Vocabulary learning at primary school:
comparison of EFL and CLIL

Elsa Tragant[a], Anna Marsol[a], Raquel Serrano[a] and Àngels Llanes[b]

[a]Dept. Filologia Anglesa i Alemanya, Universitat de Barcelona, Barcelona, Spain.

[b]Dept. d'Anglès i Lingüística, Universitat de Lleida, Lleida, Spain

Abstract

Comparative studies on CLIL (Content Language Integrated Learning) often show an advantage in favor of CLIL vs. non-CLIL students. However, oftentimes studies are difficult to interpret because the hours of instruction between the groups under comparison tend to be quite different. This study focuses on a single group of students (n=22) aged 8 who were exposed to EFL instruction in the Fall term and to CLIL instruction (Science) in the Winter term. The major objectives are to analyse the vocabulary in the materials that were used in class and to examine gains in productive lexical knowledge. Results show that students were exposed to a greater amount of words and to more abstract and technical vocabulary in the CLIL materials and that they made significant progress in vocabulary learning in both contexts. The study also reveals that learning English through Science proved to be a more challenging experience than learning the foreign language in the English class.

Vocabulary is acknowledged to be an important building block in the learning of a language and it has also been found to be closely related to language proficiency. In

vocabulary learning is crucial since vocabulary is traditionally one of the main objectives for teachers and materials developers at the early phases of second language acquisition. The present is a small-scale study with a focus on vocabulary learning by a group of primary school children who learnt English in a traditional EFL class and in a Science class through Content and Language Integrated Learning (CLIL). The study first offers an analysis of the EFL and CLIL materials that were used in the two classes to later on present a comparison of the vocabulary gains based on those materials.

Background

Before 1980, European schools that offered the possibility of learning content through a second/foreign language were limited to a minority of elite-institutions, but since the emergence of the concept of CLIL in the 1990's (Marsch, 2002), this mode of instruction has not stopped growing. In 2003, the Commission of European Communities identified it as a priority area in secondary education, and in 2006 an in-depth report was published by the European Commission about how CLIL was taking place in primary and secondary schools (Eurodyice 2006). At present the European Union continues giving support to CLIL projects through the Lifelong Learning Programme.

Vocabulary and CLIL materials

In spite of this spread of CLIL in primary and secondary education in Europe, for many years there was a dearth of specifically designed materials and textbooks (Varkuti, 2010; Wegner, 2012). Teachers had to either create their own or adapt materials designed for native speaking students. This was an undesirable situation given that ideally, CLIL materials should include specific features such as a wealth of layouts

(continuous text, tables and diagrams, bullet points, visuals) and linguistic adaptations, following Coyle, Hood and March's (2010) recommendations.

Over the past few years the picture has somewhat changed. Nowadays some major ELT (English Language Teaching) publishers have their own textbooks for a variety of subjects, both for primary and secondary education. Some of them have also produced supplementary materials (i.e., dictionaries, readers, teacher resources). Because of their novelty, these CLIL materials largely remain to be evaluated. In a study of the CLIL activities incorporated in ELT coursebooks, Banegas (2014) found that there was a simplification of content and a dominance of reading activities and lower-level thinking tasks, often related to vocabulary activities. In a comparison of an EFL (English as a Foreign Language) and a CLIL coursebook, Schenk (2013) found that there was more vocabulary repetition in the CLIL textbook under analysis, although the proportion of words in the first frequency band was similar in both coursebooks.

With the objective of further contributing to this line of research, this study will focus on the examination of vocabulary in CLIL in comparison to EFL textbooks. In the past, the evaluation of vocabulary in ELT materials has been approached in three ways: studies that deal with the role   vocabulary plays in the syllabus (e.g., O'Dell 1997), studies that analyse which aspects of vocabulary knowledge receive attention in textbooks (e.g., Brown, 2010) and studies that look into the amount of vocabulary in textbooks (e.g., Jimémez Catalán and Mancebo Francisco, 2008). The present study is going to take this last perspective by describing and quantifying the amount of vocabulary students are exposed to through a CLIL and EFL coursebook targeted to the same student population.

Vocabulary learning and CLIL

Many are the benefits reported for second language learners that participate in CLIL programmes. An e-newsletter from the European Commission (2013) lists up to 10 advantages ranging from attitudinal benefits, exposure to different classroom practices and perspectives, as well as gains in intercultural knowledge and communication, in addition to improving language competence and oral communication skills. According to Dalton-Puffer (2009), CLIL is especially beneficial for the development of linguistic competence, in particular lexical development, since students' L2 production abilities are likely to be stretched to a larger extent than in a regular EFL class.

A number of empirical studies from different research traditions confirm the lexical gains from CLIL school-aged students in Europe. There are even some studies that provide evidence for children's significant development of receptive vocabulary skills in the preschool years (Buyl and Housen, 2014). From a discourse analytic perspective, Dalton-Puffer (2007) has shown how CLIL students often notice lexical gaps in their own oral output and are able to self-initiate repair, something that is less frequent in typical EFL classrooms. Ruiz de Zarobe's (2010) work on written production indicates that after the second year, a group of secondary school students enrolled in an intensive CLIL programme obtained better vocabulary scores than students in the non-CLIL and the less intensive CLIL programmes. Quantitative analyses of written production in upper secondary education (Jexenflicker and Dalton-Puffer, 2010) also show that CLIL students use less frequent words than the non-CLIL students and they tend to show more lexical variation. Lexical variation in essays was also found to be higher among primary school CLIL students in Agustín Llach and Jiménez Catalán (2007). A different approach to vocabulary research in CLIL is the use of discrete-point vocabulary tests. Within this framework, Jiménez Catalán and Ruiz de Zarobe (2009) found that the difference in favour of the CLIL group was more evident

in the 2000 most frequent word test than in the 1000 word test in students aged 11-12. Sylvén (2004) and Seregély (2008) also found that CLIL learners outperformed traditional learners in a battery of vocabulary tests. However, and in spite of these positive findings, evidence about the CLIL advantage has generally been more conclusive in regards to receptive rather than productive knowledge and skills (Dalton-Puffer, 2007; Fusté and Miralpeix, 2013; Heras and Lagasabaster, 2014). The present study will focus precisely on the examination of the productive knowledge students gain when learning vocabulary through CLIL instruction.

The study has been designed to avoid one of the most salient problems in existing research, which is about the comparability of the CLIL and the non-CLIL groups, especially when the number of schools involved are small (see Aguilar and Muñoz, 2013; Bruton, 2011 for a discussion of other problematic aspects of CLIL research). This is the case in studies where the two groups come from different cities and /or different school systems as it occurs in, for example, Moghadam and Fatemipour (2014). The limitation sometimes comes from the fact that the CLIL groups have received considerably more hours of instruction, which is often an unavoidable confound variable, as in Jiménez Catalán and Ruiz de Zarobe (2009). Another factor that also makes comparisons less valid is when CLIL students are selected or self-selected. In fact, this was the case in Heras and Lagasabaster's work (2014), in a school where entering the CLIL program was voluntary. The lack of group comparability is in general an important challenge in L2 classroom research, but it is particularly difficult and sometimes impossible to avoid in the context of CLIL instruction. Aware of these challenges, we had the opportunity to conduct a study on vocabulary gains where the issue of comparability is not problematic because the same group of students first underwent a period of EFL instruction followed by a period of CLIL instruction.

Aims of the study and research questions

The present study focuses on a group of primary school students who followed a program that concentrated their regular English instruction time in the Fall term and their CLIL instruction time in the Winter term. The main objectives of the study are to analyse the vocabulary profile of the instructional materials used in the two instructional contexts and to compare the vocabulary learning resulting from the two instructional periods. More specifically, we aim to answer the following questions:

(1) How much and what type of target vocabulary are students exposed to from their class books after one term of EFL instruction and another term of CLIL instruction?

(2) Do students experience any gains in their productive knowledge of the target vocabulary from the beginning to the end of the two terms?

(3) Do students make more vocabulary gains in one of the two programs: EFL vs CLIL?

In the three research questions above, we will focus on the key vocabulary that appeared in the teacher's manuals of the coursebooks under analysis. This vocabulary will be referred to as 'target vocabulary' in this study, which includes the words that students are expected to learn through direct instruction.

Methodology and context

The school and participants

The school is a prestigious boys' school located in a residential area in the vicinity of a middle sized city in Catalonia. It is a small school (only one class per grade) with spacious facilities and one of its distinctive features is the importance attached to

English by its administrators. An illustration of this is that in grade 3, which is the target grade in this study, four hours a week are usually devoted to English instruction (the official requirement in Catalonia is two hours/week). In addition, three hours a week are normally devoted to Science in English at this grade level, which is when CLIL instruction is introduced in this school. For the purpose of this study, however, the school was requested to modify this scheme and concentrate the English language instruction into the Fall term (seven 55' periods/week) and the instruction of Science in English into the Winter term (seven hours/week). This scheme would allow us to examine vocabulary learning through EFL and CLIL instruction independently. In both contexts students had a class book and an activity book, and lessons took place in their regular classroom except for one lesson a week when the class was held in a computer room.

Participants in this study are 22 boys aged 8-9 (grade 3) and their English/Science teacher. One student was excluded from the study because he had serious learning difficulties and, when data was analysed, he was an outlier. Except for one father and two mothers, the rest of the parents have university degrees and ten of the children take English as an after-school activity. Even though almost all of the children (20 out of 22) report English to be an important language to learn, when asked about their three favourite subjects, English is only mentioned by three of them. These children are taught English and Science in English by the same teacher, an experienced male teacher with a good command of English and effective management skills. He is a certified English primary school teacher who also has a BA degree in English Studies.

Instructional materials and practices

The EFL lessons were based on a class and activity book called *Incredible English Kit 3* (Phillips and Morgan 2007). The class book includes a variety of materials: speaking

and reading/listening practice, songs and vocabulary, pronunciation and covert grammar activities. Vocabulary in this class book is formally introduced in the first page of each unit along with a large drawing, but it is not typologically enhanced in any way in the remaining pages of the unit. In the vocabulary activities of the activity book, words are provided against a grey background.

Most EFL sessions started as a whole class with the class book, while the last part of the class was often devoted to individual work with the activity book. One session a week took place in the computer room, where students were asked to prepare PowerPoint presentations to illustrate the vocabulary of the unit they were studying at that moment. The teacher covered most of the materials from the class book and students were usually asked to complete all the activities in the activity book. At times the teacher was also observed to provide unplanned explanations outside the textbook (i.e., similarities between L1 and L2 words, crosscultural differences). During these lessons the teacher practically always addressed the students in English and, when he asked them questions, he expected them to answer in English as well.

The Science lessons were based on a class and activity book called *Top Science 3* (Santillana, 2011). In the class book, information is presented through texts and numerous photographs and drawings, which are complemented with questions, hands on activities and written exercises. In this class book, target words are highlighted in bold type in the texts where content is introduced, and these texts are often accompanied by numerous pictures and drawings. These visuals come with arrows and labels and/or a short descriptive caption. In the activity book, words are individually highlighted in a yellow background and a black frame.

In class, most Science sessions started with a teacher presentation of content followed by practice activities that were usually done as a whole class. In the same way

as the EFL lessons, one session a week took place in the computer room and it was also devoted to prepare Powerpoint presentations on vocabulary. The teacher made a selective use of the class book since he thought some of the texts were too difficult for the students, but all the activities from the workbook were completed and the target vocabulary from each unit covered. At times students were observed to engage in spontaneous discussions or to self-initiate questions to the teacher in relation to the content of the lesson (i.e., a discussion in L1 about whether plants can move, a student's question in L1 about why salamanders do not die when they lose their tail). However, these occasions were not always picked up by the teacher. During these lessons, the teacher practically always used English but, in contrast to the EFL lessons, he sometimes accepted students using Spanish or Catalan to respond to some of his questions.

Students' self-reported evaluation of English and Science lessons were quite similar, with three fourths of the class saying that they liked them or liked them a lot and only a minority (1 or 2 students) saying they did not like them much or did not like them at all. Similarly, about the same number of students (about half the class) thought English and Science lessons were neither difficult nor easy subjects, although there were a few more students who found the EFL class easier or much easier than the Science class (26% vs 17.4% respectively).

Instruments and analysis

Textbook analysis and vocabulary tests constitute the two primary instruments in this study. Secondary sources include six classroom observations, two teacher interviews and a questionnaire with bio and attitudinal data that students completed at the beginning and end of the school year. The vocabulary analysis from the two textbooks focused on the target words of the units that were covered during the instructional

period under analysis and it was carried out with Lextutor. This online program is based on the Lexical Frequency Profiler (Cobb, 2014; Heatley, Nation and Coxhead 2002) and divides the words of texts into first and second thousand levels, academic words and the remainder or 'off-list' words.[i]

The vocabulary that was covered in the English class during the Fall term was tested at the beginning and end of that term (September-Time 1 and December-Time 2) while the vocabulary covered in the Science class during the Winter term was tested at the beginning and end of that term (December-Time 1 and March-Time 2). The test in December included items from the English and Science classes presented at random. In each vocabulary test, pupils' knowledge of30 items was assessed. Tests included small drawings, next to which students were asked to write a word so productive vocabulary knowledge was assessed. The initial of each word was provided as a prompt. A receptive vocabulary test was also administered but results proved to be unreliable and consequently this test will not used in this study. See footnote [ii] for an explanation. Because of the format of the productive vocabulary test, only concrete nouns were included while the receptive vocabulary test included adjectives and verbs as well as abstract and concrete nouns. Approximately the same number of items from each teaching unit was assessed between the productive and receptive tests. These items were chosen from the vocabulary component of the syllabus in each unit of the two textbooks. In the EFL test, we included the words that were highlighted on the vocabulary list. The CLIL vocabulary list included more items than the EFL list and they were not highlighted. In this case, we included vocabulary items that were easily depicted (and therefore interpreted). See appendix A for the list of words included in the EFL and Science tests, which were the same words for the two testing times (pre- and post tests). The tests were tailored-made for this study since no standardized tests could

be used, but they were previously piloted with children the same age to make sure that the pictures were clear and that the length of the test was adequate.

In scoring the tests, correctly produced items were awarded one point, even if words were misspelled. In analyzing them, related and independent samples t-tests were carried out. The level of significance level was set at .05.

Results

Textbook analysis: vocabulary profile

During the Fall term, the first six units from the EFL class and activity books (*Incredible English 3*) were covered in class including the following vocabulary areas: personal possessions, clothes, food, places in a town, animals, and adjectives to describe people and music. During the Winter term, five units were covered in class from the CLIL class and activity books (*Top Science 3*) including the following topics: the body, the senses, living things, vertebrate animals and invertebrate animals. The present textbook analysis focuses on the vocabulary that was encountered in these units, since this is going to be the pool of vocabulary that students will later be tested on to answer research questions two and three.

A preliminary examination of the wordlists from each class book shows some differences. The average number of lexical items per unit is higher in the Science class book than in the EFL class book (77 vs. 49 words respectively), but the proportion of multi-word lexical items are considerably lower in the Science than the EFL wordlists (11% vs. 27% respectively). The most frequent multiword units in the EFL class book are fixed expressions, that is utterances to be learnt as a whole (i.e, *Here you are, What's the matter?, what a mess!*), and productive expressions (e.g. *I can't find...., Here's....).* The most common types of multiword units in the Science class book are

phrasal verbs (e.g. *to breathe in, to cut off*) and nouns preceded by modifiers (e.g. *life cycle, magnifying glass*).

A further analysis of the two wordlists (including chunks but not productive expressions) by word frequency (tokens) with Lextutor (Cobb, 2014) was performed. This analysis is based on the total number of target words in the first six units in the EFL class book and the first five in the Science class book, which is considerably higher in the latter (a total of 256 vs 337 words respectively). Table 1 shows the proportions of the first and second thousand level words as well as the academic and off-list words in both class books, which are somewhat different but not widely different. Word types that are part of the first and second thousand levels are slightly higher in the EFL class book, while the proportion of academic and off-list words are a little higher in the Science textbook. The same trend applies for Latin cognates, with the Science class book showing slightly higher indexes.

Insert table 1 about here

A closer look at the actual words included in the off-list words reveals more relevant differences. For example, while the off-list words the EFL class book include a good proportion of lexis related to food (i.e., *asparagus, biscuits, broccoli*), clothing (i.e., *tights, trousers, gloves*) and everyday objects (i.e., *magazine, jar, torch*), the Science class book list includes a good proportion of classifying nouns and adjectives (i.e., *carnivore, canine, auditory*), words designating processes or abstract concepts (i.e., *germinate, height, vibration*) and nouns related to human and animal body parts (i.e., *iris, sternum, womb*), some of which are quite technical.

Finally, the words included in the tests that were used to measure vocabulary gains in the next section reflected the vocabulary profiles of the EFL and CLIL class books. The EFL test contained more words belonging to the first and second frequency bands (27% and 36%) than the CLIL test (23% and 27%). In contrast, the proportion of 'off-list' words in the CLIL test was higher (50%) than in the EFL test (36%).

Vocabulary gains

Gains in students' productive knowledge of EFL words from Time 1 (beginning of Fall term) to Time 2 (end of Fall term) will be examined first. The mean score at Time 1 was 8.9 for the whole sample (n=22) and it increased to 20.1 at Time 2. The related samples t-test indicates that the difference between these two scores is significant with a large effect size, according to Cohen's *d* (see Table 2). When we repeated this same test with the subsample of students not taking extra-school English lessons (n=11), results were equally significant. It should be noted that in both cases the standard deviations are quite high, showing considerable variability among students.

Insert table 2 about here

Gains in students' productive knowledge of Science words from Time 1 (beginning of Winter term) to Time 2 (end of Winter term) will be examined next. The mean score at Time 1 was 11.8 for the whole sample (n=23) and it increased to 20 at Time 2. The related samples t-test indicates that the difference between these two scores is significant with a large effect size (see Table 3). When we repeated this same test with the subsample of students not taking extra-school English lessons (n=11), results were equally significant.

Insert table 3 about here

Finally, a comparison of gains between the EFL and CLIL vocabulary tests will be presented. In order to draw such a comparisons, gains were obtained for each student by subtracting their scores at Times 1 from those obtained at Time 2. The results from an independent samples t-test show that students' mean gains are significantly higher in the EFL context and the magnitude of the difference was large (see Table 4). Whereas students in the EFL tests were able to produce an average of 11 more words at Time 2 than at Time 1, students in the CLIL test were able to produce an average of 8 more words at Time 2 than at Time 1.

Insert table 4 about here

Discussion and conclusions

The comparison of the target vocabulary students were exposed to in the EFL and CLIL contexts showed different degrees of difference depending on the analysis. The frequency analysis of the target vocabulary revealed slight differences between the EFL and Science materials, with somewhat higher proportions in the first and second frequency word bands in the former and slightly higher proportions of Latin-based words and off-list in the latter. We would expect that these differences would become more marked in higher level materials as the subject becomes more specialized and topics are dealt with in more depth.

Nevertheless, further analysis revealed clearer differences. In terms of quantity, the CLIL materials included about one fourth more items than the EFL materials, which

probably reflects the greater role of content and information in the former. In terms of

types of words, the differences that were identified are probably a reflection of the text

types included in the EFL and CLIL materials. The common presence of stories, songs

and dialogues in the EFL materials would explain why multi-word units often included

chunks and productive language (more typical of oral language), while the

predominance of expository texts in the CLIL materials would explain the presence of

nominal premodification and abstract language (more typical of descriptive language

and the scientific genre). In terms of the visual presentation of vocabulary, target words

in the CLIL materials were more often and more prominently highlighted than in the

EFL materials. An analysis of the words used in the vocabulary tests that the students

took confirmed this difference between CLIL and EFL vocabulary.

The examination of the gains students experienced from the beginning to the end

of the two terms confirmed that there was a considerable growth in the productive

vocabulary of the learners in both contexts; thus, confirming the general positive

benefits of CLIL instruction on vocabulary learning (Buyl and Housen, 2014; Jiménez

Catalán and Ruiz de Zarobe, 2009; Seregély, 2008; Sylvén, 2004). However, the

comparison of these gains between the two contexts showed that on average, students

learned a few more words from the EFL lessons than from the CLIL lessons and that

this difference was significant. The fact that students in the CLIL lessons used their L1

more often than in the EFL lessons to answer the teacher's questions could partially

explain these results. Also, the fact that the vocabulary in the EFL class book included

more frequent words probably more relevant to students' daily lives, may be another

explanatory factor. In fact, it is likely that the vocabulary in the EFL textbook could

more easily become part of the already existing every-day word network of the child's

mental lexicon. In contrast, part of the vocabulary in the CLIL textbook may not yet

have an appropriate web of words in the child's mental lexicon to easily connect to. Even though receptive knowledge could not be tested in this study (see footnote ii), and it is one of its limitations, it is possible that the impact of CLIL would have been clearer on receptive than productive vocabulary, as evidenced in past research (Ruiz de Zarobe, 2011). Additional limitations that need to be taken into account in interpreting the results are that the study involved a small sample of language learners and a few units from the textbook. The fact that the vocabulary test was restricted to concrete words made the test less representative of the vocabulary students were exposed to in their class books, this being especially detrimental in the case of CLIL. Another limitation is the fact that important factors that affect vocabulary learning were not taken into account in the analysis of the data such as word length, frequency of use, textual context, among others, etc.

Beyond the small-scale results of this study, we feel some reflections and implications for further research can be drawn from our work. First, the research design that we have employed, where the hours of instruction in CLIL and EFL remain constant, has strengthened the validity of the research. The design is also methodologically innovative, at least in CLIL comparative studies, and it has proved to be feasible in practice. Secondly, in the process of designing the vocabulary tests, we have faced major difficulties, to the point of having to discard the receptive vocabulary test. Further research should give especial attention to how to assess CLIL vocabulary with young children. The fact that some of this vocabulary typically includes abstract and technical terms is an added difficulty when it comes to assessment with this age group. Finally, some of aspects of classroom practice that are briefly described to contextualize the results of this study have revealed some intriguing issues. For example, our observation data suggests some differences in classroom interaction

between the CLIL and EFL sessions, more specifically the fact that there was more teacher-fronted teaching during the CLIL than the EFL sessions. There also seemed to be more student-initiated moves in the CLIL sessions, in spite of them being in the students' L1. Another interesting issue relates to the difficulty of the EFL and CLIL textbooks which, according to the teacher, was not comparable. A later analysis of a random unit from the two textbooks confirmed the teacher's opinion: with an average of 9 and 5.1 words per sentence in the CLIL and EFL texts and an average grade level in terms of comprehension difficulty of 4.2 and 0.6 respectively[iii]. Another issue worth exploring is the profile of the teacher and how this affects classroom practice. In our data the teacher was a language teacher not a science teacher and this could have influenced the fact that he was observed to provide more unplanned explanations during the EFL than the CLIL lessons. He also had students spend time during the CLIL sessions on vocabulary practice (powerpoint presentations), an activity which was also carried out during the EFL sessions and is more characteristic of a language than a science class. The significance of further exploring the above stated aspects of classroom practice could be of especial interest for textbook development and teacher training.

Taken together, there seems to be enough evidence from the different sources of data in this study to indicate that learning science in English was a more difficult experience for our eight-year-old learners than learning the language in the English class. The CLIL class book was considerable more difficult than the EFL class book and the teacher was more permissive with the use of L1 when teaching Science. There were also more target words to be learned from the CLIL units and the analysis with Lextutor found there were slightly more academic, Latin cognate and off-list words in the Science vocabulary. The examination of the off-list words also revealed that the Science

class book included words designating processes and abstract nouns in contrast to the more every-day topics covered in the EFL class book. The complexity of the science lessons was also reflected in students' self-reports, with a few more students finding the EFL class easier or much easier and none finding Science to be easier. Given this evidence, it would be simplistic to interpret our analysis of word gains in isolation without acknowledging that CLIL and EFL lessons offered quite different contexts of learning. Besides, it must be taken into account that vocabulary learning in the context CLIL is a more complex process, since it is not limited to learning new words but also implies understanding new concepts to perform tasks with different cognitive demands.

In sum, the present study has examined the CLIL and EFL approaches with young learners of English following a design that guarantees the comparability of groups under analysis. The analysis of the vocabulary from the class books students used revealed relevant differences, indicating that CLIL and EFL materials offer complementary sources of input. The results have also shown that both approaches offer valid contexts for the development of productive vocabulary knowledge and that Science in English was a more challenging learning experience for the children.

Reference List

Aguilar, A., and C. Muñoz. 2014. "The Effect of Proficiency on CLIL Benefits in

Engineering Students in Spain." *International Journal of Applied Linguistics* 24(1): 1-18.

Agustín Llach, M. P., and R. M. Jiménez Catalán. 2007. "Lexical Reiteration in EFL Young Learners' Essays. Effect of the Type of Instruction." *International Journal of English Studies* 17(2): 85-103.

Banegas, D. L. 2014. "An Investigation into CLIL-related Sections of EFL Coursebooks: Issues of CLIL Inclusion in the Publishing Market." *International Journal of Bilingual Education and Bilingualism*, 17(3): 345-359

Brown, D. 2010. "What Aspects of Vocabulary Knowledge Do Textbooks Give Attention to?" *Language Teaching Research* 15(1): 83–97.

Bruton, A. 2011. "Is CLIL so Beneficial, or Just Selective? Re-evaluating Some of the Research." *System* 39:523-32.

Buyl, A., and A. Housen, A. 2014. "Factors, processes and outcomes of early immersion education in the Francophone Community in Belgium." *International Journal of Bilingual Education and Bilingualism* 17(2): 178-196.

Cobb,T. *Web Vocabprofile.* Accessed 29/05/2014. http://www.lextutor.ca/vp/.

Coyle, D., P. Hood, and D. Marsh. 2010. *CLIL Content and Language Integrated Learning*. Cambridge: Cambridge University Press.

Dalton-Puffer, C. 2007. *Discourse in CLIL Classrooms*. Amsterdam: Benjamins.

Dalton-Puffer, C. 2009. "Communicative Competence and the CLIL Lesson." In *Content and Language Integrated Learning: Evidence from Research in Europe,* edited by Yolanda. Ruiz de Zarobe and Rosa María Jiménez Catalán, 197-214. Clevedon: Multilingual Matters.

Eurodyce. 2006*. Content and Language Integrated Learning (CLIL) at School in Europe*. Brussels: European Commission.

European Comission Communication. 2003. *Promoting Language Learning and Linguistic Diversity: An Action Plan 2004-2006*. Accessed 31 December 2014. http://www.saaic.sk/eu-label/doc/2004-06_en.pdf

European Commission (Languages). 2013. *Content and Language Integrated Learning.* Accessed 15 August 2013. http://ec.europa.eu/languages/language-teaching/content-and-language-integrated-learning_en.htm.

Fusté, J. and I. Miralpeix. 2013. "Assessing the Vocabulary Knowledge of Young EFL Learners in Instructional Settings." Paper presented at the 37th AEDEAN Conference, Oviedo, Nov.13t-15.

Heatley, A., I.S.P. Nation, and A. Coxhead. 2002. RANGE and FREQUENCY programs. Available at http://www.victoria.ac.nz/lals/staff/paul-nation.

Heras, A. and D. Lagasabaster. 2014. "The Impact of CLIL on Affective Factors and Vocabulary Learning." *Language Teaching Research.* doi:10.1177/1362168814541736

Jexenflicker, S. and C. Dalton-Puffer. 2010. 'The CLIL Diferential. Comparing the Writing of CLIL and non-CLIL Students in Higher Colleges of Technology'. In *Language Use and Language Learning in CLIL Classrooms*, edited by Christinae Dalton-Puffer, Tarja Nikula and Ute Smit, 169-189. Amsterdam: John Benjamins.

Jimémenz Catalán, R. M., and R. Mancebo Francisco. 2008. "Vocabulary Input in EFL Textbooks." *RESLA* 21: 147-165.

Jiménez Catalán, R. M. and Y. Ruiz de Zarobe. 2009. 'The Receptive Vocabulary of EFL Learners in two Instructional Contexts: CLIL vs. non-CLIL instruction.' In *Content and Language Integrated Learning. Evidence from Research in Europe* edited by Yolanda Ruiz de Zarobe and Rosa María Jiménez Catalán, 81-92. Clevedon: Multilingual Matters.

Marsh, D. (editor), 2002. CLIL/EMILE. The European Dimension. Actions, Trends, and Foresight Potential. Jyväskylä: University of Jyväskylä.

Moghadam, N. Z., and H. Fatemipour. 2014. "The Effect of CLIL on Vocabulary Development by Iranian Secondary School EFL Learners." *Procedia - Social and Behavioral Sciences* 98, 2004 – 2009.

Schenk, J. 2013. "A lexical Approach to Coursebook Analysis: How an EFL Coursebook and a CLIL Coursebook Approach Vocabulary Acquisition." Unpublished Master's thesis, Universitat de Barcelona.

Meara, P. 1992. *EFL Vocabulary Tests*. Swansea: Centre for Applied Language Studies.

O'Dell, F. 1997. "Incorporating Vocabulary into the Syllabus." In *Vocabulary and Language Teaching*, edited by Norbert Schmitt and Michael McCarthy, 258–78. Cambridge: Cambridge University Press.

Phillips, S. and M. Morgan. 2007. *Incredible English Kit 3*. Classbook and Activity book. Oxford: Oxford University Press.

Ruiz de Zarobe, Y. 2010. "Written Production and CLIL: An Empirical Study." In *Language use and language learning in CLIL classrooms,* edited by Christiane Dalton-Puffer, Tarja Nikula and Ute Smit, 191-209 . Amsterdam: John Benjamins.

Ruiz de Zarobe, Y. 2011. "Which Language Competences Benefit from CLIL? An Insight into Applied Linguistics Research." In *Content and Foreign Language Integrated Learning: Contributions to Multilingualism in European Contexts*, edited by Yolanda Ruiz de Zarobe, Juan Manuel Sierra and Francisco Gallardo del Puerto, 129-154. Bern: Peter Lang.

Santillana. 2011. *Top Science 3*: Textbook and activity book. Madrid: Santillana/ Richmond Publishing.

Seregëly, A. M. 2008. "A Comparison of Lexical Learning in CLIL and Traditional

EFL classrooms." Unpublished MA thesis, Universität Wien.

Sylvén, L.K. 2004. Teaching in English or English teaching? On the Effects of Content

and Language Integrated Learning on Swedish Learners' Incidental Vocabulary

Acquisition. Göteborg: Department of English, Göteborg University.

Varkuti, A. 2010. "Linguistic Benefits of the CLIL Approach: Measuring Linguistic

Competences." *International CLIL Research Journal*, 1(3): 67-79.

Wegner, A. 2012. "Seeing the bigger picture: What students and teachers think about

CLIL." *International CLIL Research Journal*, 1(4): 29-35.

Appendix A: Items included in vocabulary tests

Test items from the EFL textbook:

biscuits, bridge, butterfly, cap, cheese, coat, comb, dress, eggs, flower, glasses, gloves, grasshopper, guinea pig, ladybird, leaf, library, pillow, rabbit, river, scarf, sleeping bag, snake, strawberries, swimming pool, toothbrush, torch, towel, trainers, wood.

Test items from the CLIL textbook:

ant, bat, bear, beetle, brain, branch(es), cow, crab, dolphin, ear, eye, eyebrow, eyelash(es), fly, horse, jellyfish, lamb, mussel, nose, root, seed, skull, spider, squirrel, stems, sunflower, tear, tongue, whale, wings.

		EFL class book (n=256 tokens)	Science class book (n=337 tokens)
Frequency	1k words	39.22%	33.51%
	2k words	27.83%	21.73%
	Off-list words	31.76%	40.31%
	Academic words	1.18%	4.45%
Latin cognate index		33.91%	39.47%

Table 1. Frequency analysis of EFL and Science class books

	Time 1 mean (SD)	Time 2 mean (SD)	*t*	*p*	*d*
Whole sample (n=22)	8.9 (5.6)	20.1 (6.7)	10.4	.000	.8
Subsample (extracurricular ss excluded) (n=11)	8.1 (6.3)	20.9 (5.8)	8.7	.000	.9

Table 2. Gains in EFL vocabulary test (Fall term)

	Time 1 means (SD)	Time 2 means (SD)	*t*	*p*	*d*
Whole sample (n=23)	11.8 (4.5)	20 (4.5)	13.1	.000	.9
Subsample (extracurricular ss excluded) (n=11)	10.5 (4.6)	19.4 (4.3)	8.9	.000	.9

Table 3. Gains in CLIL vocabulary tests (Winter term)

	Mean gains (SD)	*n*	*t*	*p*	*d*
EFL Context	11.6 (4.6)				
CLIL Context	8.1 (2.9)	21	4.4	.000	.5

Table 4. Gains in EFL and CLIL tests

---

[i] 0ff-list' words are those that belong to the third frequency band and beyond and are not considered academic words.

[ii] At each testing time, students were also evaluated on their receptive vocabulary knowledge. Yes/no vocabulary recognition tests (following Meara 1992) were developed and administered together with the productive vocabulary tests. Each vocabulary recognition test included 60 words, 20 of which were imaginary. Unfortunately when these tests were scored, using a formula that adjusts for guessing, we realized that there had been too much guessing, which meant that results were not reliable and had to be discarded.

[iii] The Flesch-Kincaid Grade Level formula has been used to calculate comprehension difficulty. The formula is based on the number of words and the number of syllables per word and results correspond to the number of years in education generally required to understand a text. For more information see https://readability-score.com/.