Progress in English academic vocabulary use in writing among CLIL and non-CLIL students in Sweden[*]

EVA OLSSON
University of Gothenburg

Abstract
One of the purposes of instruction based on content and language integrated learning (CLIL) is in many cases to prepare students for higher education, but few studies have investigated the effect of CLIL on academic language. In this study, progress in English academic vocabulary use among CLIL (N=146) and non-CLIL (N=84) students in Swedish upper secondary school is compared. Two different academic wordlists, the *Academic Word List* (AWL; Coxhead 2000) and the *Academic Vocabulary List* (AVL; Gardner & Davies 2014), were used for the analysis of academic vocabulary in four writing assignments over three years. The results indicate that CLIL students use academic vocabulary to a greater extent already when they begin CLIL education but their use of academic vocabulary does not progress more than among non-CLIL students. The results also indicate that for the purpose of examining progress in academic vocabulary use, the AVL seems to be a more useful standard of reference than the AWL; its higher coverage allows for more detailed descriptions of progress. Only the development indicated by the AVL was supported by other findings.
Keywords: CLIL, academic vocabulary, writing

1. Introduction

In *Content and Language Integrated Learning* (CLIL), the overall purpose is to increase, simultaneously, proficiency in a second language (L2)[2] and subject knowledge, by using the L2 as a language of instruction (European Commission, 1995; Coyle, Hood & Marsh 2010). For many CLIL students, the goal is to prepare for higher education, where high proficiency in English is often regarded as a prerequisite for academic success also in countries where English is not a native language (Nunan 2003, Doiz, Lasagabaster & Sierra 2011). As the global use of English is expanding, high proficiency in academic English is a key factor for empowerment (Corson 1997, Norton 1997, Cummins 2008). Hence, English academic vocabulary knowledge is imperative, and investigating academic vocabulary growth among L2 learners highly relevant. However, very few studies on CLIL have focused on academic language use (Dalton-Puffer 2011).

[2] In this study, L2 refers both to second languages spoken in a country, e.g French in the English-speaking part of Canada, and foreign languages, e.g. English in Spain.

In the present study, progress in English academic vocabulary use in content-based essays written by CLIL and non-CLIL students in Swedish upper secondary school is compared over three years. The definition of academic vocabulary is, however, not clear-cut (Nation 2001, Baumann & Graves 2010). Hence, two different academic vocabulary lists were used for the analysis of academic vocabulary in students' essays: Coxhead's (2000) *Academic Word List*, the AWL, and Gardner and Davies' (2014) *Academic Vocabulary List*, the AVL. In the compilation of these lists, partially different principles were applied, resulting in lists that only overlap to a certain extent. The usefulness of the two lists for the purpose of describing progress in academic vocabulary use in students' essays is investigated and compared.

Thus, this study has a dual focus, CLIL and academic vocabulary, for the investigation of the possible impact of CLIL on progress in academic vocabulary.

2. CLIL

CLIL was established as a term in the 1990s for describing an educational approach which was assumed to enhance language proficiency among the young generation in the European Union, promoting personal and professional mobility in the Union (European Commission 1995, Coyle et al. 2010). The idea of integrating content and language instruction was not new, however, when it was introduced in the European Union; particularly in Canada, such integration had been well established since the 1960s but under a different name: *immersion* (see e.g. Bruck, Lambert & Tucker 1974, Swain & Lapkin 1982). Although there are differences between Canadian immersion and European CLIL – teachers are, for instance, more often native speakers of the target language in immersion than in CLIL – they can, nevertheless, be regarded as realisations of the same basic idea: non-language content is used as a vehicle for promoting second language proficiency (Lasagabaster & Sierra 2009, Cenoz, Genesee & Gorter 2014). Drawing on theories on L2 learning stressing the importance of meaningful input (Krashen, 1982), output (Swain, 1995, 2000) and interaction (Long, 1996) for L2 development, the theoretical assumption underlying CLIL is that if students are introduced to new and complex linguistic systems through academic subject content, meaning and language will be connected and thus learning will be enhanced (Genesee & Lindholm-Leary 2013). In practice, however, there is often a strong focus on content rather than language in CLIL classrooms; consequently, Lyster (2007) stresses the importance of a dual focus, i.e. on both language form and subject content for enhanced learning.

Research on the effects of CLIL on L2 proficiency mainly shows positive results for CLIL students in comparison with non-CLIL students: CLIL students often score higher in L2 testing and their receptive and productive L2 vocabulary tends to be larger, including low-frequency words to a greater extent (Dalton-Puffer 2011). Ruiz de Zarobe (2008, 2010) found that Spanish CLIL students in upper secondary school outperformed non-CLIL students with regard to choice and use of English vocabulary in speech and in writing. Development over time was more

positive in the CLIL groups also with regard to other aspects of writing proficiency, e.g. the organisation of texts and the use of grammar. Similar results were reported in Jexenflicker and Dalton-Puffer's (2010) study among Austrian engineering students (aged 16): CLIL students' vocabulary range, use of grammar and organisation of texts were judged significantly stronger than those of non-CLIL students. Lo and Murphy (2010) found that receptive and productive vocabulary increased significantly more among English immersion students (aged 11-15) in comparison with students who studied English in a traditional language class in Hong Kong. Thus, CLIL instruction seems to be advantageous for different aspects of L2 writing proficiency.

However, some of the positive results have been disputed by results from other studies. In a longitudinal study of English receptive vocabulary knowledge among students in upper secondary school in the Netherlands, Admiraal, Westhoff and de Bot (2006) showed that initial differences in proficiency level between CLIL and non-CLIL groups remained at the same level rather than increased; CLIL instruction did not widen the gap. Rumlich (2013) and Bruton (2011) point out that very few studies of CLIL actually include pre-tests, making it difficult to draw any firm conclusions on the effects of CLIL.

In the Swedish context, the positive impact of CLIL instruction on L2 proficiency reported above has yet to be confirmed (Sylvén 2013). Hyltenstam's (2004) survey of Swedish CLIL research concludes that none of the studies reported showed that English proficiency among Swedish CLIL students actually improved to a greater extent than among non-CLIL students. Although Sylvén's (2004) investigation of vocabulary range showed that CLIL groups scored higher on vocabulary tests already in the pre-test and that they improved their results more than non-CLIL groups, also certain background factors turned out to be influential, especially parents' level of education and students' use of English in their spare time.

A study of Swedish CLIL and non-CLIL students' vocabulary use in writing in comparison with native English students' writing showed that the CLIL group used more varied vocabulary than the non-CLIL group (Edlund 2011). In fact, the CLIL group varied their use of vocabulary to the same extent as the native English group of students. However, as development over time was not investigated in this study, no information about base-line proficiency is available.

The present study is part of a large-scale research project: *Content and Language Integration in Swedish Schools (CLISS)*, where the main objective is to study and compare the development of academic language proficiency in both Swedish and English among CLIL and non-CLIL students in upper secondary school (see Sylvén & Ohlander 2014). Of special interest here are the baseline results of the *Vocabulary Levels Test* (VLT; Nation 2001) reported by Sylvén and Ohlander (2014) and involving the same students as in the present study. The results showed that CLIL students outperformed non-CLIL students having a more extensive receptive English vocabulary already at the start of their CLIL education with a mean score of 112 in comparison with 99 for the non-CLIL group out of 150 items.

The difference between groups was significant ($p<.001$). Alongside sections covering vocabulary items at different frequency levels, the VLT included a section with 39 vocabulary items selected from the AWL (Coxhead 2000). In this section too, CLIL students scored significantly better than non-CLIL students.

3. Academic vocabulary
As academic vocabulary may be broadly defined as vocabulary that occurs more often in academic contexts than in other contexts, corpus-based academic vocabulary lists can be highly useful in studies of academic vocabulary (Nation, 2001). Academic vocabulary is often classified as either *domain-specific* or as *general* (Baumann & Graves 2010:6). Domain-specific vocabulary consists of content-specific words used in different disciplines, such as Archaeology or Medicine, whereas general academic vocabulary consists of words that appear across many or all disciplines (cf. Hyland & Tse 2007). The present study is limited to the analysis of general academic vocabulary. Since the study focuses on progress over time, it involves several writing assignments on various topics (see section 5.2); thus, it seems relevant to study vocabulary that can be found and used in different contexts rather than specialised vocabulary that varies according to context.

As already mentioned, two different academic vocabulary lists are used in the study: the *Academic Word List* (AWL; Coxhead 2000) and the *Academic Vocabulary List* (AVL; Gardner & Davies 2014). The AWL lists 570 word families[3] extracted from a corpus of 3.5 million running words of written academic text about the Arts, Commerce, Law and Science. The corpus mainly consists of texts from academic journals, university textbooks and scientific parts of corpora such as the Brown Corpus (Francis & Kucera 1982). The largest parts of the texts are from New Zealand and the rest from other English-speaking countries, e.g. the USA and Great Britain. To be included in the AWL, a word had to appear 100 times in the academic corpus – and at least 25 times in all of the four sections of the corpus (Coxhead 2000:222–227). The AWL has been widely used in a large number of studies, showing consistent coverage of approximately 10 % of vocabulary across disciplines (Coxhead 2011).

Gardner and Davies (2014), while acknowledging the significance of Coxhead's AWL in teaching and research, also point to certain weaknesses; above all, the use of word families when determining word frequencies. They specifically note that a word family may contain a large number of members with distinct meanings. The word family *react*, for instance, contains thirteen members, some of which do not share the same meaning: *react, reactionary* and *reactor* exemplify their point as these words etymologically belong to the same family but nevertheless express completely different meanings (Gardner & Davies 2014:307).

---

[3] A word family is defined as a stem plus affixed forms where suffixes and prefixes are added to the stem (Bauer and Nation 1993).

Furthermore, Gardner and Davies are troubled by the association between the AWL and West's by now partially obsolete *General Service List* (GSL) from 1953. Only words outside the 2000 most frequent word families in English as accounted for in the GSL were included in the AWL, but apparently frequently used words in 1953 may not be equally frequent today. Some words that might be expected in an academic wordlist are not listed in the AWL since they appeared in the GSL and moreover, Gardner and Davies show that the AWL actually contains many words that are highly frequent today (Gardner & Davies 2014:308-9).

To avoid some of the problems discussed above, lemmas were used instead of word families when the AVL was compiled. The AVL contains 3000 words from the academic section of the *Corpus of Contemporary American English* (COCA; Davies 2012), which includes more than 120 million words out of the total 425 million words in the COCA. The texts in the AVL corpus were published in the USA, covering nine disciplines. More than half of the words in the academic corpus come from academic journals. Other sources are topic-specific magazines and newspaper articles.

To be included in the AVL, a word had to be at least 50% more frequent in the academic corpus than in the non-academic part of the COCA. The word also had to occur in at least seven out of the nine disciplines with at least 20% of the expected frequency, based on the total number of occurrences in the academic corpus. Also, words that occurred more than three times as often as the expected frequency in any specific sub-corpus were excluded. Thus, highly domain-specific as well as highly frequent non-academic vocabulary was not included (Gardner & Davies 2014:313-16). However, a number of words in the AVL may also be used in everyday situations, e.g. *group, use, information, short, change* and *important* – hardly vocabulary items that need particular attention when learning academic writing.

Gardner and Davies tested the coverage of the AVL and the AWL on the academic sections of the COCA (from which the AVL was built) and *The British National Corpus* (BNC; Nation 2004). In both cases they found that the AVL had a higher coverage. It covered 13.8% of the academic section in the COCA and 13.7% of the academic section of the BNC in comparison with the AWL, which covered 7.2% of the section in the COCA and 6.9% of the section in the BNC.

In the present study, both lists are used as standards of reference for 'defining' academic vocabulary. However, this does not imply that either of the lists covers all academic words in the English language, but only the most frequent academic vocabulary, given certain principles for selection, in the corpora from which the lists were constructed.

4. Aims

As will already have appeared, the primary aim of this study is to investigate and compare the development of English academic vocabulary among CLIL and non-CLIL students for the purpose of investigating if there is any difference in the progress of academic vocabulary use in writing between CLIL and non-CLIL students. Since English is used as a language of instruction in CLIL in the present

study, it is a reasonable hypothesis to expect that CLIL students are exposed to, practise and learn English academic vocabulary to a greater extent than non-CLIL students.

For the analysis of progress in academic vocabulary use in students' writing, the two academic vocabulary lists already mentioned, the AWL and the AVL, were used. This gives rise to a secondary aim, namely to compare the analyses based on the two lists and to investigate the usefulness of them for the specific purpose of describing development in academic vocabulary in students' writing. A reasonable hypothesis is that analyses using either of the two lists may be expected to indicate similar development over time, on the assumption that both lists are equally valid for the analysis of progress in academic vocabulary use.

5. Material and methods
The material used in this study mainly consists of students' essays, which were analysed using corpus-based methods. In this section, the material and the methods of analysis are further described, starting with the participants.

5.1 The students
A total of 230 students, aged 16-19, from three upper secondary schools in Sweden were involved in the study: 146 CLIL students (46 male and 100 female) and 84 non-CLIL students (36 male and 48 female).

One of the schools is an international school where English is used as the language of instruction in all subjects, except in language classes (e.g. French or German). At this school, all students involved in the present study followed CLIL programmes. At the other two schools, both CLIL and non-CLIL classes participated. At these two schools, students could choose if they wanted to follow a CLIL programme where both English and Swedish were used as languages of instruction in most subjects, or a regular programme where Swedish was used as the language of instruction in all subjects, except in language classes.

In all, eight classes, five CLIL and three non-CLIL, participated in the study. The students followed programmes where the Natural Sciences, the Social Sciences or Business Management and Economics were the main subjects. All of these programmes are preparatory for higher education.

Table 1. An overview of the participating classes

CLIL (N=146)	Non-CLIL (N=84)
2 Natural Science classes 2 Social Science classes 1 Business Management and Economics class	1 Natural Science class 2 Business Management and Economics classes

English was a compulsory subject in all of these programmes[4]. Thus, non-CLIL students encountered English at school mainly in English language classes whereas CLIL students encountered English both in English language classes and in other lessons, i.e. in content and language integrated lessons.

5.2 Writing assignments
The material used for the analyses of academic vocabulary consists of 525 texts based on four different writing assignments, where students were asked to write argumentative and explanatory essays in English, covering topics mainly related to the Natural and the Social Sciences – subjects studied by all participating students (although to a greater or lesser extent, depending on the main profile of their programme). In CLIL classes, these subjects were studied at least partly in English, whereas non-CLIL classes studied them in Swedish.

The first writing assignment, related to the Natural Sciences, was administered during students' first term in upper secondary school, thus providing the baseline data. The second and third writing assignments, related to the Social Sciences, were given in the second year, and the last writing assignment, related to the Natural Sciences, was given during the students' third and final year in upper secondary school. In Table 2, the topics and text types covered by the writing assignments are shown. The number of essays collected on each occasion is also indicated. In Appendix A, a table showing the average length of the essays is provided.

Table 2. Writing assignments

Assign-ment	School Year	Title	Text type	Related subject	Number of essays
1	Year 1 autumn 2011	*For or against nuclear power*	Argumentative essay	Natural Sciences	146 (94 CLIL 52 non-CLIL)
2	Year 2 autumn 2012	*Matters of gender and equality*	Expository essay	Social Sciences	126 (80 CLIL 46 non-CLIL)
3	Year 2 spring 2013	*Ways to political and social change – violence or non-violence*	Argumentative essay	Social Sciences	138 (83 CLIL 55 non-CLIL)
4	Year 3 autumn 2013	*Biodiversity for a sustainable society*	Expository essay	Natural Sciences	115 (80 CLIL 35 non-CLIL)

The assignments were designed for the purpose of eliciting content-based texts where academic language use could be expected. The topics covered areas included

---

[4] Courses equivalent to levels B 1.2 and B 2.1 in the Common European Framework of Reference for Languages (CEFR) are compulsory. Level B 2.2 is optional.

in the Swedish curriculum for the Natural and the Social Sciences. For each of the assignments, there was a printed instruction defining the task: to argue or to explain the topic. For inspiration, there were two pages of background information attached to the instruction, for instance short factual texts, graphs or pictures. Students were allowed 90–120 minutes to finish their essays. No word limit was set with regard to text length. The essays were written on computers[5] at school.

Some of the 230 students were not present on all occasions since they had fallen ill or were absent for other reasons. Some students changed schools and thus left the CLISS project, and some new students joined it in year 2 or 3. All essays available on each occasion were collected and used in the analysis (see Table 2). 90 students wrote both the first and the last assignments and 70 students wrote all four of them.

5.3 Corpus-based analyses of students' texts

The vocabulary in the students' essays was compared to the academic word lists, the AWL and the AVL. The main purpose of using two different standards of reference was to strengthen the validity of the analysis of development: if two different measurements indicated the same kind of development, the validity would be strengthened.

Two different web-based tools were used: for the analysis of academic vocabulary covered by the AWL, *Vocabprofile,* available at http://www.lextutor.ca/vp/eng/, was used, and for the analysis of academic vocabulary found in the AVL, an interface, available at http://www.wordandphrase.info/academic/.

Both web-based tools offer analyses of academic *types* and *tokens* (Nation 2001). The percentage of academic word forms, tokens, identified by the AWL or the AVL was noted for each of the essays. When tokens are used in an analysis, repeated words are counted every time they occur. However, it was also of interest to analyse if students used different academic words and hence, the proportion of academic types (of all types in an essay) was noted (cf. Nation 2001).

The overall use of academic vocabulary is the main focus of the study, and so the discussion mainly concerns analyses based on tokens. However, some comparisons are made with analyses based on types for the purpose of describing development in a detailed manner and illuminating differences and similarities between the two wordlists.

Statistical analyses were performed, using SPSS version 21. For each writing assignment, the average percentage of academic vocabulary was calculated for CLIL and non-CLIL students, and the standard deviation noted to indicate the dispersion of scores within groups. T-tests were used for the analysis of statistical significance of differences between CLIL and non-CLIL students. Progress in academic vocabulary use from year one to year three was analysed using statistical

---

[5] Students were asked to turn off spelling and grammar check. They were not allowed to access the Internet while writing.

regression analysis, where initial scores were taken into account when analysing final scores. Thus, particular attention was paid to the first and the last writing assignments as they provided information about students' proficiency at the starting-point and at the end. In this way development may be clearly illustrated.

For the purpose of illustrating differences and similarities in the analyses using the two lists, the academic vocabulary in one student's first and last essays was compared in detail. The student, Kim, was chosen for this case study being close to average numbers in both the AWL and the AVL analyses, thus not an extreme case. Kim's first and last essays are shown in Appendices B and C.

5.4 Assessment of 30 students' essays

For the purpose of finding evidence of whether or not development traceable in the texts had taken place between the first and the last year, four external assessors were asked to rate 30 students' first and last assignments – it was beyond the scope of the study to assess all essays. The selection of students was made after sorting the final essays according to the percentage of academic vocabulary and then picking every third essay; thus, the sample included essays with various amounts of academic vocabulary. The same students' first and last assignments were compared and the assessors were asked to choose which of the two essays they rated as the better one in terms of holistic impression of language proficiency.

For the purpose of validating if the proportion of academic vocabulary may have an impact on the rating of the texts – earlier studies have indicated that lexical richness is important for a holistic impression of writing proficiency (cf. Laufer & Nation, 1995) – another round of assessment was carried out. The same 30 students' last essays were used this time and the assessors were asked to compare each of the essays with an essay that included an average percentage of academic vocabulary. The assessors were asked to judge if each of the 30 essays was weaker, at the same level or stronger than the text of comparison. The methods used in both rounds of assessment were inspired by Pollit's (2012) method of adaptive comparative judgement, although in a simplified version.

In the design of the study, great care was taken to ensure that analyses and results would be as valid and as reliable as possible. Nevertheless, there are limitations to the study that must be considered. Two of the topics of the writing assignments related to the Social Sciences and two to the Natural Sciences; this may have favoured certain students on certain occasions. Further, the assignments covered two text types, argumentative and expository essays, which may also have been of some relevance for the outcome. (See also section 7.)

6. Results

In this section, the results of the analyses are reported. Comparisons are made between CLIL and non-CLIL students' use of academic vocabulary as identified by the AWL and the AVL.

6.1 AWL analysis
In Table 3, the results from analyses of academic tokens using the AWL as standard of reference are accounted for.

Table 3. Average proportion of AWL tokens in texts by CLIL and non-CLIL students

Coxhead (2000) AWL	CLIL		non-CLIL			
	Mean %	Standard deviation	Mean %	Standard deviation	T-value	p-value
Text 1 AWL tokens %	6.55	2.29	6.34	2.87	.49	.629
Text 2 AWL tokens %	4.10	1.98	3.37	1.38	2.22	.028*
Text 3 AWL tokens %	3.23	1.77	2.29	1.36	3.34	.001*
Text 4 AWL tokens %	4.94	2.09	4.00	1.92	2.28	.025*

* statistically significant difference between CLIL/non-CLIL

As is clear from Table 3, when the AWL is used as a standard of reference for the analysis of academic tokens in the first assignment (text 1), the difference between CLIL and non-CLIL students is slight; the small difference is not statistically significant (t= .49 *p*= .629). Thus, when the AWL is used for the analysis of academic vocabulary, results indicate that CLIL and non-CLIL groups were at the same level when they began upper secondary school. In the three following assignments, the difference between CLIL and non-CLIL students' use of vocabulary covered by the AWL increases. A statistical T-test shows that, in assignments 2, 3 and 4, the difference between the CLIL and the non-CLIL groups is statistically significant (text 2: t= 2.22 *p*= .028; text 3: t=3.34 *p*= .001; text 4: t= 2.28 *p*= .025). However, the trend is clearly negative over time; the proportion of academic tokens used by students in both groups is smaller in the last assignment than in the first. In the CLIL group, 6.6% academic vocabulary was identified by the AWL in the first assigment but only 4.9% in the last, while the figures for the non-CLIL group were 6.3% in the first assignment and 4.0% in the last. Thus, the analysis of academic tokens using the AWL as standard of reference indicates that neither the CLIL group nor the non-CLIL group increased the use of general academic vocabulary through the three years at upper secondary school; instead, the results indicate that the use of such vocabulary decreased.

6.2 AVL analysis
The results from the analysis of academic tokens using the other academic vocabulary list, the AVL, are presented in Table 4.

Table 4. Average proportion of AVL tokens in texts by CLIL and non-CLIL students

Gardner & Davies (2014) AVL	CLIL		non-CLIL			
	Mean %	Standard deviation	Mean %	Standard deviation	T-value	p-value
Text 1 AVL tokens %	7.37	2.51	5.85	2.38	3.56	.001*
Text 2 AVL tokens %	9.30	3.26	7.93	3.33	2.24	.027*
Text 3 AVL tokens %	8.47	3.60	6.65	2.86	3.14	.002*.
Text 4 AVL tokens %	12.14	3.51	10.86	3.59	1.80	.075

* statistically significant difference between CLIL/non-CLIL

When the AVL is used as a standard of reference, there is a statistically significant difference in the use of academic tokens between the CLIL and the non-CLIL groups already in the first assignment (t= 3.56 *p=* .001). Thus, when the AVL is used in the analysis, the results indicate that already from the start, the CLIL group used academic vocabulary to a greater extent than the non-CLIL group. In the following two assignments, the difference between CLIL and non-CLIL groups remains significant (text 2: t= 2.24 *p=* .027; text 3: t=3.14 *p=* .002). As indicated by the figures for standard deviation in Table 4, scores for academic tokens in text 4 vary to a great extent within groups, and thus, the difference between groups is not statistically significant (text 4: t=1.80, *p=* .075). However, in the last assignment (text 4), both the CLIL and the non-CLIL groups increased their average use of academic vocabulary considerably: the CLIL students used 7.4% academic vccabulary identified by the AVL in the first assignment, compared to 12.1 % in the last assignment while the non-CLIL students included 5.9% in the first and 10.9% in the last assignment. Hence, the trend over time is clearly positive for both the CLIL and the non-CLIL groups when the AVL is used as a standard of reference.

6.3 Comparison of development over time
In this section, a summary of the analyses of academic vocabulary using the two academic word lists is offered in the form of two figures. Figures 1 and 2 illustrate the progress of academic vocabulary use over three years when analyses are based on tokens (Figure 1) and types (Figure 2) covered by the AWL or the AVL. Tables in Appendices D and E show the average proportion of academic types (AWL/AVL) in texts written by CLIL and non-CLIL students (cf. Tables 3 and 4).

Figure 1. Development of academic vocabulary use (AWL and AVL) over three years in CLIL and non-CLIL groups (tokens)
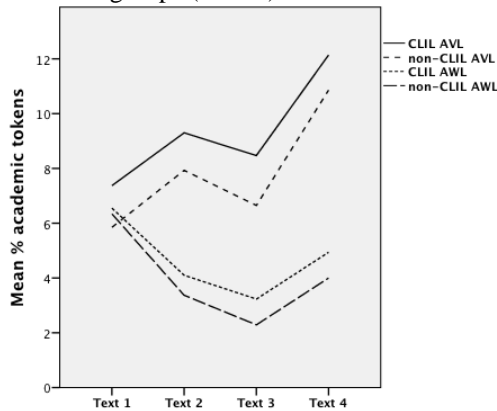

Figure 2. Development of academic vocabulary use (AWL and AVL) over three years in CLIL and non-CLIL groups (types)
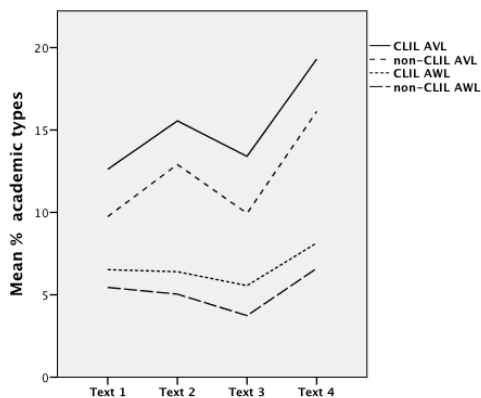

Figure 1, a visual representation of the average proportion of academic tokens reported in Tables 3 and 4, shows that there is a clear difference in development depending on which of the two different academic vocabulary lists is used in the analysis: the AVL indicates a considerably more positive development than the AWL and only the AVL indicates an initial gap between the CLIL and the non-CLIL groups. However, both the AWL and the AVL indicate a dip in academic vocabulary use in the third assignment. Different tasks may elicit academic vocabulary to a greater or lesser extent; obviously, students used a somewhat smaller proportion of academic vocabulary in the third assignment. Nevertheless, the trend is clearly more positive when the AVL is used in the analysis than when the analysis is based on the AWL.

Figure 2 shows that when the analysis is based on academic types, the AWL also indicates an initial gap between groups. Both the AVL and the AWL indicate that CLIL students use a significantly larger proportion of academic types in all four assignments than non-CLIL students (see Appendices D and E). However, also in the analysis of types, the AVL indicates a considerably more positive development than the AWL.

To draw any conclusions about the effect of CLIL on development over time, it is necessary to control for baseline differences; the final scores must be viewed in relation to the scores obtained in the first assignment. A regression analysis indicates that, regardless of which list is used, the development in the CLIL group is not more positive than in the non-CLIL group. The analysis shows that, with the baseline difference controlled for, the CLIL effect is not significant (AWL: B= .75, t=1.84, *p*= .070; AVL: B= .72, t=1.03, *p*= .308)[6]. Even if CLIL students use a larger proportion of academic vocabulary in all essays in comparison with non-CLIL students, the difference between the two groups does not increase significantly over time.

6.4 Detailed comparison of AWL and AVL: a case study

For the purpose of illustrating how the analysis of development over time can result in rather opposite conclusions – positive or negative development – the analyses of academic vocabulary in one student's, Kim's, first and last assignments are compared in detail. All words covered by the AWL and/or the AVL in Kim's essays are listed in Tables 5 and 6. After each word, the number of occurrences is indicated. At the top of each column, the number of different academic word types and the total number of academic tokens are indicated.

---

[6] The figures shown are based on the analysis of academic tokens. A regression analysis using the proportion of academic types in the texts indicates similar results: AWL: B= .97, t=1.53, *p*= .130; AVL: B= 1.77, t=1.61, *p*= .110.

Table 5 *Academic vocabulary covered by the AWL and/or the AVL in Kim's first assignment (Tokens: 8.0 % AWL / 8.2 % AVL; Types: 8.4% AWL / 13.3% AVL)*

Unique to AWL 9 types 26 tokens	**Shared** 8 types 11 tokens	**Unique to AVL** 19 types 27 tokens
area 2 energy 9 major 1 nuclear 6 odds 1 reaction 4 recover 1 relocate 1 remove 1	access 1 available 1 couple 1 reliable 1 resources 1 source 3 sources 2 specific 1	advanced 1 century 1 control 1 depending 1 example 2 gain 2 ground 1 groups 1 low 1 may 3 protection 1 provide 2 result 1 results 1 strongly 1 than 1 use 2 used 1 using 3

Table 6. *Academic vocabulary covered by the AWL and/or the AVL in Kim's last assignment (Tokens: 3.6 % AWL / 11.6 % AVL; Types: 8.1% AWL/ 21.4% AVL)*

Unique to AWL 8 types 9 tokens	**Shared** 11 types 11 tokens	**Unique to AVL** 39 types 54 tokens	(continued)
areas 1 capable 1 energy 2 generation 1 generations 1 mentally 1 physically 1 transporters 1	available 1 conclusion 1 crucial 1 cultural 1 diversity 1 factor 1 individuals 1 projects 1 requires 1 survival 1 transport 1	according 1 adopted 2 both 1 claims 1 comparing 1 conditions 1 demands 1 diversities 1 efficiently 1 exists 1 finding 1 future 2 gain 1 greatly 1 harvesting 2 helpful 1 human 4 humans 3 important 1 increase 1	increased 1 means 1 nature 1 need 6 needs 2 observe 1 part 1 population 1 produced 1 products 2 provided 1 provides 1 study 1 such 1 system 1 types 1 various 1 within 1 working 1

Table 5 shows that in the first assignment, the percentage of tokens covered by the AWL and the AVL is almost the same; the AWL identifies 8.0% of the vocabulary in Kim's first essay as academic and the AVL 8.2%. When tokens are analysed, the same word is counted each time it occurs. In Kim's essay, the word *energy,* which he uses nine times, and the word *nuclear*, which he uses six times, are identified as general academic vocabulary by the AWL but not by the AVL. This may explain why the coverage in percentage is similar for both lists in Kim's first essay even if the AVL covers a larger number of word types. When the proportion of academic types is used in the analysis of Kim's first essay, the AWL identifies 8.4% of all types as academic whereas the corresponding figure for the AVL is 13.3%. Table 5 shows that nine word types are found only in the AWL, eight are found in both word lists and 19 are found only in the AVL.

In the analysis of Kim's last essay, shown in Table 6, the difference in coverage has increased between the AWL and the AVL. Eight word types are only found in the AWL, eleven word types are found in both lists and as many as 39 word types are identified as academic only by the AVL. Thus, 8.1% of all types in Kim's last essay are identified as academic by the AWL whereas the AVL identifies 21.4% of the types as academic. The difference in proportion of academic vocabulary based on tokens is also striking: 3.4% of the words are identified as academic by the AWL in comparison with 11.6% by the AVL.

When the AWL is used for the analysis, the result is that Kim used a smaller proportion of academic vocabulary in the last essay than in the first essay. When the AVL is used in the analysis, the result is quite the opposite, i.e. Kim has substantially increased his use of academic vocabulary.

6.5 Assessment of 30 students' texts

For the purpose of validating if the development between the first and the last assignments had been positive or negative, four assessors judged the first and last essays written by 30 students, as accounted for in section 5.4. Basing their judgment on a holistic impression of language proficiency, the assessors found that in 27 out of 30 cases, the final essay was better than the first one written by the same student. Thus, the assessment indicates that a positive development has taken place.

The second part of the assessment, where comparisons were made between an essay including an average proportion of academic words and the 30 students' final essays, indicates that academic vocabulary seems to be of some importance for the holistic assessment of the essays. Essays judged as better than the text of comparison include academic vocabulary to a significantly larger extent than the rest of the essays (AWL: t= 2.25 *p*= .033; AVL: t= 2.08 *p*= .047). (See Appendix F.)

7. Discussion

The main results of this study show that CLIL students did not increase their use of general academic vocabulary more than non-CLIL students over three years. After

controlling for initial differences, whether analyses are based on the AVL or the AWL, tokens or types, the CLIL effect is not significant. Thus, higher exposure to English at school does not in itself seem to imply that CLIL students progress more in their own use of academic vocabulary than non-CLIL students. The results are to some extent unexpected since previous research has shown that CLIL tends to be beneficial for writing proficiency and for vocabulary knowledge (cf. Dalton-Puffer 2011). At the same time, the results are in line with findings reported by Admiraal et al. (2006), where initial differences in receptive vocabulary knowledge between CLIL and non-CLIL groups remained at the same level rather than increased; CLIL instruction did not widen the gap. It is, however, necessary to reflect upon the fact that the CLIL group used a larger proportion of academic vocabulary already from the start – at least when the AVL was used as the standard of reference. Possibly, it may be more difficult to increase the use of academic vocabulary from an already fairly high level, but nevertheless, the CLIL students progressed and scored higher than non-CLIL students in all assignments.

The initial gap between CLIL and non-CLIL groups was to some extent expected since CLIL is an option; students who choose to study a programme where English is used as the language of instruction have been shown to have a more positive attitude towards English and to have a higher English proficiency level already before they start CLIL in comparison with students in ordinary classes (Sylvén 2004, Lasagabaster & Sierra 2009, Rumlich 2013; Yoxsimer Paulsrud, 2014). Furthermore, baseline results reported by Sylvén and Ohlander (2014), involving the same students as in the present study, showed that CLIL students' receptive academic vocabulary was significantly larger than that of non-CLIL students. Hence, the initial difference in productive academic vocabulary between CLIL and non-CLIL groups was expected, even though receptive vocabulary is always larger than productive vocabulary (cf. e.g. Laufer & Nation 1995).

In the present study, classes with different profiles participated, which may have biased results to some extent. Even if all classes studied both the Social and the Natural Sciences, it was most likely easier to write about Biodiversity for students with the Natural Sciences as majors. Yet, since general rather than domain-specific academic vocabulary was in focus in this study, the diversity of majors was probably not as important for the outcome as it would have been if the use of domain-specific vocabulary had been investigated. Further, the assessment of students' essays showed that the proportion of general academic vocabulary seemed to influence the holistic impression of the essays, thus confirming the importance of learning such vocabulary.

However, even if the results of this study indicate that CLIL does not seem to promote development of academic vocabulary more than regular, non-CLIL education, it must be taken into account that CLIL can be practised in many different ways. As suggested by Lyster (2007) and Genesee and Lindholm-Leary (2013), the focus of many CLIL classrooms is mainly on subject content, and so language proficiency may not progress to the same extent as with a dual focus on both language form and subject content. Three different schools were involved in

the present study and it seems that further analyses of CLIL practices at those schools are necessary for the purpose of finding out how different CLIL practices may impact on progress in academic vocabulary use.

A secondary aim of the study was to compare analyses using the two different academic vocabulary lists, the AWL and the AVL. In terms of development of academic vocabulary over time, different tendencies are indicated by the two lists. The case study of one student's use of academic vocabulary clearly illustrates the difference in coverage and the opposite developmental directions indicated by the lists. Looking at the words listed from Kim's first and last assignments, it is difficult to argue that his use of academic vocabulary has not increased between the two occasions. When small samples of fairly short texts are analysed, it is necessary that the standard of reference is extensive enough – yet accurate – for progress to be detectable. The analyses using the AVL indicates that both CLIL and non-CLIL students increased their use of academic vocabulary considerably over three years. The AWL, on the other hand, indicates that both CLIL and non-CLIL students used a *smaller* proportion of academic tokens after three years in upper secondary school – clearly a depressing result from an educational perspective and definitely not in line with the holistic assessment of 30 students' first and last texts, which implied that development had actually taken place. Thus, only the AVL detected the progress indicated in the assessment of the students' essays. Consequently, the results of this study indicate that the AVL appears to be a more extensive and finely tuned instrument than the AWL for the purpose of investigating progress in students' use of academic vocabulary.

8. Concluding remarks

The main finding of this study indicates that students in Swedish CLIL education, targeting English, do not progress more in their productive use of English academic vocabulary than non-CLIL students, who follow regular education with Swedish as the language of instruction (except in language classes). Two different word lists (the AWL and the AVL) were used in the analysis and neither of them indicated any significant difference betweeen CLIL and non-CLIL groups in the progess of academic vocabulary over three years, after controlling for initial differences. Since one of the reasons why schools offer CLIL programmes is to prepare students for higher education, where proficiency in academic English is needed, the results are remarkable. The results seem to imply that attention is not paid to general academic vocabulary in CLIL to such an extent that CLIL students' productive academic vocabulary develops more compared with students in regular education. The assessment of students' essays showed that the proportion of general academic vocabulary seemed to influence judgements; thus, the relevance of knowing such vocabulary was confirmed in this study.

However, the effect of different CLIL practices was not explored here; further research in this field is clearly necessary as the focus on language may vary between schools and classrooms. Moreover, the impact of other background factors, such as extramural English, a factor shown to be of considerable importance not least as

regards vocabulary growth (Sylvén 2004, Sundqvist 2009, Olsson 2012), would also be of relevance to investigate in this connection. Obviously, there are also other aspects of academic writing proficiency, apart from general academic vocabulary use, that would be of interest to analyse, e.g. if there is any difference between CLIL and non-CLIL students in the use of grammar that is typical of academic writing and in the use of domain-specific vocabulary.

A further important finding of this study is that the AVL seems to be a more useful standard of reference than the AWL for the analysis of students' progress in productive academic vocabulary. Only the clear progress indicated by the AVL was confirmed by other findings, i.e. by the holistic assessment of students' essays. Further, the initial gap in productive academic vocabulary between CLIL and non-CLIL students indicated by the AVL (but not by the AWL) is in line with results of receptive vocabulary among the same students (Sylvén & Ohlander 2014). It seems that the higher coverage of the AVL in comparison with the AWL, due to different principles in selecting academic vocabulary, contributes to a more detailed description of students' academic vocabulary use. Hence, even though further comparison of the usefulness of the AWL and the AVL for the detection of progress is called for, the results are of methodological relevance for future research on the development of academic vocabulary use, whether in CLIL or non-CLIL contexts.

References

Admiraal, Wilfried, Gerard Westhoff & Kees de Bot (2006), "Evaluation of bilingual secondary education in the Netherlands: Students' language proficiency in English", *Educational Research and Evaluation,* 12(1):75-93. doi: 10.1080/13803610500392160

Bauer, Laurie & Paul Nation (1993), "Word families", *International Journal of Lexicography,* 6(4):253-279. http://dx.doi.org/10.1093/ijl/6.4.253

Baumann, James F. & Michael F. Graves (2010), "What is academic vocabulary?", *Journal of Adolescent & Adult Literacy,* 54(1):4-12. doi: 10.1598/jaal.54.1.1

Bruck, Margaret, Wallace E. Lambert & G. Richard Tucker (1974), "Bilingual schooling through the elementary grades: the St. Lambert project at grade seven", *Language Learning,* 24(2):183-204. doi: 10.1111/j.1467-1770.1974.tb00501.x

Bruton, Anthony (2011), "Is CLIL so beneficial, or just selective? Re-evaluating some of the research", *System,* 39(4):523-532.
doi: 10.1016/j.system.2011.08.002

Cenoz, Jasone, Fred Genesee & Durk Gorter (2014), "Critical Analysis of CLIL: Taking stock and looking forward", *Applied Linguistics,* 35(3):243-262. doi: 10.1093/applin/amt011

Cobb, Tom. *Web Vocabprofile*. Accessed 11 Dec 2015 from http://www.lextutor.ca/vp/. An adaptation of Heatley, Alex, Paul Nation & Averil Coxhead (2002). *Range and Frequency*

Corson, David (1997), "The learning and use of academic English words", *Language Learning,* 47(4):671-718. doi: 10.1111/0023-8333.00025

Coxhead, Averil (2000), A new academic word list. *TESOL Quarterly, 34*(2), 213-238. doi: 10.2307/3587951

Coxhead, Averil (2011), "The Academic Word List 10 years on: Research and teaching implications", *TESOL Quarterly,* 45(2):355-362. doi: 10.5054/tq.2011.254528

Coyle, Do, Philip Hood & David Marsh (2010), *CLIL: Content and language integrated learning*. Cambridge: Cambridge University Press.

Cummins, Jim (2008), "BICS and CALP: Empirical and theoretical status of the distinction", in Hornberger, Nancy (ed.), *Encyclopedia of Language and Education.* New York: Springer US, 487-499.

Dalton-Puffer, Christiane (2011), "Content-and-language integrated learning: From practice to principles?", *Annual Review of Applied Linguistics,* 31:182-204. doi: 10.1017/s0267190511000092

Davies, Mark (2012), *Corpus of Contemporary American English (1990-2012).* Retrieved from http://corpus.byu.edu/coca/.

Doiz, Aintzane, David Lasagabaster & Juan Manuel Sierra (2011), "Internationalisation, multilingualism and English-medium instruction", *World Englishes,* 30(3):345-359. doi: 10.1111/j.1467-971X.2011.01718.x

Edlund, Anders (2011), *Undervisning på engelska i den svenska gymnasieskolan – ett experiment med potential? En studie av tre elevgruppers engelska texter i två register.* (*Instruction in English in Swedish upper secondary school – an experiment with potential? A study of texts covering two registers involving three classes*). Licentiate thesis. Stockholm: Stockholm University.

European Commission (1995), *White paper on education and training. Teaching and learning:Towards the learning society*. Retrieved from http://europa.eu/documents/comm/white_papers/pdf/com95_590_en.pdf

Francis, W. Nelson & Henry Kucera (1982), *Frequency analysis of English usage.* Boston: Houghton Mifflin.

Gardner, Dee. & Mark Davies (2014), "A new academic vocabulary list", *Applied Linguistics,* 35(3):305-327. http://dx.doi.org/10.1093/applin/amt015

Genesee, Fred & Kathryn Lindholm-Leary (2013), "Two case studies of content-based language education", *Journal of Immersion and Content-Based Language Education*, 1(1):3-33. doi: 10.1075/jicb.1.1.02gen

Hyland, Ken & Polly Tse (2007), "Is there an 'academic vocabulary'?", *TESOL Quarterly, 41*(2), 235-253. doi: 10.1002/j.1545-7249.2007.tb00058.x

Hyltenstam, Kenneth (2004), "Engelskan, skolans språkundervisning och svensk språkpolitik" (English, language education at school and Swedish language policy), in Josephson, Olle (ed.), *Engelskan i Sverige. Språkval i utbildning, arbete och kulturliv,* Vol. 89. Stockholm: Norstedts ordbok, 36-110

Jexenflicker, Silvia & Christiane Dalton-Puffer (2010), "The CLIL differential: Comparing the writing of CLIL and non-CLIL students in higher colleges of technology", in Dalton-Puffer, Christiane, Tarja Nikula & Ute Smit (eds.), *Language use and language learning in CLIL classrooms*. Amsterdam: John Benjamins, 169–190.

Krashen, Stephen D. (1982), *Principles and practice in second language acquisition* (1. ed.). Oxford: Pergamon.

Lasagabaster, David & Juan M. Sierra (2009), "Language attitudes in CLIL and traditional EFL classes", *International Journal of CLIL Research,* 1:4–17.

Laufer, Batia & Paul Nation (1995), "Vocabulary size and use: Lexical richness in L2 written production", *Applied Linguistics,* 16(3):307-322. http://dx.doi.org/10.1093/applin/16.3.307

Lo, Yuen Y. & Victoria A. Murphy, V. A. (2010), "Vocabulary knowledge and growth in immersion and regular language-learning programmes in Hong Kong", *Language and Education,* 24(3):215-238. doi: 10.1080/09500780903576125

Long, Michael H. (1996), "The role of the linguistic environment in second language acquisition", in Ritchie, William C. & Tej K. Bahtia (eds.), *Handbook of second language acquisition*. New York: Academic Press, 413–468.

Lyster, Roy. (2007), *Learning and teaching languages through content. A counterbalanced approach* (vol. 18). Amsterdam/Philadelphia: John Benjamins.

Nation, Paul (2001), *Learning vocabulary in another language*. Cambridge: Cambridge University Press.

Nation, Paul (2004), "A study of the most frequent word families in the British National Corpus", in Bogards, Paul & Batia Laufer (eds.), *Vocabulary in second language: Selection, acquisition, and testing*. Philadelphia: John Bejamins.

Norton, Bonny (1997), "Language, identity, and the ownership of English", *TESOL Quarterly,* 31(3):409-429. doi: 10.2307/3587831

Nunan, David (2003), "The impact of English as a global language on educational policies and practices in the Asia-Pacific region", *TESOL Quarterly,* 37(4): 589-613. doi: 10.2307/3588214

Olsson, Eva (2012), *"Everything I read on the Internet is in English:" On the impact of extramural English on Swedish 16-year-old pupils' writing proficiency*. Licentiate thesis. Gothenburg: Gothenburg University.

Pollitt, Alastair (2012), "The method of adaptive comparative judgement", *Assessment in Education: Principles, Policy & Practice,* 19(3):281-300. doi: 10.1080/0969594x.2012.665354

Ruiz de Zarobe, Yolanda (2008), "CLIL and foreign language learning: A longitudinal study in the Basque country", *International CLIL Research Journal,* 1:60–73.

Ruiz de Zarobe, Yolanda (2010), "Written production and CLIL: An empirical study", in Dalton-Puffer, Christiane, Tarja Nikula & Ute Smit (eds.), *Language use and language learning in CLIL classrooms*. Amsterdam: John Benjamins, 191–212.

Rumlich, Dominik (2013), "Students' general English proficiency prior to CLIL", in S. Breidbach, Stephan & Britta Viebrock (eds.), *Content and Language Integrated Learning (CLIL) in Europe. Research perspectives on policy and practice*. Frankfurt: Peter Lang,181–202.

Sundqvist, Pia (2009), *Extramural English matters: out-of-school English and its impact on Swedish ninth graders' oral proficiency and vocabulary*. Doctoral thesis. Karlstad: Karlstad University.

Swain, Merrill (1995), "Three functions of output in second language learning", in Cook, Guy & Barbara Seidhofer (eds.), *Principle and practice in applied linguistics*. Oxford: Oxford University Press, 125–144.

Swain, Merrill (2000), The output hypothesis and beyond: Mediating acquisition through collaborative dialogue. In J. P. Lantolf (Ed.), *Sociocultural theory and second language learning* (pp. 97–114). Oxford: Oxford University Press.

Swain, Merrill & Sharon Lapkin (1982), *Evaluating bilingual education : a Canadian case study*. Clevedon: Multilingual matters.

Sylvén, Liss Kerstin (2004), *Teaching in English or English teaching? On the effects of content and language integrated learning on Swedish learners' incidental vocabulary acquisition*. Doctoral thesis. Gothenburg: Gothenburg University.

Sylvén, Liss Kerstin (2013), "CLIL in Sweden — why does it not work? A metaperspective on CLIL across contexts in Europe", *International Journal of Bilingual Education and Bilingualism,* 16(3):301-320.

Sylvén, Liss Kerstin & Sölve Ohlander (2014), "The CLISS project: Receptive vocabulary in CLIL versus non-CLIL groups", *Moderna Sprak,* 108(2):80-114.

West, Michael (1953), *A general service list of English words with semantic frequencies and a supplementary word-list for the writing of popular science and technology*. London: Longman.

Yoxsimer Paulsrud, BethAnne (2014), *English-Medium instruction in Sweden. Perspectives and practices in two upper secondary schools.* Doctoral thesis. Stockholm: Stockholm University.

Appendices

Appendix A. Average text length in number of words in texts by CLIL and non‑CLIL students

	CLIL		non-CLIL			
	Mean no of words	Stand. dev.	Mean no of words	Stand. dev.	T-value	*p*-value
Text 1	570	207	364	164	6.18	.000*
Text 2	587	239	348	147	6.13	.000*
Text 3	605	215	395	155	6.23	.000*
Text 4	525	193	379	161	3.92	.000*

* statistically significant difference between CLIL/non-CLIL

Appendix B. Kim's first essay (Assignment 1)

I am strongly against nuclear power plants, mostly because of all the polution that is caused by it but also the fact that terrorist groups now are so more advanced than they were 10 years ago and can now much easier gain access to uranium for nuclear weapons instead of using it for electricity which is it's main perpose.

A "chain reaction" is caused when uranium is used and this is very dangerous even thought the company might have protection for it. You see the "chain reaction" is the result of nuclear fission which is caused by small neutrons smashing into eachothers, this may not be a very good use for a energy source when you know that the things keeping this "chain reaction" under control, also are aquiered a big amount of energy. So for example, maybe one day we don't have that much energy left to provide the coolers for the uranium chain reaction and clearly not enough to provide the city, then we have a major problem. This is a problem only possible for nuclear energy, oil and coal and can be prevented if you use for example solar energy.

I also think that uranium should be forbidden because of all the using of it in wars causing the opposing country in the war suffer the results of nuclear attacks that the specific area may not be able to recover from because of all the radiation and destruction of the area.

Uranium is also a metal stored in the ground for thousands of years and will not reacure in a couple of years, so the amount of nuclear power is only available for maybe another century depending on if they find more resources for it while energy sources like solar energy which is a steady source because the fact that the sun won't be destroyed for another ten thousand years or so.

The power we gain from wind is also a reliable source bacause of the fact that the world always keeps turning and the sea´s still are on the earth making the temperatures rise and fall to make storms, rainy days and perhaps even hurricanes. Another thing about the uranium is that the sources for it may not always be in a wasteland, it can be under a village, a small town or if the odds are bad it might even be under a big city.

And if the amount of the uranium is getting low we would have to remove these

towns, villages and citys just to keep the energy going, but really, what is the point in having electricity if you have to relocate your home all the time?

I say, quit uranium, start using solar, geothermal and wind to make your home feel the energy.

Appendix C. Kim's last essay (Assignment 4)
Biodiversity
Water, energy and food are all something that the human race needs in order to stay alive. According to Professor John Beddington at Imperial College, we are soon to hit a point in which our supply of these are going to be of shortage. Professor Beddington claims that the world in the year of 2030 will have much bigger demands on water, energy and food because of our increased population.

We humans have already with all the other animals adopted to our surroundings in order live more efficiently. This is something that has taken a lot of time and is a part of nature that the world very badly need. We are at the moment able to transport many kind of products across the world very easy and fluidly, this is something that humanity needs. We in Europe need many of the products produced in Asia and Africa, the same goes for the other continents, and because the people over in Asia and Africa already has adopted so greatly to their certain weather conditions, they are the perfect candidates for harvesting food within their own areas. Because they are over there while the Europeans still work over in Europe, more food can be provided for all of us if we are all capable of working together. This idea requires that the country in charge of harvesting food, have to have loyal transporters and loyal buyers in order to make sure it all runs smoothly.

Another reason for why we need the biodiversity is because they make the ecological system go round. If we were to not have biodiversity we would be stuck with maybe just one bird and one kind of fish. Because we have such a diversity we have lots of fishes, and lots of birds that all carry different DNA. That means that they are all different which is very helpful for us humans because we gain strength both mentally and physically from eating very different types of food. Animals also have different ways of living, some of which we humans can observe and take advantage from and some that we can not.

Different cultural diversities can also be a very important factor in the way we look at humanity. We all think alike in some way or another and because of this we are all more or less alike. This is why I think we need to stop comparing each other and start connecting people all over the world in order to increase our chances of finding something that we are looking for. This could be anything, but at the moment we just need as much help for all projects available as possible. This could also decrease the racism for future generations that could lead to more peace in the world. This could end up in several various ways where we might find war, truce or peace. Either way the human race would profit from it.

In conclusion I would like to say that I think that biodiversity is something very crucial for the survival of the human race and that we would perish without it. It helps us to be better individuals and provides better lives for the future human

generation. We should be very great full for that it exists and I also think that it is something that we need to more deeply study.

Appendix D. Average proportion of AWL types in texts by CLIL and non-CLIL students

Coxhead (2000) AWL	CLIL		non-CLIL			
	Mean %	Standard deviation	Mean %	Standard deviation	T-value	*p*-value
Text 1 AWL types %	6.53	2.65	5.45	2.83	2.30	.023*
Text 2 AWL types %	6.40	2.80	5.04	2.03	2.89	.005*
Text 3 AWL types %	5.56	2.45	3.74	1.82	4.72	.000*
Text 4 AWL types %	8.15	3.12	6.60	2.85	2.51	.014*

* statistically significant difference between CLIL/non-CLIL

Appendix E. Average proportion of AVL types in texts by CLIL and non-CLIL students

Gardner and Davies (2014) AVL	CLIL		non-CLIL			
	Mean %	Standard deviation	Mean %	Standard deviation	T-value	*p*-value
Text 1 AVL types %	12.62	4.01	9.75	3.45	4.35	.000*
Text 2 AVL types %	15.55	4.73	12.91	4.99	2.96	.004*
Text 3 AVL types %	13.40	4.51	9.94	3.22	4.91	.000*
Text 4 AVL types %	19.30	5.29	16.12	5.01	3.02	.003*

* statistically significant difference between CLIL/non-CLIL

Appendix F. Average % of academic tokens (AWL/AVL) in texts judged as weaker, at the same level or as stronger than the text of comparison

	Texts judged as *weaker* than the text of comparison			Texts judged *at the same level* as the text of comparison			Texts judged as *stronger* than the text of comparison		
	Mean %	Stand dev	N	Mean %	Stand dev.	N	Mean %	Stand dev	N
Text 4 AWL %	3.30	2.44	6	4.18	1.78	13	5.63	2.13	11
Text 4 AVL %	10.57	3.37	6	11.05	3.53	13	13.66	3.67	11