are higher than the CLIL students. Unfortunately there is no way to determine if the difference in vocabulary scores between the EFL and CLIL students can be explained solely by the grade average of the students. The grades make it hard to separate what could be affected by the CLIL instruction and what can be explained by the better grades of the EFL students in this study. I will explore the impact of the grades further a little later in this section.

Continuing this testing we can look at the relationship between the X-Lex test and the written texts that have been analyzed. There is little purpose to separate the groups when testing for this as the sample will become far too small. Testing for the correlation between the X-Lex calculated scores and the type/token ratio we find a weak and non-significant correlation of .151. This is hardly surprising considering the small sample size that we have to work with, and it might be that with a larger sample that small correlation could prove to be greater. It would be interesting to see the effects of a study looking at this relationship between receptive and productive vocabulary in a larger sample students.

6.1.5 Analysis of Strategies to Deal with Unknown Words

While grades are a good explanation as to why the EFL students scored higher than the CLIL students in the tests, I also need to look at why they have been so successful compared to the CLIL students. Why do they have a better vocabulary and why do they have better grades as a result. Earlier in this chapter I looked at the correlation between grades and the amount of false words that were selected, discovering that the students with higher grades would select fewer false words than students with lower grades. This could imply that these students have developed better strategies for dealing with unknown words. Looking at the correlation between grades and the how often a student would meet unknown words reveal an interesting result that might shed some light on this. Those students who said they met unknown words often are also students who generally got poorer grades, as we can see from the negative correlation of $r = -.259$, $p < .05$, $N = 74$. This negative correlation between grades and how often they meet unknown words is perhaps unsurprising, but it is also a clear indication that the

stronger students could probably be challenged more in their English. In table 6.10 below I correlate the question “How often do you meet unknown words” with the other questions regarding unknown words in the survey.

Table 6.8 Correlation of How often do you meet unknown words with the other questions in the survey for the students in the sample who completed both the X-lex vocabulary test and Survey, N=71

Correlation of How often do you meet unknown words?	Pearson Correlation	Sig. (2-tailed)
How often do you stop when reading English because of words you do not understand?	.459	.000
How often do you use an English Dictionary?	.533	.000
How often do you guess the meaning of a word from context?	.052	.658
How often do you ask the teacher?	.204	.081
How often do you ask your parents or others at home?	.321	.005
How often do you ask your friends or classmates?	.344	.003
How often do you ignore the word and continue reading?	.298	.010
How often do you give up reading completely?	.580	.000

There is a significant correlation between how often students meet unknown words and how often they will use strategies such as looking up words in a dictionary, asking friends, and asking parents or others at home. Interestingly there is no correlation between students meeting many unknown words and those asking the teacher for help. So while other strategies for dealing with unknown words such as looking up in a dictionary and asking friends strongly correlates with students who meet many unknown words, asking for help from the teacher is something that both strong and weaker students will do.

Another interesting point about students who meet many unknown words is how strongly this correlates with students who will often give up reading a text completely. With a correlation of .580, $p < .001$ $N = 74$ this is very clear. Although this seems like common sense it goes some way to confirming a longstanding belief in regards to vocabulary which is that if 95% of a text is known a student should be able to understand the text completely. Although it is certainly possible to read a text with less known words this could make it easier for the student to give up completely as the meaning of the text becomes less clear. How many unknown words can be in a text before it becomes so difficult as to make the student give up is not clear of course, and this might be more a test of will than anything else. If we look at the strategies we can see that none of these correlate strongly at a significant level with giving up on reading a text completely. Although it is not significant it can almost look like these students employ fewer strategies than others. In fact, contextual understanding has a very weak insignificant correlation with students who give up on a text completely, while students who ignore words they do not understand and keep on reading correlate at $r = .321$, $p < .005$ $N = 74$. These students do not seem to make an attempt at understanding the words, choosing not to deal with them at all. Also, these students have a significant strong negative correlation with grades, at $r = -.522$, $p < .01$ $N = 74$. Table 6.11 below looks further at the correlation of grades with the questions regarding unknown words that were asked in the survey.

Table 6.11 Correlation of Grades with the answers to questions regarding unknown words for students in the sample who completed both the X-lex vocabulary test and survey N=71

Correlation of Grade in Written English with questions about unknown words	Pearson Correlation	Sig. (2-tailed)
How often do you meet unknown words?	-.259	.027
How often do you stop when reading English because of words you do not understand?	-.051	.670
How often do you use an English Dictionary?	-.015	.899
How often do you guess the meaning of a word from context?	.052	.662
How often do you ask the teacher?	-.072	.545
How often do you ask your parents or others at home?	.047	.694
How often do you ask your friends or classmates?	-.150	.205
How often do you ignore the word and continue reading?	-.332	.004
How often do you give up reading completely?	-.522	.000

As we can see from table 6.11 above, there are some more interesting correlations. Strong students, or at least those with good grades, do not ignore words that they do not understand with a negative correlation of $r = -.332$, $p < .005$. There is no significant correlation between good grades and specific strategies however, indicating that none of the strategies that the students were asked about are exclusive to students with good grades. One can choose to interpret these results as a sign that these students are not being challenged enough considering the weak negative correlation of $r = -.259$, $p < .05$ between grades and how often the students meet unknown words. The results could simply imply that these students do not need to employ these strategies because they are not often in a situation where they encounter unknown words.

These two groups, the students with high grades and the students with lower grades are seemingly at odds with one another. The stronger students are seemingly not being challenged enough and the weaker students seem to be challenged too much, and are giving up without trying. On a more positive note, the scores of the stronger students in the X-Lex test are quite high, in fact the mean score of 3607,97 is fairly high for this age group, which is a sign that Norwegian students might have a fairly strong general vocabulary.

Summary of Data Analysis

I have now presented and analyzed the most important and significant data that was found in this study. The results of the vocabulary tests showed that the EFL students did better than the CLIL students. When correlating vocabulary scores with the grades of the students I found that there is a strong correlation, helping to explain why the EFL students scored better than the CLIL students. Using the rest of the data collected in the survey I looked at the correlation between how often a student met unknown words and what strategies they would use to solve this with. Discovering that the students who met unknown words often would often use a dictionary, and would also often give up on reading completely. This confirms what Hellekjær (1996) mentioned in chapter 2 that some students lack the reading strategies needed to read a text effectively. However, continuing the analysis, I looked at the correlation between the grades of the students and the questions regarding unknown words. There was no significant correlation between how often a student would look something up in a dictionary and their grade. There was a strong correlation between the grades and giving up on a text completely. In the next section I will discuss the interviews held with the teachers and students in the CLIL class to attempt to shed some light into the process of teaching CLIL, and further explain why we might be seeing the results of the vocabulary tests.

Introduction to Interview Section

In the following sections I will be discussing and analyzing the interviews conducted with the CLIL teachers and students during the study. These interviews will be used to help explain why the CLIL students did not score as highly as I hoped on the vocabulary tests. The most important function of these interviews is to give an insight into how the teachers thought about CLIL, and to what extent they valued vocabulary and language in their teaching. The interviews with the students will also help by asking questions regarding their use of English outside of school, and their view on English. The interviews touch on several interesting topics, and further study could certainly be useful, I will, however, mostly focus on the topics that help paint a better picture of why the EFL students scored better on the vocabulary tests than the CLIL students.

Interview with CLIL Teacher A, General Section

The teacher at School A teaching CLIL (Referred to as Teacher A from now on) taught her CLIL students in Social Studies. This was her fifth year as a teacher having completed a teaching degree four years previously. The subjects that she taught in included Norwegian, English and Social Studies. She explained that she had no prior experience with CLIL, having never even heard of it before she was asked to teach it. It was the principal at the school that had originally asked her to teach a CLIL class and she saw it as an interesting opportunity to improve on the students' English. She went on to explain that the students did not really have a choice on whether or not they wanted to have CLIL. It was, however, optional to use English in class. "...it's optional if they want to write in English, speak English, or if they want English or Norwegian texts but it is not optional that I speak English...". The class had what she explained as period teaching, where they would have an increased amount of teaching hours in the subject during a three-week period. "We have eight hours a week, so it has been pretty intense. So when we've had about the Cold War in English, we had eight hours a week for three weeks." They would then continue with another three-week course of English that

was in progress as this interview was being held. This would total approximately one third of the total hours of the subject for that year, which is the necessary amount of hours needed for it to qualify as a CLIL class. She continued to explain that this was really all the time they had available to teach the CLIL method. “We had to do it this way since this was the tenth grade... ..After this they have to work towards their exams, since they are graduating this year.”.

When asked about their use of English in class Teacher A explained:

They answer quite a few assignments related to the texts, and those assignments are in Norwegian... ..They get written tasks pretty much every class, and they can choose if they want to write in English or Norwegian. More and more have started writing in English regardless of if they have read the text in Norwegian or English.

When asked if the students had attempted to do tests in English as well, she mentioned that they had at first seemed eager, asking if they could answer their tests in English, but no one had ultimately done so. When asked about how much reading they did in the class Teacher A explained that in the last period (The Cold War) they had some English texts for the students, but most of the texts that they would read were mainly in Norwegian.

They way we have understood it we are only supposed to present the material, then they can choose for themselves (Norwegian or English). And then there have been few who would choose English texts, though they have been available to them all the time, both online and printed out.

She had not considered the English material as something that should be forced onto the students, but instead something the students should be allowed to decide for themselves if they wanted to use. When asked about what textbook they had been using in the CLIL class she said that they had used the ordinary Norwegian textbook. This was due to the difficulty in finding a suitable replacement book that she felt was at the level of the students. They had considered switching to another topic because of this, one that would be easier to find texts for, but she felt that since this was their tenth and final grade that they needed to get through

the curriculum together with the other classes. She was also asked if the CLIL and English class would cooperate in any way, which she explained had proven difficult to accomplish. Social Studies and English did not have many comparable topics, making such a cooperation difficult. She again pointed out that the students were concerned about their grades, and the teachers did not want to damage their chances, so they were unwilling to change either subject to coordinate with the other.

6.1.6 Interview with CLIL Teacher A, Vocabulary Section

When asked to explain her thoughts on the issue of vocabulary when selecting materials for the class Teacher A explained:

All the normal vocabulary, the day to day stuff, you get that in the regular English lessons... ..And now, I think, when they are in the tenth grade they have to be able to discuss all kinds of things. Warfare, politics and such. So that you learn words related to those things.

Teacher A continued to explain that she felt that the students needed to learn words related to subjects such as communism, the arms race, and other words that are important to the Cold War. Subject words were the words she felt were most important to learn for the students. She did not focus on more ordinary words, leaving these to the English lessons. She felt that they needed these subject-specific words to be able to understand what they were reading and writing about, because without them they would not be able to understand the subject at all. When asked about what strategies they had attempted to implement to learn new words Teacher A responded that they did not really have any complicated strategies. They would often consult a dictionary or ask their teacher or others in the class. She also explained that she would encourage them to try and understand words from their context in a text. The only direct vocabulary teaching Teacher A would employ was related to important subject words. She would attempt to have these words repeated throughout a class, introducing them with a definition at the beginning of the class and then using it consistently in the class itself.

6.1.7 Interview with CLIL Teacher A, Difficulties and Thoughts

As noted above, Teacher A stressed the difficulty of finding suitable materials for the class to use. Not only was this a problem for the students, she had also felt that it impacted her ability to teach effectively because of subject words that she struggled to find a good English equivalent to the Norwegian word.

You feel kind of stupid too, but at the same time I could not find the word. I am sure there is a word for it, but I had no idea. So there are always some words that crop up. But if I had an English textbook then I feel that the problem would vanish.

She also said that it was time-consuming to find good, suitable texts for the students. When teaching about the cold war this was not all too problematic, but in relation to their current topic, which she said that had been very difficult to find anything that was usable for. “That really is the problem, and I feel a lot of issues would be solved if we had good materials.” At the end of the interview Teacher A added that “It might be that we have done everything wrong, but we have not got any kind of instruction or anything. So we just assume some things and then see what works.”

Interview With CLIL Teacher B, General Section

The CLIL teacher at School B (Referred to as Teacher B from now on) stated that she had a general teaching degree and had been working as a teacher for a total of fourteen years. She also said that she had no experience with CLIL before this year. She had however heard about it before and become quite interested in it, and she was eager to attempt it as having lived in Australia she felt her English was fairly good. However she expressed that she had no previous experience as an English teacher. “I feel fairly confident in English, at least speaking, but I have not got a degree in English.” She was currently teaching in Norwegian and Religion, and decided that Religion was a good subject to test CLIL in. While CLIL was

not strictly speaking optional for the students, they had the possibility of switching to her other Religion class that she was teaching at the same time if they wished. However, none of the students had expressed any desire to do so. In total the class was supposed to have half their classes in English since it was only a two-hour per week course. However due to an assortment of other events and arrangements this number was somewhat lower, Teacher B believed it to be closer to 30% of the classes that had been had in English.

When talking about the use of spoken and written English Teacher B explained:

It is mostly oral, and then they read in Norwegian and I make the presentations in English... ..and then they have got a few texts in English, but not many. So I would say that approximately 80% of the written is in Norwegian and 20% in English. Bot orally it has been about 60-40 English.

She went on to explain that when they are writing they can choose for themselves if they want to write in English or in Norwegian. Discussing how often they would choose to write in English Teacher B said:

It varies from topic to topic, when we have had subjects that they have felt have had a lot of difficult words then they have chosen to write in Norwegian. It has a lot to do with vocabulary, having enough vocabulary... ..If they think it takes too long to find the right words then they choose to write in Norwegian. But when it comes to speaking they usually use English, and then they will just use Norwegian words on the words they do not know and keep going.

She felt that they were much more confident speaking than they were writing English as well, explaining that it was not as embarrassing to say something wrong, although it had been like that before. When asked about what materials they used she, as Teacher A had, explained that they used the Norwegian textbook for the most part. While she would find some English texts online, but this had not been something she had focused to much on. With regard to the level

of cooperation between Teacher B's CLIL class and the English class she said that they had succeeded in working together earlier in the year on the topic of ethics. The students were tasked with presenting a presentation in English about an ethical dilemma. This was worked on in both the English and Religion classes, and the students got a grade in both subjects which she felt worked quite well. This was because the English teacher would mostly focus on correcting their English while Teacher B could focus on the content.

6.1.8 Interview with CLIL Teacher B, Vocabulary Section

When asked about how Teacher B taught vocabulary in her CLIL class she stated: “I do it as I go, we start with a word-wall with a few words on it, and then we collect words as we go, and keep building on it throughout the topic.” The word-wall was a simple vocabulary list that they built as they went, adding more and more vocabulary to it. The word-wall would in part be made by the students and in part by her. She also focused on the content words, important subject words that the students would need to understand the subject itself. She felt that “If there are recurring mistakes I will try to correct them, but for me the content is what is important. If I am too picky it might be disheartening for the students.” She also pointed out that her subject was such a small one that time constraints were a real issue. Therefore she could not spend too much time on pure English mistakes, or teaching basic vocabulary. When the students were given English texts to read they would often use markers to outline words that they were unsure of or did not know, and she also encouraged them to attempt to understand words from the context surrounding them. “I have experienced that they are incredibly different, some have a very strong vocabulary, while others have a much more primitive one in English.” While she had no specific ways of measuring their vocabulary she explained that since it was an oral subject, and focused on their oral output and their ability to speak and be comfortable in doing so.

6.1.9 Interview with CLIL Teacher B, Difficulties and Thoughts

As with Teacher A, Teacher B felt that the greatest difficulty in her CLIL teaching was that the texts she had found online were often too challenging with regard to their vocabulary.

“Those texts that are online are often quite difficult with an advanced language. Not very accessible for the students, so that is a challenge. Not having teaching materials that are well suited.” Teacher B felt that those texts were too challenging for the students, but reiterated that she had not focused too much on texts as it was an oral subject and she was focused on that. Her goal with teaching CLIL was to make the students a little bit more confident about their English, so they would use it more later. She also stressed at the end of the interview, just as Teacher A, that teaching CLIL was a new experience and that they had only been working with it for a few months. She was eager to continue, but with younger students, and to have CLIL over several years.

Comparison of the Teacher Interviews

When comparing what the two teachers said we can see that they both encountered many of the same problems and challenges when teaching CLIL. Chief among them was the subject of materials. Both Teacher A and Teacher B found it difficult to find good, suitable English language materials for their students in English. This was mostly due to the difficult vocabulary and wording that these texts would have compared to the students' normal textbooks. In addition to this, and somewhat in relation to their search for materials, were the time constraints they were under. Teacher A repeatedly explained how she felt that they needed to get through the necessary material for the exams, and that therefore English was often forced to take a back seat. Teacher B did not state this explicitly but pointed to the fact that she only had two hours of Religion with the class a week, and many of these classes would not occur due to other activities the school had. Both of them were also focused on teaching subject words, the normal day to day English vocabulary was of a secondary concern. They stated that the subject words were necessary to understand the subject itself. If you wanted to talk about the cold war for example you would need a vocabulary that contained certain topical words. The teachers were not occupied by correcting more normal English mistakes, as this was seen as something that might demoralize the students should it be focused on too much. The importance was placed on the content, not the language.

To sum up, neither class can be said to have had a large amount of CLIL lessons, with both teachers estimating the number of CLIL lessons to be about 30% of the total lessons in the subject. While teacher A had taught her CLIL lessons in a shorter time period, because of the period plan her class was on, Teacher B had spaced the CLIL lessons out, having one CLIL lesson for every normal lesson. Another common theme was concerned the amount of reading, writing, and speaking. Both CLIL classes had mostly focused on speaking English, reading was more problematic in part due to the difficulty in obtaining suitable texts in English. Teacher B seemed more focused on the oral part of the subject than Teacher A, saying that very little of what the students read was in English compared to their use of English when speaking. When looking at their backgrounds neither teacher had any prior experience with CLIL before the current project began. Teacher A had taught English before however, while Teacher B had never taught English before. One interesting difference between the two is that Teacher B was actively planning to teach CLIL before the project they participated in began, while Teacher A was approached by her principal and asked to do it. Teacher B had already heard about CLIL, and was planning on doing it on her own, which might have given her some more time to prepare.

Interviews with CLIL Students

In this section we will be looking at the interviews with some of the students from the two CLIL classes. There are three students in each interview, which will be referred to as 1, 2, and 3 A or B depending on their school. The students were interviewed together and were asked questions regarding their use of English outside of School, their thoughts on English, and what vocabulary strategies they employed.

6.1.10 Interview with CLIL Students from School A

The students from school A consisted of two boys, A1 and A3, and one girl A2. The students were first asked about how often they used English outside of a school setting. All of them

explained that they would use English quite often outside of school, two of the students, students 1A and 3A used it frequently when playing computer games, while student A2 used it more on YouTube and other internet sites. When playing computer games they would write and speak English, and all three stated that they spoke a fair bit of English outside of school. None of the three students had read any English books or anything similar. The students were also asked about what they would do when encountering an unknown word in English. Student A3 immediately responded saying “Either ask the teacher or use a dictionary.” When asked if he did this often he admitted “No, not really.” A2 then said that she would often try to guess the meaning of a word from context, also confirming that she did this outside of school as well. A3 also mentioned briefly that he would use Google to get the meaning of a word. They did not have any other methods of dealing with unknown words.

All three of them seemed somewhat unsure if having a strong English vocabulary would be important to their future education, however, both A2 and A3 believed that being able to use English would be useful for them later in life. While somewhat contradictory this might imply that they were considering other things than their education. When they were asked if they thought their English vocabulary had improved from having CLIL, they were not very confident that it had. A2 and A1 both felt that they might have improved a little bit. A3 went on saying “No, it is mostly just the teacher speaking English.” When he was then asked if he did not answer in English he said “Some do, but I do not.” Continuing the questioning they were asked if they spoke in Norwegian because they lacked the necessary vocabulary, or if it felt strange talking in English. To this they replied that they felt it was both, it was just easier speaking in Norwegian. All three of them were confident in saying that speaking English was easier than writing, mostly because they spoke more in class than they would write.

6.1.11 Interview with CLIL Students from School B

The students from school B consisted of one boy B1, and two girls B2 and B3. They were also asked about their use of English outside of school, to which B1 responded that he used it when he played computer games. B2 said that she used it when she travelled each year, as well as hearing it through television shows and movies. B3 said she used it some when on the

computer but not very often. When asked if any of them read any English books only B2 said that she did this. None of them spoke much English outside of school, B1 saying that he did not speak any English at all. They were then asked about what they did when they encountered unfamiliar words, to which B2 responded that she would want to find the meaning, and would often ask other students for it. In contrast to this, B3 said that she would usually just ignore the word and keep reading, and B1 said he would attempt to understand the word through contextual clues.

Some interesting answers came up when the students were asked about their opinion on how important English was for their education. Student B3 said “I have really given up on English. I think it's rather hard, so I'm hoping not to have it for very much longer.” However when asked how important she thought English was for her education she reasoned that “Yes, I think it actually is to be honest.” It is interesting that she identifies the need for English and yet she does not desire to continue with it. B2 Had a somewhat different response to the question, saying that she loves English and that she had always been good at it. Finally B1 had a more practical view on the use of English “It might be important if you go out to sea or something, but other than that I can't really see the importance of it.” It should be noted that the most likely meaning of going out to sea that B1 meant was to work on an oil platform, which many of the boys at the school claimed to be planning on doing. They were also asked if they felt that their language, and especially their vocabulary had improved through their CLIL class. B2 explained that she had improved on her grade but was unsure if this was because of an improved English or something else. B2 continued saying “We've learned a lot of religious words, but there wasn't anything more,” showing her doubt that the CLIL instruction had any effect. B3 pointed out that “ We did get to speak English and got more practice you know.” They both continued explaining that they felt equally good at speaking and writing English. Still, when asked whether they felt most comfortable when they were writing or speaking English they responded differently. B2 saying she felt more comfortable when she spoke English and B3 thought she found it most comfortable to write in English, to which B1 chimed in in agreement.

Summary of Interviews

The interviews revealed several interesting points that will be useful to help explain the results of the vocabulary tests. The most important thing that came to light in the interviews was the teachers explaining the amount of CLIL they had taught and their problems with finding English texts for the students to read. They also talked at length about how worried they were that the language would hurt the content in the subject. They were focused on teaching the subject, not the language.

The student interviews also helped reveal some interesting topics that are important for the subject. They mentioned that they did little reading outside of school, with most of the English was done in other forms than reading books. They also had little knowledge about their improvements in English, seeming to not be very aware of their own progress in the language.

Overall Summary of Results

In this chapter I have given an overview of the results of this study. First presenting the results from the most important part of the study, the X-lex test. These results show that the CLIL students did not score better on the vocabulary test than the EFL students. The text analysis of their written material resulted in the same, although there seemed to be less difference between the two groups here. In an attempt to explain and explore why the scores on the vocabulary tests are the way they are I analyzed the survey that was conducted. The grades of the students were an especially telling point in explaining the vocabulary scores here. Looking at the correlation of grades with the questions about unknown words I found that this correlated with strategies dealing with new vocabulary. This goes some way to confirm that the grades are a good indicator of a strong vocabulary.

The grades were an answer to why the vocabulary scores were the way they were, but this did not explain why it seemed like the CLIL instruction had not impacted the vocabulary of the students very much, if at all. Using the interviews I have explored why the CLIL instruction was not as successful as one could hope it would be. The main explanations for this were found in the teacher interviews. They explained that CLIL teaching was limited to 30% of the total subject hours, and that they used mostly Norwegian texts. The two teachers had also been occupied with teaching content, not worrying too much about the language. These points help explain why the CLIL instruction seems to have had little effect on the vocabulary scores of the students. The student interviews did not reveal much that could explain the test scores, they did however reveal a few things that are interesting. They reported that they did not read any English books, only one student saying that she read in English. They did use English outside of school however, mostly on the computer. Another interesting thing that came up in the student interviews was their awareness of their own vocabulary, English skill, and improvement in English. Their lack of awareness over their own English development was interesting, they struggled to say if they had really improved much through CLIL. In the next chapter I will be using these results to discuss the research question and also discuss them together with the theory presented in this study.

7) Discussion

Introduction

In this chapter I will first review and see whether, and to what extent, I have answered my research question. Then I will present the findings from my study, starting with the vocabulary tests, and supplementing these results with the findings from the survey and interviews. After this I will be discussing the findings in light of the theory presented in chapter 4 regarding incidental vocabulary learning. Then I will discuss the findings compared to other studies, and finally I will be giving some more thoughts on the validity of this thesis.

First of all in this chapter I would like to return to my research question, which is «Do classes with EFL & CLIL score better on vocabulary tests than classes with only EFL?» My findings, in the present study, show that they did not do this. In the following I will therefore be looking at the reasons why the CLIL students did not score better on the vocabulary tests. I will first give a brief summary of the most important findings from the current study. I will then discuss these findings in light of the relevant theory on incidental vocabulary acquisition. After this I will discuss my results in light of relevant results from a previous study on vocabulary acquisition in CLIL. Finally I will be giving some more thoughts on the validity of the study.

Relevant Findings from the Study

The aim of this thesis is as mentioned to see whether students who are receiving CLIL instruction in English, as well as ordinary EFL would score higher on vocabulary tests than those students who have received only EFL instruction. To find this answer I tested four

classes from the final grade of lower secondary school with two tests. Two of these classes received CLIL and EFL, and the other two received only EFL. This was because it was necessary to have a control group for the CLIL classes, as I needed to compare the results from the classes receiving CLIL to classes who received only EFL. In addition to this I wanted to collect additional information to help explain the results of the tests, and discover if there were other factors influencing the results of the vocabulary tests than the CLIL instruction. The results of the tests and the findings from the survey and interviews are presented in chapter 6 of this thesis. In the following sections I will give a brief summary of the findings, starting with the vocabulary tests, which I consider to be the most important findings of the study. After this I will then present the findings from the survey and the interviews with the CLIL teachers and students that help explain and expand on the findings of the vocabulary tests. As I am using an embedded design for my mixed methods approach I intend to use the other findings collected to help explain the results of the vocabulary test.

The Vocabulary Test Findings

In this section I will give an overview of the findings from the vocabulary tests, summarizing the most important and interesting results of the tests. Presented below are the conclusion of the two tests that were conducted, the X-lex test and the text analysis:

- The EFL students scored higher than the CLIL students on the X-lex test.
- The EFL students had a slightly higher type/token ratio than the CLIL students on the text analysis.
- The overall scores on the X-lex vocabulary test was fairly high.

These vocabulary tests show that the EFL students performed better, which was disappointing considering the hope that the CLIL instruction would have yielded better results on the

vocabulary tests. The mean scores of the X-lex vocabulary test that can be seen in table 6.3 show that the mean score for the CLIL students was 300 below the EFL scores. This means that, not only did the CLIL students not perform better, they actually performed worse than the EFL students. The text analysis gave a similar result, with the EFL students having a slightly higher type/token ratio than the CLIL students, which can be seen in table 6.5. Unfortunately the sample for this is very small, and so I will not be discussing this result any further. On a more positive note, the scores on the X-lex vocabulary test are fairly high, which suggest that many of these students have a very good grasp of some of the most common words in the English language. These results have thus answered my research question in a negative way, showing that the CLIL students do not score better than EFL students on vocabulary tests. The question still remains as to why the CLIL instruction does not give the CLIL students better scores on the vocabulary tests than the EFL students.

Findings from the Survey and Interviews

In this section I will be looking for why the CLIL instruction seemed to have little effect on the vocabulary of the students. I will be presenting the findings from the survey and interviews that I feel explain this in the following table, and then continue to explain why I believe these findings prove why the CLIL students did not score as well as the EFL students on the vocabulary tests.

Table 7.1 Important Findings from the Survey and Interviews

A) Survey Findings	B) Teacher Interview Findings	C) Student Interview Findings
The EFL classes had a higher grade average than the CLIL classes.	Lack of English texts, using mostly Norwegian texts in the CLIL classes.	Students would use Norwegian instead of English in the CLIL classes.
The CLIL students ask their teacher about unknown words more often than the EFL students.	Only 30% of the teaching hours of the subject were dedicated to CLIL.	Lack of awareness of their own progression in English.
The CLIL students encounter more unknown words than the EFL students.	Content was prioritized over the language, and English was voluntary.	Use of English outside of school was mostly restricted to use on the computer, music and TV. Very little reading of English books.

The findings in table 7.1 help explain the results of the vocabulary tests. Perhaps the most important finding to explain the vocabulary scores is the grade average of the EFL students being higher than the average grade of the CLIL students. In chapter 6 I found a significant correlation between grades and vocabulary scores, students with better grades seem to have a larger vocabulary as well. This helps explain why the results show that the EFL students scored better than the CLIL students on the tests. This would be an adequate explanation if I were comparing two EFL classes, but as one class has received CLIL instruction we should expect that their vocabulary would be at least equal to the EFL students. I believe that the other findings presented in table 7.1 indicate why this is not the case. I will now be looking at why the CLIL class did not work better than it has.

The most telling finding giving an indication as to why the CLIL instruction has not been as successful as had been hoped can be found in the teacher interviews. Both CLIL teachers explained that only 30% of the teaching hours in the subject were dedicated to instruction in English. In addition to this the lack of English texts, and the use of the Norwegian textbook meant that the amount of English that the students were being subjected to was not very large at all. This amount of English might be too little to have any effect on the English vocabulary of the CLIL students. Additionally, the CLIL teachers were both mostly concerned with

teaching the content of the subject rather than the language. They would not risk slowing down the progress in teaching content for the sake of language instruction.

The survey also had some more explanations as to why the CLIL instruction might not have been as successful as one could have hoped. The CLIL students diverged from the EFL students on two interesting questions regarding unknown words (see appendix 3 for complete survey). These questions were about how often they encountered unknown words, and how often they would ask the teacher when they encountered an unknown word. These two questions seem to indicate a demand for a strong teacher in CLIL, with a very good vocabulary. One of the teachers also mentioned how it was fairly difficult to find a good English equivalent for certain words. It is not ideal if the teacher is forced to spend time attempting to explain the meaning of English words to the students constantly.

The CLIL student interviews were less revealing, but there were some interesting points that might help explain the vocabulary results. They confirmed what their teachers had said, that students would sometimes choose to use Norwegian instead of English. One of the CLIL students saying that he preferred to answer in Norwegian rather than in English, as this was easier. This helps decrease even further the amount of English that the CLIL students are subjected to in CLIL. They were also asked about their use of English outside of school, to which most replied that they used it a little, though this was mostly restricted to games, chatting, listening to music, and watching TV. Only one of the students said that she read English books. These things give the impression that the amount of outside English influence on the CLIL students was not all that large. Finally I noticed that the students were not aware of their own progression in English, they could not say for certain if the CLIL instruction had improved on their language. While this might not be very surprising I found it interesting, as I believe that having an awareness regarding your vocabulary will aid vocabulary development.

In this section I have presented the most important findings that I believe can explain why the CLIL instruction in these classes have not been as successful as I had hoped. In the following

section I will be discussing these findings in light of the theory that has been presented in this thesis.

Theoretical Discussion of Findings

I have already established that the CLIL students did not score higher on the vocabulary tests than the EFL students, and now we will be looking more closely at why they did not do this. I will start by giving a brief overview of the two most important theoretical aspects that I will use to explain the results from the vocabulary tests.

7.1.1 Overview of Theory

In chapter 4 I presented a few theories on how vocabulary is acquired and retained through incidental contact with the language. Krashen's (1989) input hypothesis suggests that through input alone new vocabulary can be acquired if the vocabulary is not too demanding. I have challenged this idea in chapter three presenting several alternative views on what is required for successful incidental vocabulary acquisition. Most importantly I have looked at the Involvement Load Hypothesis by Laufer and Hulstijn (2001), and the effects of repetition on incidental vocabulary acquisition as tested by Folse (as cited in Lee and Hirsh 2012).

The Involvement Load Hypothesis devised by Laufer and Hulstijn (2001) claims that the learning of a word does not depend on simply exposure to the word, it instead relies on the need and involvement on the part of the learner. Simply telling the learner what a word means would involve little to no involvement on the part of the learner which would mean it, according to the Involvement Load hypothesis, would be less likely to be retained by the learner. In simpler terms this means that a learner needs to be challenged in their efforts when understanding the meaning of new words.

Repetition is claimed by Folse (as cited in Lee and Hirsh 2012) to be more important than the involvement of the task when it comes to vocabulary acquisition. Quantity should therefore be focused on instead of creating one demanding task it would be more beneficial with several less involving tasks repeating the same word in all of the tasks. This is an unsurprising discovery perhaps, as repetition has been used in many fields of teaching with success. Nation (1990) has also pointed out that using texts to help in the repetition of vocabulary can be very effective, allowing for a large amount of vocabulary to be repeated.

To summarize, there are two important aspects to the retention of vocabulary, quality and quantity. While Folse (as cited in Lee and Hirsh 2012) recommends that it is better to focus on quantity, I would argue that using both methods could yield far greater results than using only one method exclusively. The involvement Load Hypothesis seems to suit the more important words in a text, while repetition is more suitable to the high-frequency words in a text. Low frequency words will be difficult to repeat in texts, and they could benefit from a more focused quality approach to being acquired by the learner. Following are two sections looking at what I have dubbed the quantity and quality aspects. These look at the amount of vocabulary and the quality of the vocabulary tasks that are evident from the findings in the study.

7.1.2 The Quantity Aspect

Looking at the amount of hours the two CLIL classes received we can see that they did not have much more than the required amount of English instruction that is needed for it to be called CLIL in Norway. Only approximately 30% of the instruction was in English, if we couple this with the fact that the textbooks used were the Norwegian ones we can assume that the amount of English input for the students was not very large at all. As was mentioned above, the usefulness of texts that repeat words as a form of vocabulary teaching can be very important. In the CLIL classes almost no reading in English has been done, which will have severely impacted the amount of incidental vocabulary learning in the classes. Using only

some English texts will not give the necessary amount of repetitions needed to obtain the vocabulary. Additionally it is not certain that the vocabulary is repeated, as the texts might not be about the same topic areas, therefore containing a different vocabulary in each text. As was explained above, the amount of repetitions can greatly impact the vocabulary learning of the students. In chapter 1 we presented a study by Hellekjær and Hopfenbeck (2012) where they suggested that for CLIL to be successful it was necessary with 50% of the instruction being in the foreign language. This seems to suggest that the amount of CLIL was simply too little to see much of an effect on the students' general vocabulary. The quantity of English, and the quantity of vocabulary in the CLIL classes seem to be far too little to have had the necessary impact on the vocabulary retention of the students.

7.1.3 The Quality Aspect

Looking at the quality of the vocabulary teaching and tasks involved in the CLIL classes is somewhat more complicated than the quantity. The quality does not imply a lack of skill or quality in the teachers, instead it suggests that there has not been enough of a focus on teaching vocabulary in the classes, and spending time ensuring that vocabulary is acquired. If we base the quality of the vocabulary teaching and tasks on the Laufer and Hulstijn's (2001) Involvement Load Hypothesis then we can gain a greater understanding of how effective the teaching of vocabulary has been in these CLIL classes. While the Involvement Load Hypothesis is designed for use with individual tasks we can use it to gain a picture of the role of vocabulary in the class as a whole.

The need dimension is related to how important the word is for the student to complete the task they are set to. If a learner is writing an essay on polar bears then the learner will need to know the word for polar bear if they are going to be able to write anything at all. Looking at the need aspect of the Involvement Load Hypothesis, being absent (0), moderate (1), or strong (2) we can see what need there existed in the CLIL classes to learn the necessary vocabulary. Intuitively I would suggest that the need aspect should be high in a CLIL class as the vocabulary is important to be able to express yourself on the subject matter. If we look at the practice in the two classes however we can see a slightly different picture. From the reports of

the teachers and the students they explain that the students are never forced into using English. Answering all their tests in Norwegian the need created from being necessary for a grade ceases to exist. In addition to this the students could also answer questions in Norwegian should they prefer to do so. One of the teachers explained that they would choose to write in Norwegian if they found that they lacked the vocabulary. Therefore it seems that the need to learn new vocabulary is almost non-existent in these classes. Scoring the general need for vocabulary in the classes, could suggest a score of absent (0) in the Involvement Load Hypothesis. The accepted use of Norwegian in the two CLIL classes makes it easier for the students to ignore the holes in their vocabulary knowledge rather than attempt to fix and learn the vocabulary needed.

Moving on to the second aspect of the Involvement Load Hypothesis is the Search aspect, either being absent (0) or present (1). The search component simply demands that the student must ascertain the meaning of the word in some way, how this is achieved is not very important, just that the learner is consciously attempting to learn the word. As was seen in the results section, the students answered questions regarding how often they would employ different strategies when encountering unknown words. The most significant difference between the EFL and CLIL groups were the amount of students who would ask their teacher often, and the amount of students who would ask their parents often. The CLIL students used these two strategies more than the EFL students who chose to attempt to understand the words contextually instead. This confirms that there seems to exist a search component in the CLIL classes, which means we can check this off as present (1). There is some reason to be cautious of this search however. In the interviews with the students they were questioned about what they would do when they encountered a new word. One student answered that he would ask the teacher or look up the word in a dictionary. When he was then asked if he did this often, he said that he did not. Although this is just one student it is suggestive that the actual amount of times that the students will search out the meaning of a word might be less than what might first be apparent from the survey results. As the survey results contain quantifiers that are vague by nature, such as sometimes and often, this means that they are always going to be subject to interpretation. We can however assume that most of the students who received CLIL instruction would ask their teacher often or more than often about unknown words. This means that the vocabulary knowledge of the teacher becomes quite important to the success of

the subject. While this thesis is not going to consider the English vocabulary skills of the teachers involved it can be important to note that one of the teachers had no formal English education, and the other teacher expressed how difficult she had found the subject-specific terminology. Finally it must be noted, again, that the students were never forced to use English and could freely switch to Norwegian, which was confirmed by the students in the interviews as something they would sometimes do. This will remove the search component a lot of the time, which means that the search component cannot always be claimed to be present in the tasks given to the students. Also, both teachers mentioned that they would supply some words to the students making the search component absent in these cases.

The final aspect of the Involvement Load Hypothesis is the evaluation aspect, this aspect can be either absent (0), moderate (1), or strong (2). The evaluation aspect measures the degree of assessment that the student will use when determining whether a word is right for the situation. It is the learner who makes a decision on when the word should be used. This is somewhat more difficult to address in the CLIL classes as it can be assumed that it will vary depending on the tasks involved in the class more greatly than the other aspects. However, both teachers said that they did not use any advanced vocabulary tasks in their teaching. They would correct smaller mistakes but they were not worried about language, focusing on content instead. The subject words were the only words that they considered to be important for the students to use correctly. Claiming that the classes received vocabulary instruction that were evaluation was absent (0) might be incorrect, but we can fairly safely state that the amount of evaluation on part of the students was not strong (2). There were few vocabulary specific tasks given to the students, nor where these tasks primarily focused on vocabulary learning. This is further compounded by the fact that English was considered optional in the classes.

To summarize the discussion on quality of vocabulary teaching and learning in the class we can say that it seems somewhat lacking according to the Involvement Load Hypothesis. This lack of focus on the quality of the vocabulary teaching means that the incidental vocabulary learning by the students was probably not as great as it could have been. Mostly this decrease in quality was done because of time-constraints. The teachers did not want the language to come at the expense of the content. Hill and Laufer (2003) tested the Involvement Load

Hypothesis and also looked at the time it took to complete the different vocabulary tasks. The hypothesis being that more involving tasks would take longer to complete for the students, therefore become less practically applicable. They found that this was not the case however. I would argue that in CLIL it might even save time, as better retention of vocabulary would remove vocabulary teaching down the line. The more vocabulary you know, the easier it is to obtain new vocabulary as discussed in chapter 2.

7.1.4 Conclusion to the Theoretical Discussion

In the theoretical discussion I have attempted to show that the CLIL instruction that the students received lacked the necessary focus on both quantity and quality. It is especially the quantity of the CLIL instruction that seems the most significant. With no textbooks, and only 30% of the hours being taught in English this makes for a very small amount of CLIL instruction. The effect of the CLIL treatment in the present study is therefore, not surprisingly, fairly small or even negligible. The lack of English textbooks for the students is especially detrimental to the vocabulary learning in the CLIL classes. This removes one of the simplest and most effective vocabulary learning techniques available, reading. The value of a textbook that contains a well-thought out vocabulary that is repeated could be of great importance to the vocabulary improvement in CLIL classes. The quality is debatable as there are many factors that are not presented in this study that could affect them, from the data available. However, it can be concluded that more could have been done to increase the quality of the teaching in the CLIL classes, especially with regard to the optional nature of English in these classes. Most of the written production of the students was in Norwegian, and a fair amount of the discussion in the class was also in Norwegian. Not only did this hurt the involvement load of the incidental vocabulary learning in the class, it would also decrease the quantity of incidental vocabulary learning. In the next section I will present the works from some other studies looking at incidental vocabulary learning, including one regarding the incidental vocabulary acquisition of CLIL students.

Studies on Incidental Vocabulary Learning

In Paribakht and Wesche's study (1997) that looks at the effects of reading on incidental vocabulary learning (see section 4.3) they found that the effects of incidental vocabulary learning was increased by additional vocabulary tasks. They discovered that the learners gained a greater understanding of the words when they received additional tasks. The learners who only read the texts they were given could recognize the words, but little else. The findings from their study showed that the effects of additional tasks would greatly increase the amount of learning a student would enjoy. Another study that tested the Involvement Load Hypothesis discussed above was Hill and Laufer's study (2003). They discovered much the same as indicated by the study by Paribakht and Wesche (1997). With the amount of tasks relating to the word in question affecting the knowledge of the word. The studies by Folse (as cited in Lee and Hirsh 2012) and Lee and Hirsch (2012) continued to test the Involvement Load Hypothesis and they found that repetition had a greater effect than the involvement. As has been discussed in section 7.4, I argue that both the number of repetitions and the involvement are important. These studies indicate that there is more to incidental vocabulary teaching than Krashen's (1989) input hypothesis.

7.1.5 Studies on CLIL

Liss Kerstin Sylven (2010) conducted a study looking at the implicit vocabulary development of Swedish learners in CLIL. The study was done over three test rounds spanning two years (Sylven 2010: 9). She found that the students who reported that they read English books outside of school would score as well as the CLIL students who did not read. This indicates that reading might be as effective as CLIL in some cases (Sylven 2010, p.109). While all of the CLIL classes scored better than the control groups in the study, she could only attribute the increase in vocabulary to CLIL in one of the classes. In the other classes there are other factors that influence the results more than the CLIL instruction. Her conclusion was that the most important reason for implicit vocabulary learning was the reading habits of the students. She pointed to the fact that many CLIL teachers in Sweden lack a sufficient English

proficiency and education as a problem with the current teaching of CLIL (Sylven 2010, p.226).

As has been mentioned earlier in the chapter Hellekjær & Hopfenbeck (2012) conducted a study on CLIL. While they did not look specifically at vocabulary learning in CLIL, the results are still interesting for the current study. Looking at the effects of reading in upper secondary school in Norway they discovered that there were certain demands for CLIL to be effective. Most importantly, as already mentioned, was that at least 50% of the hours in the CLIL class were in English. They also strongly recommended the use of English materials and texts being used instead of Norwegian (Hellekjær and Hopfenbeck 2012, p.119).

7.1.6 Comparison with the Present Study

The studies that I have detailed in this thesis looking at the vocabulary acquisition of learners in chapter 3 and summarized again here in section 7.5 seem to be in accordance with the results of this study. With insufficient amount of time and focus was spent on teaching vocabulary, and little time spent using the language at all, it is not surprising that the CLIL students did not score better than the students receiving only EFL on the vocabulary tests. The studies presented here have advocated either a focus on repetition, or on quality of involvement, neither of which has been present to a very large degree in the classes in this study.

Sylven's (2010) study on CLIL also shows how problematic it is to measure CLIL effectively. As she discovered that reading habits had a much greater effect than the CLIL instruction. This should not necessarily mean that CLIL is not useful, rather that CLIL instruction needs more focus on teaching language, and on using the language more within the class. In fact, Sylven's (2010) study supports the idea presented by Nuttall (as cited in Coady) in section 3.4 regarding the vicious circle of reading. The fact that reading habits outside of school have such a great effect on vocabulary should be seen as an encouragement to ensure that weaker students are able to obtain the necessary vocabulary to break this vicious circle of reading.

Sylven's (2010) study reaches much the same conclusions as can be drawn from this study. The conclusion being that the CLIL instruction has not had a very large impact on the students' vocabulary, instead it seems that the grades in English for the CLIL and EFL students have been the biggest factor in influencing the test scores. As with Sylven's (2010) study, this indicates that the CLIL instruction is in need of something more to prove effective, as right now, the CLIL instruction does not seem to have the desired effect.

Validity of the Study

In chapter 5 I discussed the external validity of this study, and potential threats to it. As was mentioned in section 5.9 the sample size present in this study is not large, nor is it representative. This is a significant factor when discussing the findings from this study. My conclusion from section 5.9 was that the study has clear limitations with regards to external validity. In this section I will therefore look a bit closer at the validity of the vocabulary testing in the current study and then discuss the validity of the study in light of the other studies presented in section 6.3.

The main focus of this thesis has been on the vocabulary scores of students on the X-lex vocabulary test, the validity of this test is the most important for the validity of the study as a whole. As was mentioned in section 4.5.1 variants of this test have been used successfully to test vocabulary. First we should look at the concurrent validity of the test. The concurrent validity of a test is measured in comparison to already established tests, checking to see if it can be compared favourably (Ary et al 2010, p.228). Checking the concurrent validity of this test we can see that results from the X-lex test have been compared favourably to CEFR levels (Milton 2009, p.191), indicating that the test can be used as a relevant measure for vocabulary knowledge. Another issue to look at when discussing the relevance of the X-lex test is construct underrepresentation. This refers to assessment that is too narrow, not taking into account other important dimensions (Ary et al 2010, p.225). As discussed in section 4.6, there are two forms of vocabulary tests that each measure different aspects of vocabulary

knowledge. Having settled for using a breadth test since this gave a better overview of the vocabulary of the students I could not measure the depth of knowledge that students had regarding the words they were questioned about. This might be considered construct underrepresentation, and therefore be an issue with the validity of the study. However, there is evidence that receptive vocabulary can give an estimate of the productive vocabulary of a learner, which means that this is less of a concern (Melka 1997, p.92).

The other studies presented in this chapter have shown results that concur with the results from this study. This strengthens the external validity of this study somewhat. The results from the studies on incidental vocabulary learning help explain why the CLIL classes did not see better results on the vocabulary tests that were conducted in this study. More importantly still, the two studies on CLIL, one by Sylven (2010) and the other by Hellekjær & Hopfenbeck (2012), both come to the same conclusions that I have come to in my study. This indicates that it is necessary with a certain amount of both quality and quantity in the CLIL classes to have an effect on the vocabulary of the students. The fact that the results from this study coincide with the results from the other studies strengthens the validity of the study somewhat.

To conclude the discussion on validity there is some evidence from the other studies presented that strengthen the validity of this study a little. I would also argue that the methods used in the study are reasonable valid, with the exception of the text analysis where too small a sample precludes it from being of much value. However, I would like to again stress that in particular the sample size makes it hard to generalize the results of the study to the whole of the student population.

Summary of the Chapter

In this chapter I have given a brief overview of the most important findings from the study, and discussed these in light of the main theories I chose to base my study on, namely the Involvement Load Hypothesis and the research on repetition of implicit vocabulary teaching. I have attempted to explain why the CLIL students did not score better on the vocabulary tests than the EFL students in this study. There are several studies looking at incidental vocabulary learning that come to many of the same conclusions as this study, showing that incidental vocabulary learning is more effective when given more involving tasks or more repetitions. The studies stress the importance of more than simply being subject to the language, encouraging the use of specific vocabulary exercises. Sylven's (2010) study on vocabulary in CLIL also agrees with the results from this study. As was discovered in this study, the average grades of the students was the most important factor when correlating the results of the vocabulary tests. This is much the same as what Sylven (2010) discovered, along with the students reading habits being more important for vocabulary development than CLIL. In the next chapter I will be making some concluding remarks, including some suggestions for further research.

In the next and final chapter I will discuss the implications of this study, and mention what further research could be conducted in this area.

8) Conclusion

Implications of the Study

In this study I have examined if CLIL students score better on vocabulary tests than students who are only receiving EFL instruction. The findings showed that the CLIL students did not score better on the vocabulary tests. Examining the possible reasons why the CLIL students did not do this I found that there was a lack of both quantity and quality of vocabulary teaching in the CLIL classes. In chapter 7 I argued that an incidental approach to vocabulary learning needs to be conducted with considerable amount of thought. It will either need good repetition of vocabulary, through reading and tasks, or quality of involvement. It is important to create a need, desire, and use for English vocabulary, as well as encouraging the students to improve upon their vocabulary. An incidental approach to vocabulary will become far more effective if taught correctly. CLIL Teachers need to be aware of vocabulary in their teaching, and the value in tasks that help with the retention of vocabulary if they want to see an improvement in their students' vocabulary.

One of the main issues to actually doing this, which I found in this study, was the willingness of the teachers to properly implement the language portion into the CLIL class. The teachers were very wary of focusing too much on the language, as they feared it would harm the content portion. This is a problem, as it could lead to a devaluation of language in the CLIL classes. If the language is not seen as part of the goal of the subject, then it is essentially rendered useless. Therefore I would advocate for a better integration of the language portion of CLIL. It seems as if the language portion is viewed as an obstacle, rather than a goal. When teaching CLIL, part of the purpose has to be teaching and improving language. While there is an understandable fear of the content suffering because of this, I would argue that if content is your primary concern with the subject then you should not use the CLIL method. CLIL will inevitably take time away from the content, and the return of this needs to be an improved language for the students. Achieving this is not an easy task however.

In the interviews with the CLIL teachers they explained that they had chosen to use the Norwegian textbook because they felt that the language used in English equivalents was too advanced for their students, and contained too many new words. This is a legitimate concern, too much new vocabulary will make it very difficult for many students to read. There are two things I would like to address with this however. Firstly, there seems to be a lack of precision when it comes to the English vocabulary of students in the Norwegian school system. The teachers claimed that the vocabulary in many English texts was too challenging for the students, while they had few ways of properly measuring the vocabulary of their students. As I have attempted to show in this thesis, there are some important benchmarks, such as knowing the first 2000 words of the English language that are incredibly useful for using English. This is something that should be measured, and implemented into the school system. As I mentioned in chapter 3, the subject-curriculum in English is quite vague when discussing vocabulary, this is something that could benefit from being altered. Focusing more on precise vocabulary knowledge, and making attempts at measuring this could help improve the level of English in Norwegian students. This is especially important for weaker students, finding the holes in their vocabulary knowledge and fixing these would be possible with more accurate measurements of vocabulary.

The second part of the issue with using a Norwegian textbook that I would like to mention is that without an English textbook, the CLIL instruction becomes severely weakened. What I feel is important to point out is that the teachers in this study spent much of their time searching for articles and texts, and finding little of real use. They also spent much of their time translating important subject-specific words that were needed. This time could have been saved with a proper textbook, and their time could have been better spent on other matters. I believe that time is a very real problem for CLIL teachers. Since they worry that the language will hurt their ability to teach the proper amount of content, leaving them feeling strapped for time. While I think it is advisable, and possible, to find an equivalent English book to use as the main textbook in CLIL, I would argue that CLIL needs a better infrastructure around it. Having properly developed textbooks for CLIL instruction would be very valuable, especially for the vocabulary learning in CLIL. While I understand that many teachers do not feel they need to use textbooks in their teaching, I would argue that textbooks can be extremely useful

when it comes to vocabulary. This is especially true in CLIL, where a textbook that manages to weave together language learning with content learning could be an efficient and useful tool. Vocabulary learning needs repetition and involvement for the best effect, which such a book could possibly achieve. A good textbook could also lighten the load of the CLIL teacher, as it is a difficult task at the present time to teach CLIL properly.

Further Research

As I have stated in this thesis, the extent to which this study can be generalized is very limited considering it only contained two small CLIL classes who were tested along with two EFL classes. Which makes for a very small sample, in addition to this I did not conduct a pre- and post-test, which means I could not measure their vocabulary learning and retention of individual words.

Therefore, a goal for further research should be to continue testing the vocabulary development in CLIL classes. Using pre- and post-tests would allow for a much better picture of the actual vocabulary development of the students in CLIL classes. With this both a breadth and a depth test could be used to test vocabulary knowledge. If any such research is to be done I would also stress the importance of using a mixed-methods approach, as this can reveal important information that would not be found using only vocabulary tests. This is important because there are many factors that could influence the vocabulary development of CLIL students, such as their reading habits outside of school or other use of English. Additionally, the teacher is an important factor in CLIL, since the teaching can vary so much between teachers it is important to pick up on this as well. As the teacher could be a very big influence on the vocabulary learning of the students.

It would also be of interest to look at the materials used in CLIL classes, testing classes who use English textbooks in CLIL, looking at the vocabulary in these books and using that data

together with the data from vocabulary tests could yield some very interesting results. The comparison of the vocabulary of the book with the vocabulary of the students could prove very interesting. Again I would stress the use of a mixed-methods approach, as it would be important to pick up on how the students dealt with new vocabulary, and how students who encountered too much new vocabulary dealt with this. A simpler study testing CLIL classes who use an English textbook with CLIL classes who do not use an English textbook could also be interesting, and again mixed-methods would be vital to ensure that factors impacting on the vocabulary learning are caught.

Concluding Remarks

Having worked on this thesis I feel that I have come to a new understanding regarding vocabulary learning in English and in CLIL. I will certainly be eager to try and measure and develop the vocabulary of my students in my future career. I firmly believe that CLIL can be an incredibly useful and exciting way to teach English, and help improve the vocabulary of students. In my own teaching I hope to get the opportunity to teach CLIL at some point, and use what I have learned in this thesis to help improve the language and vocabulary of the students. I believe that vocabulary is at the very core of the English language, and without it we are rendered helpless in the language.

Litteraturliste

Ary, D., Jacobs, L.C., & Sorenson, C., (2010) *Introduction to Research in Education*. (8th edition) Belmont, CA: Wadsworth Cengage Learning

Calfee, R., & Sperling, M. (2010) *Mixed Methods: Approaches to Language and Literacy Research*. New York, NY: Teachers College Press

Coady, J., (1997) *L2 Vocabulary Acquisition Through Extensive Reading*. In Coady, J. & Huckin, T. (Eds), *Second Language Vocabulary Acquisition* (pp.225-237). Cambridge: Cambridge University Press

Dalton-Puffer, C., (2007) *Discourse in Content and Language Integrated Learning (CLIL) Classrooms*. Amsterdam: J Benjamins

Gass, S., (1988) *Integrating Research Areas: A Framework for Second Language Studies*. *Applied Linguistics*, 9, (2), 198-217.

Hellekjær, G. O., (1996) *Easy Does It: Introducing Pupils to Bilingual Instruction*. *Språk og språkundervisning*, 3, s.9-14

Hellekjær, G. O., (2005) *The Acid Test: Does Upper Secondary EFL Instruction Effectively Prepare Norwegian Students for the Reading of English Textbooks at Colleges and Universities*, Oslo: Universitetet i Oslo (doctorate).

Hellekjær, G. O., & Hopfenbeck, T., (2012) *Rapportene fra FoU-prosjektene, Fokus på Språk, CLIL: Kombinert fag- og engelskopplæring i videregående skole*, 28, 84-124.

Hill, M., & Laufer, B., (2003) *Type of Task, Time-on-task and Electronic Dictionaries in Incidental Vocabulary Acquisition*, *IRAL- International Review of Applied Linguistics in Language Teaching*, 41 (2) 87-106.

Hulstijn, J., & Laufer, B., (2001) *Some Empirical Evidence for the Involvement Load Hypothesis in Vocabulary Acquisition*. *Language Learning* 51 (3), 539-558.

Krashen, S., (1989) *Language Acquisition and Language Education: Extensions and Application*. New York, NY: Prentice-Hall International

Lee, Y., & Hirsh, D., (2012) *Quality and Quantity of Exposure in L2 Vocabulary Learning*. In Hirsh, D. (Ed.), *Current Perspectives in Second Language Vocabulary Research* (pp.79-116). Bern: Peter Lang AG

Lightbown, P. M., & Spada, N., (2006) *How Languages are Learned*. Oxford: Oxford University Press

Matsuoka, W., (2012) *Searching for the Right Words: Creating Word Lists to Inform EFL Learning*. In Hirsh, D. (Ed.), *Current Perspectives in Second Language Vocabulary Research* (pp.151-177). Bern: Peter Lang AG

Melka, F., (1997) *Receptive vs Productive Aspects of Vocabulary*. In Schmitt, N & McCarthy, M. (Eds), *Vocabulary Description Acquisition and Pedagogy* (pp.84-102). Cambridge: Cambridge University Press

Milton, J., (2009) *Measuring Second Language Vocabulary Acquisition*. Bristol: Multilingual Matters

Moon, R., (1997) *Vocabulary Connections: Multi-word Items in English*. In Schmitt, N & McCarthy, M. (Eds), *Vocabulary Description Acquisition and Pedagogy* (pp.40-63). Cambridge: Cambridge University Press

Nation I. S. P. (1990) *Teaching and Learning Vocabulary*. New York, NY: Newbury House

Nation, P., & Newton, J., (1997) *Teaching Vocabulary*. In Coady, J. & Huckin, T. (Eds), *Second Language Vocabulary Acquisition* (pp.238-254). Cambridge: Cambridge University Press

Paribakht, T. S., & Wesche, M., (1997) *Vocabulary Enhancement Activities and Reading for Meaning in Second Language Vocabulary Acquisition*. In Coady, J. & Huckin, T. (Eds), *Second Language Vocabulary Acquisition* (pp.174-200). Cambridge: Cambridge University Press

Read, J., (1997) *Vocabulary and Testing*. In Schmitt, N & McCarthy, M. (Eds), *Vocabulary Description Acquisition and Pedagogy* (pp.303-320). Cambridge: Cambridge University Press

Read, J., (2000) *Assessing Vocabulary*. Cambridge: Cambridge University Press

Schmitt, N., (2000) *Vocabulary in Language Teaching*. Cambridge: Cambridge University Press

Svenhard, B. W., (2012) *Fokus på Språk, CLIL: Kombinert fag- og engelskopplæring i videregående skole*, 28

Sylven, L. K., (2010) *Teaching in English or English Teaching? On the Effects of Content and Language Integrated Learning on Swedish Learners' Incidental Vocabulary acquisition*. Gøteborg: Intellecta Infolog Gøteborg

Curriculums:

Kunnskapsdepartementet. (2010). *National Curriculum for Knowledge Promotion in Primary and Secondary Education and Training. English Subject Curriculum*. Oslo: Ministry of Education and Research.

Appendices

Appendix 1: X-lex test

Appendix 2: Survey

Appendix 3: Interview Guides

Appendix 1

	Jeg vet hva dette ordet betyr	Jeg er usikker på hva ordet betyr
Ornament		
Hardly		
Stamp		
Rough		
Migrate		
Mature		
Go		
Vickery		
Criminal		
Warm		
Risk		
Apartment		
Westfold		
Puzzle		
Cage		
Man		
React		
Attract		
Probable		
Cardboard		
Anxious		
Stand		
Gentle		
Gumm		
Just		

	Jeg vet hva dette ordet betyr	Jeg er usikker på hva ordet betyr?
Eckett		
Wedge		
Upset		
Thick		
Vain		
Bring		
Vital		
Causticate		
Look		
Glory		
Persuade		
Grand		
Avoid		
Tindle		
Collect		
Scatter		
Evening		
Budget		
Wife		
Gallimore		
Steam		
Enclose		
Previous		
Worry		
Manly		

	Jeg vet hva dette ordet betyr	Jeg er usikker på hva ordet betyr
Dozen		
Relation		
Grass		
Stillhard		
Samphirate		
New		
Organise		
Accuse		
Victory		
Commerce		
Sense		
Hammond		
Start		
Item		
Reaction		
Fierce		
Moreover		
Brighten		
Group		
Hyslop		
Drum		
Trick		
Sack		
Easy		
Fertile		

	Jeg vet hva dette ordet betyr	Jeg er usikker på hva ordet betyr
Baldry		
Arrange		
Complain		
Muscle		
Tail		
Crop		
Hear		
Offense		
Peritonic		
Anyone		
Ridall		
Ridiculous		
Splash		
Remind		
Steel		
Park		
Limp		
Daily		
Fishlock		
Feel		
Deny		
Obsolation		
Solemn		
Select		
Cow		

	Jeg vet hva dette ordet betyr	Jeg er usikker på hva ordet betyr
Curious		
Sudden		
Restore		
Snowy		
Bring		
Effactory		
Rain		
Inform		
Waygood		
With		
Century		
Oak		
Stream		
Military		
Sandy		
Lessen		
Cliff		
Both		
Darroch		
That		

Appendix 2

Teacher Interview:

- 1. What is your background?**
 - How long have you been teaching?**
 - Any previous experience with CLIL**
- 2. How do you experience teaching CLIL/EFL**
 - Are there any particular challenges in teaching CLIL?**
 - Is the program voluntary?**
- 3. What materials do you use in your class?**
 - What did you consider when selecting the textbook**
 - Was vocabulary a concern when selecting materials?**
 - Do you consider new vocabulary when selecting materials?**
 - Do you try to introduce specific vocabulary through the materials?**
- 4. How do you teach vocabulary in your class?**
 - When you start a new topic do you introduce the class to any new terminology?**
 - Do you teach specific strategies when learning vocabulary?**

Examples?

 - Do you teach vocabulary directly? Examples?**
 - Do you teach vocabulary indirectly? Examples?**

- (CLIL) Are there any specific challenges regarding vocabulary in CLIL?

5. Is there anything else you would like to add?

Appendix 3

Student Interview:

- 1. Can you tell me a bit about your English education?**
 - How many years have you had English?**
 - Do you have any English speaking family?**
 - Which class are you in currently?**
 - Why did you want to participate in a CLIL class?**
- 2. How often do you use English outside of school?**
 - Do you read many English books at home?**
 - Do you speak English at home or outside school?**
 - Do you write English at home or outside school?**
- 3. What do you do when you encounter an unfamiliar word in an English text?**
 - Do you have any other ways of dealing with this? Other strategies?**
 - Are you more**
- 4. Do you feel that your vocabulary has improved through CLIL?**
 - Do you feel more confident when you write or speak English?**
- 5. Is there anything else you would like to add?**

Appendix 4

Spørreskjema til CLIL-klassen på 10.trinn

Spørsmål om din lesing av engelske tekster i CLIL-faget

Under følger noen spørsmål om hvordan du opplever lesing av tekster i CLIL-faget. Gi kun ett svar per spørsmål.

1. Hvor raskt leser du engelske tekster i CLIL-faget?

Veldig sakte Middels Raskt og enkelt

1 2 3 4 5

2. Hvor ofte opplever du en CLIL-tekst som utfordrende?

Aldri Noen ganger Ofte Nesten Alltid Alltid

3. Hvis du opplever at en CLIL-tekst er utfordrende, fortsetter du å lese den?

Ja Nei

4. Hvis ja, hvor ofte opplever du at du forstår teksten etter hvert?

Aldri Noen ganger Ofte Nesten Alltid Alltid

5. Hvor ofte gir du opp å lese fordi teksten er for vanskelig?

Aldri Noen ganger Ofte Nesten Alltid Alltid

6. Hvor ofte opplever du at læreren din oppfordrer deg til å arbeide med utfordrende tekster?

Aldri Noen ganger Ofte Nesten Alltid Alltid

7. Hvor ofte opplever du at læreren din hjelper deg med å mestre utfordrende tekster?

- Aldri Noen ganger Ofte Nesten Alltid Alltid

Spørsmål om din bruk av lesestrategier i CLIL-faget

Det er flere måter å arbeide på for å forstå tekster. Nedenfor følger noen spørsmål om hvordan du leser for å få med deg innholdet i tekstene. Gi kun ett svar per spørsmål.

8. Når jeg leser en tekst, leser jeg gjennom den kun en gang

- Aldri Noen ganger Ofte Nesten Alltid Alltid

9. Når jeg leser en tekst leser jeg raskt gjennom den en gang, for deretter å lese nøye gjennom den etterpå

- Aldri Noen ganger Ofte Nesten Alltid Alltid

10. Når jeg leser en tekst første gang stopper jeg opp ved ukjente ord

- Aldri Noen ganger Ofte Nesten Alltid Alltid

11. Jeg leser tekster nøye for å få med meg alle detaljer

- Aldri Noen ganger Ofte Nesten Alltid Alltid

12. Når jeg leser en tekst understreker jeg ord eller skriver ned viktige stikkord

- Aldri Noen ganger Ofte Nesten Alltid Alltid

13. Når jeg leser skriver jeg sammendrag av teksten

Aldri Noen ganger Ofte Nesten Alltid Alltid

14. Når jeg leser forsøker jeg å organisere teksten ved å for eksempel lage et tankekart eller lignende

Aldri Noen ganger Ofte Nesten Alltid Alltid

15. Når jeg leser forsøker jeg å forstå stoffet bedre ved å knytte det til noe jeg kan fra før

Aldri Noen ganger Ofte Nesten Alltid Alltid

16. Jeg tar ofte pauser mens jeg leser for å tenke over hva jeg har lest

Aldri Noen ganger Ofte Nesten Alltid Alltid

17. Når jeg har lest ferdig kontrollerer jeg hvor mye jeg har forstått av teksten

Aldri Noen ganger Ofte Nesten Alltid Alltid

Spørsmål om lesetesten du nettopp hadde

Under følger noen spørsmål om leseprøven du nå har hatt. Svar på hvor enig eller uenig du er i følgende utsagn. Kun ett svar per spørsmål.

18. Jeg var motivert for å gjøre mitt beste på leseprøven

Svært uenig Uenig Enig Svært Enig

19. Det var viktig for meg å gjøre det bra på leseprøven

Svært uenig Uenig Enig Svært Enig

20. Jeg opplevde teksten i leseprøven som mer utfordrende enn tekster jeg leser i CLIL-faget?

Svært uenig Uenig Enig Svært Enig

21. Jeg arbeidet med å forstå teksten uten å gi opp, selv om den var vanskelig

Svært uenig Uenig Enig Svært Enig

22. Jeg brukte lesestrategier for å forstå teksten i leseprøven

Svært uenig Uenig Enig Svært Enig

23. Jeg gjorde mitt beste på leseprøven

Svært uenig Uenig Enig Svært Enig

Spørsmål om Ukjente ord

Nedenfor følger noen spørsmål om hva du gjør når du møter ukjente ord, svar på hvor ofte du gjør følgende. Kun ett svar per spørsmål.

24. Hvor ofte møter du nye og ukjente ord?

Aldri Noen ganger Ofte Nesten Alltid Alltid

25. Hvor ofte stopper du opp når du leser engelsk på grunn av ord du ikke forstår?

Aldri Noen ganger Ofte Nesten Alltid Alltid

26. Hvor ofte slår du opp i en engelsk ordbok?

Aldri Noen ganger Ofte Nesten Alltid Alltid

27. Hvor ofte gjetter du hva ordet betyr ut i fra sammenhengen?

Aldri Noen ganger Ofte Nesten Alltid Alltid

28. Hvor ofte spør du læreren?

Aldri Noen ganger Ofte Nesten Alltid Alltid

29. Hvor ofte spør du foreldrene dine eller andre hjemme?

Aldri Noen ganger Ofte Nesten Alltid Alltid

30. Hvor ofte spør du venner eller medelever?

Aldri Noen ganger Ofte Nesten Alltid Alltid

31. Hvor ofte overser du ordet og fortsetter å lese

Aldri Noen ganger Ofte Nesten Alltid Alltid

32. Hvor ofte gir du helt opp å lese?

Aldri Noen ganger Ofte Nesten Alltid Alltid

Litt om din egen bakgrunn

33. Gutt Jente

34. Hvilke språk er ditt førstespråk (morsmål)?

Norsk Engelsk Annet

35. Hvis du IKKE er født i Norge, hvor gammel var du da du kom til Norge?

Alder: _____

Hvis du var yngre enn 12 måneder, skriver du null (0)

Hvis du er født i Norge, hopper du over dette spørsmålet.

36.1 Er en eller begge av foreldrene dine fra Norge?

Ja

Nei

36.2 Hvis nei på 27.1, er en eller begge av foreldrene dine fra engelskspråklige land?

Ja

Nei

37. Hvor ofte snakker du engelsk hjemme?

Aldri

Sjelden

Månedlig

Ukentlig

Daglig

38. Har du bodd i et engelsktalende land?

Nei

Ja, i mindre enn 12 måneder

Ja, 12 måneder eller mer

39. Har du gått på skole utenfor Norge?

Nei

Ja, i et engelsktalende land

Ja, i et ikke-engelsk talende land

Litt om din skolebakgrunn

40.1 Har du undervisning i engelsk i et *ikke*-språkfag, som for eksempel historie, samfunnsfag eller reiligon?

Ja

Nei

40.2 Hvis ja kan du anslå i prosent hvor mye av undervisningen som er på engelsk?

_____ %

41. Hvilken karakter fikk du i standpunkt i engelsk skriftlig på 9.trinn?

1 2 3 4 5 6

42. Hvilken karakter fikk du i standpunkt i engelsk muntlig på 9.trinn?

1 2 3 4 5 6