



Age and type of instruction (CLIC vs. traditional EFL) in lexical development

MARÍA PILAR AGUSTÍN-LLACH*
Universidad de La Rioja

Received: 14/02/2015. Accepted: 3/09/2015.

ABSTRACT

The present paper compares the vocabulary development of a group of CLIL and of traditional EFL learners along three years. The observation that a CLIL approach might provide with larger benefits in the long run vocabulary is the starting point of this study. We had learners in the two groups complete a letter-writing task. These writings were then scrutinized for L1 influence in the form of borrowings and lexical creations. The frequency of the words in the letters was also object of analysis. Results revealed that CLIL learners perform slightly better but non-significantly better than traditional EFL along the three years. Furthermore, the evolution of L1 influence and word use also followed an expected improvement pattern as learners went up grade. However, our results do not provide evidence of a growing CLIL advantage with increasing experience. The young age and low proficiency of learners in the present study might be blocking this possible advantage found elsewhere.

KEYWORDS: type of instruction, CLIL vs. EFL, age constraints, vocabulary acquisition, L1 influence.

1. INTRODUCTION

The observation that Content and Language Integrated Learning, henceforth CLIL, is better in the long run and that its effects start becoming visible after some experience with the approach (cf. Celaya & Ruiz de Zarobe, 2010; Pfenninger, 2014), sustains the present research. This claim is, therefore, the starting point of our study. Specifically, we looked at the production of borrowings and L1 influenced lexical creations which have been identified to change their patterns of appearance as learners gain in proficiency and get older (e.g. Celaya, 2008). We wonder, thus, how the evolution of borrowings and lexical creations

**Address for correspondence:* María Pilar Agustín-Llach. Universidad de La Rioja, Dpto. Filologías Modernas, C/ San José de Calasanz, 33, 26004, Logroño, La Rioja, Spain; e-mail: maria-del-pilar.agustin@unirioja.es.

would be along three years in traditional EFL and CLIL learners. In order to provide a more complete picture of lexical development, we also submitted learners' writings to lexical analysis of word frequency. Here, we wanted to observe how CLIL instruction affects use of lower frequency words or of specific words (e.g. from the CLIL subject) as experience with the approach increases. Again, we use traditional learners' productions as a reference.

2. VOCABULARY DEVELOPMENT: LONGITUDINAL STUDIES

Longitudinal studies dealing with lexical learning have shown the incremental nature of vocabulary acquisition (e.g. Schmitt, Schmitt & Clapham, 2001; Terrazas Gallego & Agustín-Llach, 2009). This means that as learners grow older, their vocabulary knowledge increases, in number of words incorporated to the lexicon as well as in number of word aspects added to each lexical entry, e.g. semantic, pragmatic, morphological, syntactic and/or phonological information. It is a widely acknowledged fact that vocabulary develops in stages, with specific lexical features or word classes (Marsden & David, 2008; Schmitt, 1998), word learning processes and associations established (Jiang, 2000; Meara, 1996; Robinson & Ellis 2008), word frequencies (Milton, 2009), or lexical inconsistencies (Agustín-Llach, 2011; Celaya & Ruiz de Zarobe, 2010) associated to each stage. The development of vocabulary knowledge is complex and to a certain extent unstable.

L2 proficiency, age, and hours of instruction are alluded to as reasons that delimit each stage (cf. Agustín-Llach, 2011). However, the weight of each variable on lexical development is not fully clear. Likewise, intralexical characteristics can also influence the "learnability" of words (e.g. Ellis & Beaton, 1993; González Álvarez, 2004; Laufer, 1990, 1997).

Agustín-Llach (2011) showed that not only do learners write longer compositions, but they also produce fewer lexical errors in general terms as they go up grade. This study also revealed that the lexical profile of learners changes over time with complex relations between age, L2 proficiency, and hours of instruction with production of lexical inconsistencies. Although not specifically aimed at describing lexical learning, Ruiz de Zarobe (2008) found that vocabulary acquisition progressed from non-CLIL to CLIL learners; this advantage increased with increasing CLIL experience.

Longitudinal studies of vocabulary acquisition have focused on the development of specific linguistic features such as verb meaning acquisition (Saji, Imai, Saalback, Zhang, Shu & Okada, 2011), phonological processing abilities which contribute to L2 vocabulary development (Nicolay & Poncelet, 2013), organization of the mental lexicon (Cui, 2009), breadth knowledge and vocabulary fluency (Zhang & Lu, 2013), or acquisition of lexical phrases (Li & Schmitt, 2009), among others. Lexical fluency has been found to relate with writing, so that a higher fluency leaves room for students to concentrate on other composing

aspects (Schoonen, van Gelderen, De Glopper, Hulstijn, Simis, Snellings & Stevenson, 2003). Moreover, fluency relates to word frequency, with higher frequency words being accessed faster. Age and exposure are identified in these studies as crucial variables for lexical development, together with intra-lexical features (e.g. word length, word origin, word frequency, word polysemy, among some others; cf. Laufer, 1990), which show complex developmental patterns difficult to systematize.

Lexical knowledge develops with proficiency but as previous studies reviewed in this section reveal, the nature of this development varies, it shows instability and is influenced by different interrelating factors (cf. Caspi & Lowie, 2013), being age a crucial one.

Longitudinal studies tend to focus on the development of a specific aspect by few learners. Longitudinal studies with CLIL are rare, especially in the area of vocabulary acquisition, not to speak of comparisons with traditional learners. The present paper intends to fill that gap.

3. VOCABULARY AND CLIL

The simultaneous instruction of content and foreign language in the CLIL class has been observed to exert a positive influence in some linguistic areas such as vocabulary, communicative abilities (creativity, fluency, risk taking), receptive abilities, morphology, or motivation, whereas other areas like syntax, writing, or pragmatics show no advantage from CLIL approaches (Dalton-Puffer, 2008). In a more recent review, Ruiz de Zarobe (2011) concludes that clear gains are observed for CLIL learners in reading, receptive vocabulary, speaking, some morphological phenomena, emotive and affective outcomes, fluency and complexity in writing and partially in listening. Syntax, productive vocabulary, informal language accuracy in writing and pronunciation were not found to be favorably affected by CLIL. Although the exact nature of the relationship between CLIL and lexical learning has not been teased out in an undisputable way, it seems generally acknowledged that they are positively related.

The acronym CLIL (Content and Language Integrated Learning) has been used to refer to all those teaching approaches in which the foreign language is used as the vehicle for content transmission and, although attention is paid to language issues in general, the focus is on meaning (cf. Dalton-Puffer, 2011: 186). This dual focus or integration of foreign language and content is the essence of CLIL, and we believe it is what makes CLIL so adequate for vocabulary teaching and learning, since it provides real and meaningful input for the learner in the form of subject content and language for classroom management (Muñoz, 2007). The CLIL approach is not a foreign language teaching approach, but a general pedagogical approach in which the foreign language is used to transmit content knowledge. As a result,

foreign language competence is believed to benefit from this vehicular foreign language use in context. Additionally, CLIL teachers are required to focus on language issues when necessary, i.e. with new, difficult or important linguistic aspects. Students in the CLIL approach still attend EFL classes. Therefore, what makes traditional learners different from CLIL learners is that whereas the former are exposed to the FL only in the FL class, CLIL learners also attend content classes held in the FL (content class + EFL class). In this sense, one could claim that the quantity and the quality of the FL exposure in both approaches differ.

In this sense, vocabulary is the aspect most dealt with in CLIL classes (Matiasek, 2005, in Dalton-Puffer, 2008: 145), which happens to be a very, if not the most, appropriate method for vocabulary development (Morgan, 1999, in Sylvén, 2010: 33). Sylvén (2010) found that in the CLIL class learners learn mostly vocabulary, display risk-taking behaviors in lexical use, develop technical and academic (i.e. subject specific) vocabulary, and establish semantic relationships. When compared to counterparts in traditional approaches, CLIL learners have been found to display better lexical knowledge (Lasagabaster, 2008; Ruiz de Zarobe & Celaya, 2009; Xanthou, 2011). The longer, and more meaningful and contextualized instruction in CLIL classes has revealed itself beneficial for raw lexical retention measured through bilingual lists (Xanthou, 2011), general vocabulary knowledge (Ruiz de Zarobe & Celaya, 2009), receptive vocabulary (Canga Alonso, 2013; Jiménez Catalán & Ruiz de Zarobe, 2009), rich and deep vocabulary (Jiménez Catalán, Ruiz de Zarobe & Cenoz, 2006; Moreno Espinosa, 2009), and a more appropriate and accurate lexical use with fewer instances of lexical inconsistencies (Ackerl, 2007; Agustín-Llach, 2009; Celaya, 2008).¹ But the picture is far from being complete and further research contrasting both approaches in lexical learning is needed.

In relation to this, Celaya and Ruiz de Zarobe (2010) found out that the type of instruction provided by CLIL favors L2 vocabulary development, the tendency towards lexical creations being increased and the possibilities of lexical borrowing being diminished. Lexical creations are typically produced by high proficient learners, whereas borrowings are typically seen at beginning stages of acquisition (Celaya & Ruiz de Zarobe, 2008; González Álvarez, 2004). Age and L2 proficiency might be influencing this dynamics of L1 influence, as well (Celaya & Ruiz de Zarobe, 2008, 2010). In a seminal paper on L1 influence and bilingualism, Poulisse and Bongaerts (1994) already established that L1 influence is a matter of combined L1 and L2 proficiency, with higher proficient learners being less permeable to L1 lexical intrusions. Few studies investigate the role of L1 influence in CLIL learners. Among those, Celaya (2008) and Agustín-Llach (2009) found that CLIL and traditional EFL learners seem to differ in the type of L1 influence instances. In this sense, the communicative nature of the CLIL approach leads us to believe that transfer phenomena will be rarer in this context, which somehow simulates natural second language acquisition, than in traditional EFL or English-as-a-school-subject contexts. Thus, Rokita (2006) also noticed in an analysis

of code-mixing episodes in very young early bilinguals and L2 learners that whereas the former conceived English as a tool to communicate, for the latter it was something they had to learn to please their parents, and never really used English to interact. Similarly, Agustín-Llach (2014) found that CLIL learners display few instances of L1 borrowing and some more of L1-based coinages, but produced much higher numbers of L1-based phonetic renderings. This is in line with Celaya and Torras (2001), Dewaele (1998, 2001), Gabryś-Barker (2006), and Ringbom (2001), who showed that as proficiency increases meaning-related transfer becomes more common. The communicative approach used for the instruction of the CLIL learners may also serve as evidence for the meaning-related transfer, which is more common than form-related L1 influence (cf. Ecke, 2001). Results of CLIL and non-CLIL settings parallel those between proficient and less-proficient FL learners.

Additionally, one would assume that CLIL learners' lexical production is characterized by including low frequency words which stem from the CLIL subject. We have not found, however, systematic examination of this issue. In general, previous studies examining the relationship between CLIL and vocabulary acquisition frequently omit comparisons with traditional learners (e.g. Canga Alonso, 2013; Jiménez Catalán & Ruiz de Zarobe, 2009), and if they do include those comparisons, then they are mostly made controlling for hours of instruction rather than age (e.g. Celaya, 2008; Celaya & Ruiz de Zarobe, 2010). In this sense, the present study compliments previous research.

Accordingly, we believe that the CLIL vs. non-CLIL variable could be very useful and advantageous in lexical studies, since it could help tease out the lexical learning process. Nevertheless, there are, to our knowledge, few studies investigating how CLIL and lexical learning relate with the reference of traditional EFL general vocabulary acquisition. Longitudinal studies in this field are even rarer, especially with young primary school learners. The present study intends to fill this gap in research and to contribute to disentangling the role of CLIL instruction in general vocabulary acquisition in the L2 along the last 3 years of primary education.

In the present study, we were interested in examining the lexical development of young EFL learners in a traditional instructional approach and in a CLIL approach. Comparison cohorts have been traditionally chosen on the basis of similarity of hours of exposure to the FL, not of age and grade similarity. On most occasions, research deals with secondary school students (Celaya, 2008; Celaya & Ruiz de Zarobe, 2010), and only rarely are primary school learners the focus (e.g. Agustín-Llach, 2014; Canga Alonso, 2013), as they are here. We wanted to ascertain whether CLIL exerts a positive influence in lexical learning and what the evolution of this influence is. Thus, we set out to answer the following specific research questions:

1. How does the general English lexical profile of traditional learners develop along 3 years?

2. How does the general English lexical profile of CLIL learners develop along 3 years? Do CLIL-induced lexical gains transmit to general L2 communication as observed in a writing assignment with a general topic?
3. Are these developments comparable or similar in both groups of learners across the three years tested?

4. METHOD

This study presents a longitudinal and cross-sectional research which follows two groups of students along three years. Students are, therefore, compared both longitudinally (within their group) and cross-sectionally (between groups). This design allows us to obtain data of the same group of students relative to their lexical development, but it also allows us to establish stage-wise and longitudinal-wise comparisons of two different groups of learners.

Several reasons have persuaded us to establish comparisons controlling age instead of hours of exposure to the FL. First, this is done for the sake of ecological validity, since a study that compares learners of the same age and school grade is representing school and educational reality more faithfully than when learners at different grades and ages are compared. Also, this comparison design allows us to better paint the picture at the end of primary education, mirroring what is happening now in our primary schools in Spain. Furthermore, learners at the same grade find themselves at the same stage in the EFL curriculum (English as a subject), and this is especially important for comparisons of general English. This design is also supposed to ensure that no learner cohort profits from cognitive advantages due to age (cf. Pfenninger, 2014). Finally, the fact that young learners are “slow learners” (cf. Nikolov, 2014) led us to think that hours of exposure to the FL at such an early age might not be so relevant as with older learners.

4.1. Participants

Two groups of learners participated in this study. One group, hence the traditional or non-CLIL group, was made up of 61 learners. The other group, hence the CLIL group, comprised 68 learners. They had been learning English as a school subject since the 1st grade in three weekly sessions of 50 to 60 minutes. Additionally, the CLIL group had been receiving extra-exposure to English through Science for an additional two hours a week since their 1st grade. Accordingly, we followed all 4th graders for 3 years (4th, 5th, and 6th grades) in the school, first when the school only had English as a school subject, and some years later when CLIL tuition was introduced. At the first data collection moments, participants were between 9–10 years old and attended 4th grade of Primary education; at the end of the study, in the third data collection sessions, these learners were between 11 and 12 years old and attended 6th grade of

Primary. The participant groups differed in the kind of instruction they received, i.e. CLIL vs. non-CLIL, and consequently, in the number of hours of exposure to English FL. Learners in the non-CLIL group were exposed to English through the English FL school subject exclusively. However, learners in the CLIL group received, apart from the weekly EFL lessons, input in English in the school subject Natural Sciences, which was taught through the medium of English. Thus, traditional learners had received approximately between 105–110 hours of exposure to EFL on a yearly basis since 1st grade of Primary. The CLIL group had received these 105–110 hours plus 72–74 more hours in CLIL science, also since 1st of Primary. Table 1 illustrates the approximate number of hours of exposure students had received by the three times of data collection.

Both CLIL and traditional EFL learners follow the official curriculum imposed by the authorities for Natural Sciences in English and for the general EFL class. Every student enrolled in the school participated in the CLIL program with no exceptions or extra selection of pupils.

Since this is a longitudinal study, only learners who were present at the three data collection moments were considered for this study. Learners for whom we only had one or two data sets were eliminated from the study. Both student samples attended the same school in a northern region of Spain but some years apart; therefore, the sample is homogeneous as regards their socio-economic and cultural background. Data was first collected when the school was monolingual and again three years later when the CLIL program had been introduced and established for the whole student population entering 1st grade. Students also shared Spanish as their mother tongue (L1). Furthermore, learners who are beginners in EFL (A1 level) were attending same-proficiency-level EFL courses at school:

GRADE	AGE	CLIL (N= 68)	NON-CLIL (N= 61)
4 th Primary	9–10	714	419
5 th Primary	10–11	839	524
6 th Primary	11–12	944	629

Table 1. Hours of exposure to English FL (accumulated exposure)

4.2. Data-gathering instruments

We had students complete a writing assignment for 30 minutes with no word limit or other linguistic constraints.² They had to write a letter to a prospective host family in England. In this letter, they had to introduce themselves and talk about their hometown, school, family, and any other interesting thing about themselves. No help from the teacher or other classmates was allowed nor the help from grammars or dictionaries. We used a composition because we wanted to examine free production of general vocabulary. Considering the design of the study with a longitudinal comparison plus the cross-sectional CLIL vs. non-CLIL comparison, the free composition was deemed most appropriate in order to explore the

evolution of vocabulary use in spontaneous production and to compare between groups. Furthermore, this task and general letter topic was chosen to find out whether CLIL had an impact on general lexical development beyond the subject-specific or technical domain of the CLIL course. Additionally, the topic imposed few linguistic challenges, and learners could make use of the structures or lexical items that they had at their disposal and somehow adapt the topic to align with their own linguistic knowledge. This topic was also perceived to be of interest to the young learners.

4.3. Procedures and analysis

Letters were written during regular class time at six moments of data collection: during 3 consecutive years for the non-CLIL group, and later for another 3 consecutive years for the CLIL participants. Compositions were handwritten and later typed into computer-readable files and submitted to an analysis in two main phases. The first one wanted to examine learners' production of lexical errors as the result of the use of their L1. Lexical transfer has been found as a reliable way to look into lexical learning (cf. Celaya & Ruiz de Zarobe, 2010). The second phase included a submission of the learners' writings to a lexical profiler program where the frequencies of the words used were identified. This allowed us to determine learners' lexical progression (cf. Milton, 2009).

4.3.1. Phase 1. Lexical error analysis: L1 influence

In order to examine learners' recourse to their L1, we identified and classified instances of L1 influence into borrowings and lexical creations.

Borrowings are mere insertions of L1 words in the L2 syntax without any attempt at adaptation (Celaya & Torras, 2001). The lack of lexical knowledge generally lies behind this normally conscious communication strategy. Traditionally, young and low proficiency learners have been found to recur to borrowing more frequently (Celaya & Ruiz de Zarobe, 2010). The following example from our data illustrates the phenomenon of borrowing:

- (1) My favourite *comida* is spaghetti, strawberries, ham, chicken. (Sp. 'comida' for Eng. 'food')

Lexical creations, on their part, are adaptations of L1 words to the L2 morpho-phonological rules (cf. Celaya & Torras, 2001). They imply higher L2 linguistic knowledge and metalinguistic awareness, since L2 rules come into play. They are, therefore, more typical of the discourse of higher proficiency learners. Learners generalize L2 phonemic, morphological, or spelling rules to adapt an L1 word so that it looks or sounds English. An example illustrating an instance of lexical creation follows:

- (2) My rabbit is small, very *divert*. ('*divert*' from Sp. '*divertido*' for Eng. 'funny')

Measures of L1 inconsistencies were taken and raw numbers converted into error density or percentage of L1 errors every 100 words. Thus, this measure takes into account composition length and is more accurate than absolute measures of instances produced.

4.3.2. Phase 2. Lexical analysis: Vocab. profile

We decided to conduct some analysis with the VocabProfiler (version for kids) available at www.lextutor.ca (Tom Cobb) to look for word frequencies. This test allows the researcher to establish a measure of the type of words and frequencies of words learners produce. Specifically, in order to probe possible finer-grained differences in vocabulary knowledge (frequency levels), we decided to run VP-Kids,³ an instrument designed to gauge young learners' vocabulary sizes. It has ten smaller frequency bands plus an off-list category made up of 250 words taken from a children corpus (<http://www.lextutor.ca/vp/kids/>). It covers the first ten levels of frequent words, i.e. from the one thousand most common, the second thousand, and so on until the ten thousand most common words list. In the results section, we will focus on analysis of the least frequent levels, because higher and more fine-grained differences are to be expected in the words of the three thousand least frequent words. Off-list words are words included in the English corpus or dictionary but which are not collected in any of the frequency lists the VocabProfiler works with. On the contrary, off-list unknown words are those which cannot be found in either dictionaries or corpora, for instance errors, misspellings, or borrowings.

5. RESULTS

We were interested in comparing the lexical profiles of CLIL and traditional EFL learners by first examining the evolution of L1 influence along three years in their written productions. Specifically, we concentrated on identifying instances of borrowings and lexical creations.

The following table (Table 2) presents the mean figures of error density or instances of error every 100 words for each of the L1 transfer types for each group at each grade level:

Grade	NON-CLIL (N= 61)		CLIL (N= 68)	
	Borrowing	Lex. creation	Borrowing	Lex. creation
4 th Primary	1.21	0.62	0.69	0.94
5 th Primary	0.87	0.90	1.16	0.96
6 th Primary	1.43	1.02	0.65	1.16

Table 2. Borrowings and lexical creations every 100 words (in means) in traditional and CLIL learners across their last years of primary education

These results were slightly surprising, since they did not fully match our expectations based on previous research. We were surprised in particular by two observations. First, the increase in borrowing production from 5th to 6th grade in the traditional EFL group and from 4th to 5th in the CLIL group seems to run counter to previous research findings (e.g. Celaya & Ruiz de Zarobe, 2010), which show that as proficiency and grade increase, borrowing production decreases. Second, the fact that for 5th grade CLIL learners produce more borrowings than the traditional group was also surprising, because they run counter to previous research and, therefore, to our expectations (cf. Agustín-Llach, 2009, 2015; Celaya, 2008; Celaya & Ruiz de Zarobe, 2010). These two results might be related, since they involve both 5th grades.

In order to ascertain whether these differences were statistically significant, we conducted non-parametric tests of means comparison. Tests selection was made on the basis of normality assumptions. Thus, since the sample did not meet the normality assumption, Mann-Whitney tests for two independent samples and Wilcoxon tests for two related samples were carried out. Significance value was set at $p < 0.05$. Results are organized in two sets. First, we looked at differences between traditional and CLIL learners in their production of borrowings and lexical creations, i.e. a between-groups comparison. Results revealed that, except for borrowings in grade 6, where CLIL learners exhibited a lower rate than non-CLIL learners, differences were not significant.

Table 3 shows the figures for all cases:

	borrowing_4	borrowing_5	borrowing_6	creation_4	creation_5	creation_6
Mann-Whitney U	1928.5	1939.5	1634.0	1779.0	2001.5	1780.0
Sig. (two-tailed)	.447	.477	.018	.122	.702	.137

Table 3. Inferential statistical test results for traditional vs. CLIL learners

Second, we were also interested in examining differences between the evolution of borrowings and lexical creations along the three years tested for both groups, i.e. a within-groups comparison. Results of the Wilcoxon signed rank test for two related samples reveals that for traditional learners no significant differences could be found among the borrowings and lexical creations produced over the three years. Accordingly, the production of borrowings did not experience significant changes from 4th to 5th grade, from 5th to 6th, nor even from 4th to 6th. The same conclusion stands for lexical creations. Table 4 offers these statistical results:

	borrowings_5 – borrowings_4	borrowings_6 – borrowings_5	borrowings_6 – borrowings_4	creations_5 – creations_4	creations_6 – creations_5	creations_6 – creations_4
Z	-.404	-1.308	-.238	-1.132	-.184	-.927
Sig. (two-tailed)	.686	.191	.812	.258	.854	.354

Table 4. Inferential statistics for traditional learners in 4th, 5th, and 6th grade

For CLIL learners, the picture becomes a bit more complex. Lexical creations show the same pattern commented above of non-significant differences across grades. However, borrowings significantly increase from 4th to 5th grade, to again significantly decrease from 5th to 6th grade. No significant differences were found in borrowing production between 4th and 6th grade. Table 5 shows this set of results:

	borrowings_5 – borrowings_4	borrowings_6 – borrowings_5	borrowings_6 – borrowings_4	creations_5 – creations_4	creations_6 – creations_5	creations_6 – creations_4
Z	-2.063	-2.645	-.768	-.232	-.883	-1.027
Sig. (two-tailed)	.039	.008	.442	.817	.377	.304

Table 5. Inferential statistics for CLIL learners in 4th, 5th, and 6th grade

A more in-depth description of learners' lexical profiles was also intended to in this study. We wanted to look into the frequency of the words produced and the specific words produced by members of each group. Results of VP-Kids comparisons revealed an even more complex picture of the interrelations between grade and teaching approach in the frequency bands of the words used. Tables 6 and 7 below offer the figures for traditional and CLIL learners, respectively. Specifically, we have included the percentage of coverage of words from the 10 first levels of frequency (see above method).

Again these figures show similar results for CLIL and traditional learners at grade 4. If we go through the figures of the percentage of words belonging to each frequency level (from 1 to 10), we can observe that all over the 10 levels figures are comparable, especially at the higher frequency levels. The figures show a complex picture with a general increase of off-list known words (lower frequency words). Likewise, CLIL learners also display higher number of these lower frequency words, which tentatively might be pointing to better vocabulary knowledge on part of the CLIL learners. Moreover, and also pointing in the direction of a tentative CLIL advantage, the figures for off-list unknown words is higher in traditional EFL learners' writings. These might be the result of higher numbers of misspellings, lexical inconsistencies, and lexical errors in these writings, and show thus indirectly lower levels of proficiency in vocabulary:

FREQUENCY LEVEL	4 TH GRADE	5 TH GRADE	6 TH GRADE
Kid250-1	67.64	62.76	69.78
Kid250-2	8.95	9.51	10.05
Kid250-3	4.06	3.79	3.52
Kid250-4	2.35	2.50	2.65
Kid250-5	0.75	1.06	0.88
Kid250-6	1.52	1.50	1.28
Kid250-7	0.88	0.79	0.80
Kid250-8	0.67	0.72	0.84
Kid250-9	0.30	0.40	0.27
Kid250-10	0.36	0.19	0.19
Off-list known	0.69	0.94	1.45
Off-list unknown	11.84	15.85	8.29

Table 6. Percentage of word coverage for traditional learners

FREQUENCY LEVEL	4 TH GRADE	5 TH GRADE	6 TH GRADE
Kid250-1	65.99	71.10	63.06
Kid250-2	11.82	12.67	10.57
Kid250-3	4	3.66	3.30
Kid250-4	2.87	3.53	3.07
Kid250-5	0.97	0.75	0.63
Kid250-6	1.11	1.35	1.10
Kid250-7	0.84	1.23	0.63
Kid250-8	0.53	0.85	0.74
Kid250-9	0.33	0.28	0.23
Kid250-10	0.20	0.30	0.25
Off-list known	1.44	2.08	1.79
Off-list unknown	9.91	2.21	14.65

Table 7. Percentage of word coverage for CLIL learners

In order to get an even finer vision of the similarities and differences between traditional and CLIL learners, we looked into the specific words of the three lowest frequency levels (level 8, 9, and 10) for each grade (Tables 8, 9, and 10). Shared words are marked in bold and the number of participants who produced those words is found between brackets when it is different from one. Only correctly used words are considered here. For example, the word ‘lecture’ appears in learners’ writings; however, when it is misused and confused with the word ‘reading’ (‘lectura’ in Spanish), it is eliminated from these counts. Analysis of the writings allows us to obtain this information:

	TRADITIONAL	CLIL
Frequency levels 8, 9, 10	fan, fine, jungle, market, thin , pear , ping-pong, traffic-light, vase, watermelon, taxi, umbrella, homework (2) , toilet (2), magazine (2), hundred (2), skirt (2), swimming-pool (2) , lemon (3), March (3) , July (3) , hobby (3) , adventure (5), vegetable (6) , biscuit (7) , violin (7), zebra crossing (7), guitar (8) , piano (10) , hall (16), spaghetti (24).	July , pear , December, fan , guitar , March , piano , rat, thin , underground, age, president, important, vegetable , salad, scissors, angry (2), busy, calendar (2), biscuit (2) , expensive (2), earth (2), pencil sharpener (2), restaurant (3), thousand (3), homework (3) , hobby (4) , rice (5), swimming-pool (11) .

Table 8. Same and different low frequency words for 4th grade

	TRADITIONAL	CLIL
Frequency levels 8, 9, 10	hall, onion, president, someday, steak , sponge, strawberry, terrible, mushroom, watermelon , crisp, hundred, united states, dining-room, expensive , zebra crossing, insect, scout , stadium , violin , young (2), biscuit (2), map (2), lazy (2) , March (2), market (2), swimming-pool (2) , guitar (3), July (3) , skirt (3), homework (4) , pear (4), salad (4), hobby (5) , piano (5) , theatre (5), rice (5) , important (6) , parent (6) , spaghetti (8) , restaurant (10) , thin (12), vegetable (31) .	calendar, collection, designer, instrument, lazy , steak , traffic, violin , volcano, watermelon , grape, thousand, address, adventure, list, scout , expensive , homework (2) , July (2) , guitar (2), age (2), stadium (3) , piano (3) , spaghetti (3) , swimming-pool (3) , cafeteria (3), restaurant (4) , rice (4) , vegetable (4) , important (6) , parent (8) , hobby (32) .

Table 9. Same and different low frequency words for 5th grade

	TRADITIONAL	CLIL
Frequency levels 8, 9, 10	December , adventure, biscuit, fan, guinea-pig, instrument, worse, crisp, expensive, enjoy, exercise , foggy, pear , strawberry , taxi, theatre , polar, skateboard , snail, stadium , weather (2) , date (2), goldfish (2), factory (2) , ham (2), hundred (2), important (2) , salad (2), July (2) , palace (2), restaurant (3) , rice (3) , thousand (3), spaghetti (4), steak (4), ping-pong (5), piano (6) , thin (7), March (7), guitar (8) , hall (8) , holiday (8) , homework (8) , vegetable (13) , swimming-pool (16) , parent (23) , hobby (29) .	airport, desert, December , grape, holiday , jam, mustache, newspaper, punish, sugar, toilet, underground, theatre , Africa, age, alarm, cafeteria, diamond, homework , list, separated, skateboard , stadium , during (2), exercise (2) , mystery (2), strange (2), strawberry (2) , July (2) , violin (2), weather (2) , young (2), hall (3) , March (3), thin (3), pear (3) , scout (4), factory (5) , guitar (5) , vegetable (5) , important (5) , rice (5) , restaurant (6) , hobby (7) , parent (7) , swimming-pool (9) , piano (23) .

Table 10. Same and different low frequency words for 6th grade

As can be seen from the three tables above, CLIL and non-CLIL learners display similar numbers of infrequent words. Likewise, they display very similar numbers for shared and non-shared low frequency words. The frequency of these words in students' letters is

very similar as well, if the number of different students who produces them is considered. This does not allow us to reach undisputable conclusions in favor of a CLIL advantage.

Additionally, we also searched for words related to the field of science, the CLIL subject, in both groups of learners. We identified manually in the data from the letter corpus words related to animals, the environment, plants, or food.

	TRADITIONAL	CLIL
4th grade	Fingernails, monkey, rabbit, monster, eagle, science , north, lion (2) , dinosaurs (2), plants (2), dolphin (2) , snake (4) .	Venus, Mercury, Neptune, Pluto, Scorpio, asteroid, lion , wolf, snake , spider, tiger, turtle, world, sea, dolphin , rat, underground, earth (2), Mars (2), Uranus (2), constellation (2), planet (2), Saturn (3), bat (3), metal (3), science (4) .
5th grade	Flower, mouse, rock , insect, brain, chest, eagle, snake, spider, daisy, elephant, shark, seed, sand, planet , wildlife (2), dinosaurs (2), north (2) , spring (2), world (2), sun (3), plant (4) , sea (4) , rabbit (5) river (6) .	Mediterranean, underworld, reptiles, amphibians, monkey, rock , plastic, volcano, inventor, survive, star, planet , plant , river , sea , fox, grass (2), turtle (2), squirrel (2), world (2), Recycling (2), giraffe (2), north (2) , science (6).
6th grade	Architecture, bee, galaxy , bear , duck, land , rabbit, snake , star , tiger, wood, farm, shark , moon, planet, guinea-pig, nature, fog, polar, snail, eagle , north (2) , goldfish (2), geography (2), monument (2), rock (2) , plastic (2), turtle (2) , dolphin (2), plant (2), temperature (3), world (4) , ice (6), science (10) .	Dinosaur, dolphin, galaxy , liver, minecraft, bear , hill, land , island, snake , space, star , eagle , north , elephant, forest, shark , wind, tear, sea , underground, industry, liquid, universe, lion (2), monkey (2), rock (2) , locate (2), robot (3), ice (3), world (4) , turtle (6) , science (8) .

Table 11. Science words

As can be seen from the words in Table 11, learners in the traditional and CLIL groups produce roughly similar numbers of words related to the field of Science. Likewise, in both groups we find some shared words and also some non-shared or idiosyncratic words. As learners go up grade, they write more words of this field, but differences between the instruction approaches are small. This result might point to a lack of transfer from the CLIL subject to general English use.

Vocabulary related to school and classroom activities and management was also very frequent in both groups of learners. In this sense, we could not find idiosyncratic vocabulary distinguishing both groups. They all produce words like ‘playground’, ‘homework’, ‘pencilcase’, ‘bookshelf’, ‘blackboard’, ‘sharpener’, ‘ruler’, and ‘rubber’ along the three years tested. This might point to these words being learned in the EFL class under classroom discourse.

Previous research on language learning from a cognitive perspective (cf. Ruiz de Mendoza, 2013), revealed that production of general words of the type ‘thing’, ‘put’ or ‘make’ is very infrequent in younger learners, probably because of a lack of a fully developed

cognitive system. Banking on this observation, we wanted to check in our learners' production and found the low presence of these general words. Results in table 12 show that they are more frequent with increasing grade, and also more frequent for CLIL over traditional learners (measures taken every 100 words):

	'THING'	'PUT'	'MAKE'
4 th traditional	0	0	1
5 th traditional	7	0	4
6 th traditional	23	3	4
4 th CLIL	7	0	4
5 th CLIL	12	0	4
6 th CLIL	15	3	10

Table 12. Production of general words for traditional and CLIL learners

6. DISCUSSION

In the present study, we wanted to examine the evolution of vocabulary development in a group of traditional learners and another of CLIL learners, and compare them. Our results lead to inconclusive findings, some of them rather unexpected. Our first objective was to compare their lexical profiles through examination of lexical errors caused by L1 influence along three years for CLIL and traditional learners and compare this evolution.

The production of lexical creations follows the expected patterns. First, CLIL learners produce more instances of lexical creations than traditional learners for all the three data collection times. Second, the production of lexical creations increases with grade for both learner groups. Differences are, nonetheless, non-significant in either case, but they show a clear tendency. Lexical creations are related to higher level of proficiency, so that they have been found to be more frequent as learners' proficiency increases (Celaya & Ruiz de Zarobe, 2008; González Álvarez, 2004). In this sense, CLIL learners can be said to indirectly display higher levels of EFL mastery. The fact that learners in both groups find themselves at the same stage in the general EFL class might also be exerting some influence and contribute to explaining lack of significant differences.

The case of borrowings in this specific study is puzzling for unexpected. In general terms, CLIL learners produce fewer instances of borrowing than traditional learners and these tend to decrease with increasing proficiency. However, we have two exceptions here. Traditional 6th graders and CLIL 5th graders produce surprisingly many borrowings. This deviation from the general tendency is hard to explain and has no apparent reason to be, except for the instability of the interlanguage of this low intermediate stage in 5th grade. Borrowings are typical of low level learners and can be generally traced back to overall lack of lexical knowledge in the L2 (Celaya & Ruiz de Zarobe, 2010; Celaya & Torras, 2001) and

probably to the inability of learners to generalize L2 rules, i.e. lack of metalinguistic awareness. The fact that words become more difficult and more abstract with increasing grade, and thus learners are more prone to use the borrowing strategy, might be a reason for this. These results, however, do not allow us to conclude that CLIL gains benefits in the long run or with increasing experience, since differences with traditional learners remain stable across grades, except for borrowing production, with CLIL learners in 6th grade producing significantly fewer than traditional EFL learners. This might be a faint light in the direction of CLIL benefits, which again may differ in the various areas of vocabulary examined, as its influence changes from language competence to language competence (cf. Dalton-Puffer, 2011; Ruiz de Zarobe, 2011). In this sense, our results coincide with previous findings that accuracy in writing is not favorably affected by CLIL (cf. Ackerl, 2007; Ruiz de Zarobe, 2011).

We also wanted to examine the frequency of the words produced in learners' writings. In the results section, we have detailed the percentage of coverage of the different frequency bands; if we specifically look at the off-list words (known), we see that CLIL learners' figures are slightly higher than those from traditional learners. This might again point to CLIL learners deploying higher levels of L2 vocabulary mastery, faintly supporting, thus, Várkuti's findings (2010) that CLIL learners are better at performing academic tasks, such as the letter-writing task in the present study (see also Meara & Bell, 2001; Vermeer, 2004). Again, we are unable to indisputably conclude in favor of CLIL increasing advantages in the longer run. A lengthy immersion time might be necessary for (possible) CLIL-induced lexical gains to transmit to general L2 communication tasks. Again, our results here confirm previous findings in that informal or non-technical vocabulary is not affected positively by CLIL, nor is productive vocabulary (cf. Ruiz de Zarobe, 2011).

However, as Singleton (1999: 51) highlights, having knowledge of the words in isolation is no guarantee of the ability to recognize or adequately deploy those words in context. To put it shortly, it might be the case that CLIL learners have higher levels of knowledge of isolated words as a result of their longer exposure (in the form of spoken and written input in the explanations and activities, but also in the form of output in the realization of practice exercises) to the FL, in particular within the field of science. But despite this, they might be unable to transfer this knowledge and use these words in context. (Low) L2 proficiency and (young) age have been found to constrain transfer abilities (e.g. Cabaleiro González, 2003), and this is precisely what might be happening here.

Finally, we looked at specific word production and focused on four main areas: least frequent vocabulary, science (CLIL subject) vocabulary, school vocabulary, and general words. In the particular case of low frequency words, we saw that the number of words is very similar between both groups, with an increasing tendency with ascending grade. Figures for least frequent words are very low. This might most likely respond to the fact that learners have a low proficiency level, but even with increasing proficiency learners prefer the use of

frequent and smaller error risk vocabulary (cf. Ringbom, 1998), maybe because high-frequency words have been found to be accessed faster (Zhang & Lu, 2013), they are learned earlier, and learners are more familiar with them. Additionally, we could attest the presence of shared and non-shared words to similar extents.

Regarding the semantic field of science, the CLIL subject, we could not find CLIL learners displaying specific science vocabulary not found in the productions of traditional learners. We could not find evidence of transfer of vocabulary from the CLIL subject into general use of English, except for some occasional words such as ‘minecraft’, ‘volcano’, ‘liquid’, or ‘universe’. These might be pointing to an incipient transfer. In this sense, our results do not support Zhang and Lu’s (2013: 18) statement that “the frequency of words in the specific context of L2 learning and use may have a stronger effect on learners’ vocabulary acquisition than general word frequency”. Nevertheless, we need to bear in mind that the latter topic is far too general and might be preventing learners from displaying all the word knowledge they have, especially as concerns science words, or academic or infrequent vocabulary. Another more specific composition topic could give rise to more instances of vocabulary transfer from the CLIL class. This stands as a major limitation for the present study. Furthermore, learners’ subjective willingness to produce vocabulary according to opportunity and wish might also be playing a role (cf. Nation, 2001; Zheng, 2012).

Likewise, we could not find differences in the vocabulary of the field of school between the two groups of learners. This might be explained by the fact that this vocabulary is learned in the EFL class, so it is shared by both groups and used as classroom language probably in the EFL class.

Finally, the lack of general words in the learners’ production might answer the still limited cognitive development of learners which is linked to age and which is thus more independent from instruction approach (cf. Ruiz de Mendoza, 2013). Despite the fact that learners in this study share age, it is still a very relevant factor, since it helps us explain lack of differences found between the two groups, one of which had received more input in the L2.

CLIL is more demanding in cognitive terms, so that might explain lack of differences in English vocabulary production in favor of CLIL learners despite their extra exposure hours to the FL. CLIL learners need to pay attention to subject content and not only to the new L2 words and structures. We think this might slow down the process of vocabulary acquisition in general English and CLIL learners need more exposure time to catch up in general English.

Our results and interpretation point to two related conclusions. First, as concerns the theoretical import of the study, we did not find compelling evidence in favor of CLIL being beneficial in the long run when young, low proficient learners are concerned, although some faint tendencies were observed. Differences in the amount of input between the two groups might be insufficient or too small for them to reveal larger advantages. As Nikolov (2014) suggests, young learners are slow learners and they need substantial amounts of input to

profit from the beneficial effects of continued exposure. This result is related to Pfenninger (2014), who also found that CLIL learners cannot gain the positive effects of the combination of the implicit and explicit EFL learning approaches until they have gained cognitive maturity and enough experience with the CLIL teaching approach.

Second, the results of this study have an important effect on our knowledge of L2 vocabulary acquisition. Namely, we believe that age imposes a strong constraint in L2 lexical development, which is even more relevant a factor than exposure time per se. Again, Pfenninger's results (2014) point to the same direction, with age at testing being a relevant factor influencing learners' L2 performance. Cognitive development, plus a combination of continued, long, and massive exposure, and implicit and explicit instruction, are key factors in successful foreign language development.

7. CONCLUSION

The present study aimed at comparing the lexical production across three years of a group of traditional EFL learners and another of CLIL learners. Specifically, we focused on two main aspects: lexical transfer from L1 influence and word frequency. Interpretation of the results obtained is not an easy task, since they do not wholly meet our expectations based on previous research that CLIL advantage manifests with increasing CLIL experience in particular.

Generally speaking, our findings show that CLIL learners are slightly better foreign language vocabulary users than members of the traditional group, since the former display lexical transfer behaviors typical of more proficient learners, with fewer borrowings and more lexical creations. Likewise, as learners in both groups go up grade, their increased proficiency is reflected in their lexical transfer patterns as well. However, these are only faint tendencies, since we could not find significant evidence of a growing advantage of the CLIL approach over time.

Similar conclusions can be drawn for results related to word frequency. Again, in general terms, CLIL learners produce slightly more words of lower frequency than traditional learners and this tends to increase with grade, age, and proficiency. However, the expected larger benefits of the CLIL approach in the long run could not be attested with our data.

Learners' young age is alluded to as the overriding factor explaining these results. Age seems to be a better predictor of lexical behaviors over time than longer exposure time, and instructional approach. The young age and low proficiency of learners might be blocking further CLIL and exposure time advantages. Unfortunately, we have little information about how the CLIL courses were implemented; and this stands as another main limitation of the study. Apart from knowing that teachers were Spanish L1 speakers, primary school teachers

with a B2 level of proficiency in English, we are actually unaware of how much English was used in the class by the teachers and pupils, or any other qualitative feature of the specific CLIL classes. Without further probing into the nature and quality of the CLIL program under examination, our conclusions remain speculative.

In the future, it would be interesting to examine learners' textbooks (both EFL and science ones), to look into the frequency of the different words in the input and compare it to the learners' productions. We think that exploring the *Common European Reference Framework* (2001) level of the words produced by members of both groups might also be a rich future avenue for research.

ACKNOWLEDGEMENTS

This research was funded by grant no. FFI 2010–19334 by the Spanish Ministry of Science and Competitiveness.

NOTES

- 1 In most of these studies participants are learners of the same age but who have received more exposure hours. This is not an irrelevant detail, because age might be influencing the possible benefits of extended exposure, as we claim to be the case here, and similarly, CLIL learners might be favored by more exposure hours rather than by the different type of exposure provided by the CLIL approach. Teasing out both amount and type of exposure is a relevant issue.
- 2 This letter-writing task topic had been previously used in other national research projects funded by Spanish Ministry of Science and Education of which this research forms part (reference numbers: BFF 2003–04009–C02–02; HUM 2006–09775–C02–02). This allows for comparison among learners with different characteristics and for comparison of different variables.
- 3 The VP-Kids list project was initiated by Prof. Roessingh in Canada, who used a list of words from *Word Express: The first 2500 words of spoken English* (Stemach & Williams, 1988). This list was a composite of several word lists for pre-school aged native English speaking children (10 word lists: 10 x 250 words) (Pinchbeck, personal communication). We think this list is the best instrument to date to obtain the lexical profile of texts written by young CLIL and traditional EFL primary school learners.

REFERENCES

- Ackerl, C. (2007). Lexico-Grammar in the Essays of CLIL and non-CLIL Students: Error Analysis of Written Production. *Views*, 16(3), 6–11. Special issue: *Current Research on CLIC 2* (U. Smit & C. Dalton-Puffer (Eds.)). Retrieved September 3, 2015 from https://anglistik.univie.ac.at/fileadmin/user_upload/dep_anglist/weitere_Uploads/Views/Views_0703.pdf.
- Agustín-Llach, M. P. (2009). The role of Spanish L1 in the vocabulary use of content and non-content EFL learners. In Y. Ruiz de Zarobe & R. M. Jiménez Catalán (Eds.), *Content and Language*

- Integrated Learning: Evidence from Research in Europe* (pp. 112–129). Bristol: Multilingual Matters.
- Agustín-Llach, M. P. (2011). *Lexical errors and accuracy in foreign language writing*. Bristol: Multilingual Matters.
- Agustín-Llach, M. P. (2014). Exploring the lexical profile of young CLIL learners: towards an improvement in lexical use. *International Journal of Immersion and Content-based Learning*, 2(1), 53–73.
- Agustín Llach, M.P. (2015). The Effects of the CLIL Approach in Young Foreign Language Learners' Lexical Profiles. *Journal of Bilingual Education and Bilingualism*. doi: 10.1080/13670050.2015.1103208.
- Cabaleiro González, M. B. (2003). La escritura en L1 y L2: estudio empírico. *RESLA*, 16, 33–52.
- Canga Alonso, A. (2013). The receptive vocabulary of Spanish 6th grade primary school students in CLIL instruction: A preliminary study. *Latin American Journal of Content and Language Integrated Learning (LACLIL)*, 6(2), 22–41. doi: 10.5294/laclil.2013.6.2.2.
- Caspi, T. & Lowie, W. (2013). The Dynamics of L2 Vocabulary Development: A Case Study of Receptive and Productive Knowledge. *RBLA, Belo Horizonte*, 13(2), 437–462.
- Celaya, M. L. (2008). I study natus in English: Lexical transfer in CLIL and regular learners. In R. Monroy & A. Sánchez (Eds.), *25 Years of Applied Linguistics in Spain: Milestones and Challenges* (pp. 43–49). Murcia: Servicio de Publicaciones de la Universidad de Murcia.
- Celaya, M. L. & Ruiz de Zarobe, Y. (2008). CLIL, Age, and L1 influence. Paper presented at the *XXXII AEDEAN Conference*. University of the Balearic Islands, Palma de Mallorca, November, 13–15.
- Celaya, M. L. & Ruiz de Zarobe, Y. (2010). First languages and age in CLIL and non-CLIL contexts. *International CLIL Research Journal*, 1(3), 60–66.
- Celaya, M. L. & Torras, M. R. (2001). L1 influence and EFL vocabulary. Do children rely more on L1 than adult learners? In M. Falces Sierra, M. M. Díaz Dueñas & J. M. Pérez Fernández (Eds.), *Proceedings of the 25th AEDEAN Conference* (pp. 1–14). Granada: Universidad de Granada.
- Cobb, T. (n.d.). *Compleat Lexical Tutor*. Available at <http://www.lextutor.ca>.
- Council of Europe. (2001). *Common European Framework of Reference for Languages: learning, teaching, assessment*. Cambridge: Cambridge University Press.
- Cui, Y. (2009). The development of lexical organization in Chinese EFL learners at tertiary level. *Arizona Working Papers in SLA & Teaching*, 16, 57–73.
- Dalton-Puffer, C. (2008). Outcomes and processes in Content and Language Integrated Learning (CLIL): current research from Europe. In W. Delanoy & L. Volkman (Eds.), *Future Perspectives for English Language Teaching* (pp. 139–157). Heidelberg: Carl Winter.
- Dalton-Puffer, C. (2011). Content and language integrated learning: from practice to principles. *Annual Review of Applied Linguistics*, 31, 182–204.
- Dewaele, J. M. (1998). Lexical inventions: French interlanguage as L2 versus L3. *Applied Linguistics*, 19 (4), 471–490.
- Dewaele, J. M. (2001). Activation or Inhibition ? The Interaction of L1, L2 and L3 on the Language Mode Continuum. In J. Cenoz, B. Hufeisen & U. Jessner (Eds.), *Cross-linguistic Influence in Third Language Acquisition: Psycholinguistic Perspectives* (pp. 69–89). Clevedon: Multilingual Matters.
- Ecke, P. (2001). Lexical retrieval in a third language: evidence from errors and tip-of-the-tongue states. In J. Cenoz, B. Hufeisen & U. Jessner (Eds.), *Cross-linguistic Influence in Third Language Acquisition: Psycholinguistic Perspectives* (pp. 90–114). Clevedon: Multilingual Matters.
- Ellis, N. & Beaton, A. (1993). Psycholinguistic Determinants of Foreign Language Vocabulary Learning. *Language Learning*, 43(4), 559–617.
- Gabryś-Barker, D. (2006). The Interaction of Languages in the Lexical Search of Multilingual Language Users. In J. Arabski (Ed.), *Cross-linguistic Influences in the Second Language Lexicon* (pp. 144–166). Clevedon: Multilingual Matters.
- González Álvarez, E. (2004). *Interlanguage Lexical Innovation*. München: Lincom Europa.

- Jiang, N. (2000). Lexical representation and Development in a Second Language. *Applied Linguistics*, 21(1), 47–77.
- Jiménez Catalán, R. M. & Ruiz de Zarobe, Y. (2009). The receptive vocabulary of EFL learners in two instructional contexts: CLIL vs. non-CLIL instruction. In Y. Ruiz de Zarobe & R. M. Jiménez Catalán (Eds.), *Content and Language Integrated Learning. Evidence from Research in Europe* (pp. 81–92). Clevedon: Multilingual Matters.
- Jiménez Catalán, R. M., Ruiz de Zarobe, Y. & Cenoz, J. (2006). Vocabulary profiles. *Views*, 15(3), 23–27. Special issue: *Current Research on CLIC* (C. Dalton-Puffer & T. Nikula (Eds.)). Retrieved September 3, 2015 from https://anglistik.univie.ac.at/fileadmin/user_upload/dep_anglist/weitere_Uploads/Views/views15_3_clil_special.pdf.
- Lasagabaster, D. (2008). Foreign language competence in content and language integrated courses. *The Open Applied Linguistics Journal*, 1, 31–42.
- Laufer, B. (1990). Why are some words more difficult than others? Some intralexical factors that affect the learning of words. *IRAL*, 28, 293–307.
- Laufer, B. (1997). What's in a word that makes it hard or easy: some intralexical factors that affect the learning of words. In N. Schmitt & M. McCarthy (Eds.), *Vocabulary. Description, Acquisition and Pedagogy* (pp.140–155). Cambridge: Cambridge University Press.
- Li, J. & Schmitt, N. (2009). The acquisition of lexical phrases in academic writing: A longitudinal case study. *Journal of Second Language Writing*, 18, 85–102.
- Marsden, E. & David, A. (2008). Vocabulary Use during Conversation: A Cross-sectional Study of Development from Year 9 to Year 13 amongst Learners of Spanish and French. *Language Learning Journal*, 36(2), 181–198.
- Meara, P. (1996). The dimensions of lexical competence. In G. Brown, K. Malmkjaer & J. Williams (Eds.), *Performance and Competence in Second Language Acquisition* (pp. 33–53). Cambridge: Cambridge University Press.
- Meara, P. & Bell, H. (2001). P_Lex: A simple and effective way of describing the lexical characteristics of short L2 texts. *Prospect*, 16(3), 5–17.
- Milton, J. (2009). *Measuring second language vocabulary acquisition*. Bristol: Multilingual Matters.
- Moreno Espinosa, S. (2009). Young Learners' L2 Word Association Responses in Two Different Learning Contexts. In Y. Ruiz de Zarobe & R. M. Jiménez Catalán (Eds.), *Content and Language Integrated Learning. Evidence from Research in Europe* (pp. 93–111). Clevedon: Multilingual Matters.
- Muñoz, C. (2007). Age-related differences and second language learning practice. In R. DeKeyser (Ed.), *Practice in a Second Language. Perspectives from Applied Linguistics and Cognitive Psychology* (pp. 229–255). Cambridge: Cambridge University Press.
- Nation, P. (2001). *Learning vocabulary in another language*. Cambridge: Cambridge University Press.
- Nicolay, A.C. & Poncelet, M. (2013). Cognitive abilities underlying L2 vocabulary acquisition in an early L2-immersion educational context: A longitudinal study. *Journal of experimental child psychology*, 115, 655–671.
- Nikolov, M. (2014). What we have learnt from students' and teachers' feedback on diagnostic tests: Implications for classrooms. Plenary talk presented at the *International Conference on Child Language Acquisition*, Universidad del País Vasco, Spain, November, 16–17.
- Pfenninger, S. (2014). The misunderstood variable: Age effects as a function of type of instruction. *Studies in Second Language Learning and Teaching*, 4(3), 529–556.
- Poullisse, N. & Bongaerts, T. (1994). First language use in second language production. *Applied Linguistics*, 15, 36–57.
- Ringbom, H. (1998). Vocabulary frequencies in advanced learner English: A cross-linguistic approach. In S. Granger (Ed.), *Learner English on Computer* (pp. 41–52). London: Addison Wesley Longman.
- Ringbom, H. (2001). Lexical Transfer in L3 Production. In J. Cenoz, B. Hufeisen & U. Jessner (Eds.), *Cross-linguistic Influence in Third Language Acquisition: Psycholinguistic Perspectives* (pp. 59–68). Clevedon: Multilingual Matters.
- Robinson, P. & Ellis, N. C. (Eds.). (2008). *Handbook of cognitive linguistics and second language acquisition*. New York, NY / London: Routledge.

- Rokita, J. (2006). Code-mixing in Early L2 Lexical Acquisition. In J. Arabski (Ed.), *Cross-linguistic Influences in the Second Language Lexicon* (pp. 177–190). Clevedon: Multilingual Matters.
- Ruiz de Mendoza, F. J. (2013). Pedagogical grammar and meaning construction. Plenary Lecture presented at the *International Conference: Constructionist Approaches to Language Pedagogy CALP 2013*. Université Saint-Louis, Belgium, November, 8–9.
- Ruiz de Zarobe, Y. (2008). CLIL and foreign Language learning: A longitudinal study in the Basque Country. *International CLIL Research Journal*, 1(1), 60–73.
- Ruiz de Zarobe, Y. (2011). Which Language Competencies Benefit from CLIL? An Insight into Applied Linguistics Research. In Y. Ruiz de Zarobe, J. M. Sierra & F. Gallardo del Puerto (Eds.), *Content and Foreign Language Integrated Learning: Contributions to Multilingualism in European Contexts* (pp. 129–154). Bern: Peter Lang.
- Ruiz de Zarobe, Y. & Celaya, M. L. (2009). CLIL instruction and foreign language competence in two bilingual communities. Paper presented at the *XXVII International AESLA Conference*. Universidad de Castilla-La Mancha, Ciudad Real, March 26–28.
- Saji, N., Imai, M., Saalback, H., Zhang, Y., Shu, H. & Okada, H. (2011). Word learning does not end at fast mapping: Evolution of verb meanings through reorganization of an entire semantic domain. *Cognition*, 118, 45–61.
- Schmitt, N. (1998). Tracking the incremental acquisition of second language vocabulary: A longitudinal study. *Language Learning*, 48(2), 281–317.
- Schmitt, N., Schmitt, D. & Clapham, C. (2001). Developing and Exploring the Behaviour of two New Versions of the Vocabulary Levels Test. *Language Testing*, 18(1), 55–88.
- Schoonen, R., van Gelderen, A., De Glopper, K., Hulstijn, J., Simis, A., Snellings, P. & Stevenson, M. (2009). First language and second language writing: The role of linguistic knowledge, speed of processing and metacognitive knowledge. *Language Learning*, 53(1), 165–202.
- Singleton, D. (1999). *Exploring the Second Language Mental Lexicon*. Cambridge: Cambridge University Press.
- Stemach, G. & Williams, W. (1988). *WordExpress: The first 2500 words of spoken English*. Novato, CA: Academic Therapy Publications.
- Sylvén, L. K. (2010). *Teaching in English or English Teaching? On the effects of content and language integrated learning on Swedish learners' incidental vocabulary acquisition*. Gothenburg: Acta Universitatis Gothoburgensis.
- Terrazas Gallego, M. & Agustín-Llach, M. P. (2009). Exploring the increase of receptive vocabulary knowledge in the foreign language: A longitudinal Study. *International Journal of English Studies*, 9(1), 113–133.
- Várkuti, A. (2010). Linguistic Benefits of the CLIL Approach: Measuring Linguistic Competences. *International CLIL Research Journal*, 1(3), 67–79.
- Vermeer, A. (2004). The relation between lexical richness and vocabulary size in Dutch L1 and L2 children. In P. Bogaards & B. Laufer (Eds.), *Vocabulary in a second language: Selection, acquisition and testing* (pp. 173–190). Amsterdam / Philadelphia, PA: John Benjamins.
- Xanthou, M. (2011). Current trends in L2 vocabulary learning and instruction: Is CLIL the right approach? In A. Psaltou-Joycey & M. Mattheoudakis (Eds.), *14th International Conference on Applied Linguistics. Advances in Research on Language Acquisition and Teaching: Selected Papers* (pp. 459–471). Thessaloniki: GALA. Retrieved September 3, 2015 from <http://www.enl.auth.gr/gala/14th/Papers/English%20papers/Xanthou.pdf>.
- Zhang, X. & Lu, X. (2013). A Longitudinal Study of Receptive Vocabulary Breadth Knowledge Growth and Vocabulary Fluency Development. *Applied Linguistics*, 35(3), 283–304. doi: 10.1093/applin/ amt014.
- Zheng, Y. (2012). Exploring long term productive vocabulary development in an EFL context: The role of motivation. *System*, 40, 104–119.