

ANALYSIS OF THE EFFECT OF CLIL PROGRAMMES ON THE WRITTEN COMPETENCE OF SECONDARY EDUCATION STUDENTS

Ana Cristina Lahuerta Martínez
Universidad de Oviedo

RESUMEN

El presente estudio compara la competencia escrita de estudiantes de programas de Aprendizaje Integrado de Contenidos y Lenguas Extranjeras (AICLE) y aquellos que no siguen estos programas. Un total de 400 estudiantes de educación secundaria obligatoria procedentes de colegios públicos de Asturias participaron en el estudio. La muestra consistía en un grupo de estudiantes de programas AICLE de primero y cuarto de enseñanza secundaria obligatoria y otro grupo de estudiantes de primero y cuarto de enseñanza secundaria obligatoria no pertenecientes a estos programas. Los datos de este estudio proceden de una actividad de composición que los participantes realizaban como parte de sus tareas de clase. Las medidas utilizadas para medir la competencia escrita fueron: fluidez, precisión, complejidad gramatical y complejidad léxica. Los resultados muestran que los programas AICLE ejercen un efecto positivo en todos los aspectos de la lengua examinados.

PALABRAS CLAVE: escritura, fluidez, precisión, complejidad, programa AICLE.

ABSTRACT

The present study aims to analyse and compare the written competence of two groups of secondary education students: one group enrolled on a CLIL programme and another group enrolled on a non-CLIL programme, including grades in the analysis. Four hundred secondary education students from several state schools in Asturias participated in the study. They were in the first and the fourth year of compulsory secondary education. The data for the present study comes from a written composition activity, which was administered to participants in their own classroom. The writing measures used to analyse the written production of the learners were fluency, accuracy, grammatical complexity and lexical complexity. The results obtained showed that CLIL programs exert a positive influence on all the language aspects measured.

KEYWORDS: writing, fluency, accuracy, complexity, CLIL programmes.



1. INTRODUCTION

Content and Language Integrated Learning (CLIL) relates to the teaching of a content-based subject by means of a language which is not the mother tongue of the students in the classroom. It is a teaching approach in which an additional language is used for the teaching and learning of subjects with a dual focus on language and content (Heras and Lasagabaster 2015: 71). The CLIL approach has been praised on many different grounds (Coyle 2008; Coyle *et al.* 2010; Marsh 2008). CLIL is considered to be an alternative path to conventional English as a Foreign Language teaching. It is believed to foster implicit and incidental learning by focusing on meaning and communication, and as a result, it is also believed to improve overall language competence in the target language (Lasagabaster 2008; Heras and Lasagabaster 2015).

Written competence is a subset of learners' language competence, with an emphasis on writing-specific abilities such as the production of a variety of genres and rhetorical features, but also including language-specific abilities such as the use of a range of vocabulary and syntactic structures (Wolfe-Quintero, Inagaki and Kim 1998: 2). Following Larsen-Freeman (2006), we conceptualize language development as a complex dynamic process. Written competence as a subset of language competence is also complex and cannot be totally accounted for by performance in any one subsystem. In this study, written competence is characterised by three dimensions of language proficiency: fluency, accuracy and complexity. Research has shown that these three dimensions are robust indicators of a learner's written competence (Baba and Nitta 2014; Wolfe-Quintero *et al.* 1998). Wolfe-Quintero, Inagaki and Kim (1998) review several studies that have used measures of fluency, accuracy, grammatical complexity and lexical complexity to analyse written data in order to, for example, compare writers at different proficiency levels, or to examine the relationships between errors and holistic ratings of second language writers.

We follow this line of research and examine the written competence of first and fourth grade secondary education students enrolled on CLIL and non-CLIL programmes. This paper intends to shed some light on the question of the effectiveness of CLIL programmes focusing on one skill, writing.

2. A REVIEW OF THE LITERATURE

Research on the benefits of CLIL on written development is inconclusive. While some studies suggest the existence of limited progress regarding writing in CLIL classrooms, others show the benefits of CLIL on written competence. Among the former studies, Llinares and Whittaker (2006) found that their secondary Spanish CLIL participants learning social science through English hardly ever used resources such as modality or clause expansion through elaboration in their compositions. For their part, Pladevall-Ballester and Vallbona (2016) showed that in contexts of minimal and equal exposure, CLIL had no remarkable effects, although



significant intragroup progress was observed in the two contexts examined (Formal Instruction and Formal Instruction + CLIL).

Among the research works that show the benefits of CLIL on written competence, Lasagabaster (2008) examined foreign language written competence in CLIL contexts, taking on a holistic perspective. The sample used included a group made up of non-CLIL students in the fourth year of secondary education, another group composed of CLIL students in the fourth year of secondary education and a third group made up of CLIL students in the third year of secondary education. Competence in the foreign language was measured via four English tests corresponding to grammar, listening, speaking and writing (2008: 35). Results (2008: 36-38) revealed that the CLIL groups significantly outperformed their non-CLIL counterparts in every single test and in the overall English competence score. The third year CLIL group not only caught up with their fourth year non-CLIL counterparts and but also surpassed them in overall foreign language competence. His findings were in line with Navés and Victori (2010), who focused on the general language proficiency as well as writing skills of primary and secondary education CLIL and non-CLIL learners. Students' language proficiency was analysed by means of a listening test, a grammar test and a dictation in English. The writing test consisted of a composition, which was analysed for accuracy (error-free sentences), fluency (number of words), syntactic complexity (subordinate clauses) and lexical complexity (word variation) (2010: 34). As these authors explain (2010: 41-43), results of both studies showed that CLIL learners outperformed non-CLIL learners in most of the tests administered. The first study also found that seventh grade CLIL learners obtained results similar to those of non-CLIL learners one or two grades ahead for each of the measures analysed: dictation, reading comprehension, grammar and listening skills. In the second study, CLIL learners' writing at lower grades was observed to be as good as or even better than that of older learners a few grades ahead. From both studies the authors conclude (2010:47) that when learners are at grades 7 and 9 and have received CLIL instruction they achieve a level equivalent to or even higher than learners a couple of grades ahead in many of the domains of a language. All these findings were in line with those of Ruiz de Zarobe (2010), who focused on written skills using the same holistic methodology but a different design, which included controlling exposure to the target language. Her findings showed that 15 to 16-year-old secondary CLIL learners scored significantly higher than their non-CLIL counterparts in vocabulary, language use and mechanics. In a similar vein, Roquet (2011) used holistic and analytical complexity, accuracy and fluency measures to examine the effect of type of instruction and age over English as a Foreign Language (EFL) development. She also found greater benefits for CLIL secondary students in overall writing

The study by Jexenflicker and Dalton-Puffer (2010) used analytical complexity measures to examine specifically the effect of CLIL provision on written competence. They found (2010: 180) that CLIL students as compared to students who followed a traditional EFL curriculum showed a wider range not only of lexical but also morphosyntactic resources, which they deployed in more elaborate and more complex structures.



For their part, Lorenzo and Rodríguez (2014: 64) focussed on the evolution of syntax in CLIL writing. They analysed a corpus of historical narratives of subjects from the third year of secondary Education to the second year of post-compulsory secondary Education (Baccalaureate) from 4 secondary schools where a content and language integrated approach was set up. Results (2014: 68) showed that learners in the lowest grades in a CLIL setting produced an amalgamated language, characterized by a lack of dependent clauses, t-units and coordinate phrases. However, this language skill was consolidated in higher grades. As these authors (2014: 71) conclude, CLIL contexts in these grades and settings enhance the complex, discursive, academic functions, as expressed in narratives.

Among the research works that have shown benefits in writing, there are some longitudinal studies too. Whittaker *et al.* (2011) examined the linguistic resources used to create coherence and appropriate register in the CLIL students' written texts. Texts from history classes were collected annually over the four-year obligatory junior secondary education program from the same students (aged 12 to 16), in two state schools (2011: 348). Results (2011: 358) showed development in the control of textual resources, as well as some increase in nominal group complexity, over the four years.

Like Whittaker *et al.* (2011), Merisuo-Storm and Soininen's (2014: 75) longitudinal study suggests that CLIL settings provide suitable contexts in which to develop written discourse. They investigated the effects of CLIL on the development of children's literacy skills from the beginning of the first grade to the end of sixth grade. The test group studied, from the very beginning of the first grade, different school subjects in Finnish and in English. The students in the control group studied all school subjects in Finnish and started to learn English as a second language in third grade (2014: 71). The students' writing tasks consisted of writing from dictation, and writing a story. Results showed that after two study years, the reading and writing skills of the test group were significantly better than those of the control group. After four school years the children's creative writing skills had also benefited from bilingual teaching (2014: 78).

Gené *et al.* (2014) and Pérez-Vidal and Roquet (2015) used analytical complexity, accuracy and fluency measures in their longitudinal studies of the effect of CLIL on writing development. Gené *et al.* (2014) investigated the performance of CLIL and Formal Instruction learners in a written composition at four data collection times over three years. They concluded that the CLIL approach succeeded in developing written competence in secondary school students. Similar results were obtained by Pérez-Vidal and Roquet (2015), who examined the linguistic progress achieved over one academic year by CLIL secondary education learners, enrolled in an English-medium Science course in contrast with a Formal Instruction programme developed in the same school. Results obtained confirmed that larger relative gains were obtained by the Formal Instruction + CLIL programme on their writing ability, and particularly so their accuracy which showed higher relative gains.

Longitudinal studies by Knoch *et al.* (2014, 2015) examined ESL writing development in students who had spent some period of study abroad. They found significant writing development but limited to certain measures. Knoch *et al.* (2014)



examined students' ESL writing proficiency following a year's study in an Australian university. To this end, they used global writing scores, as well as measures of accuracy (error-free clauses and error-free T-units), fluency (number of words), grammatical complexity (words per T-unit, clauses per T-unit, words per clauses), and lexical complexity (percentage of academic word list, lexical sophistication and D-value, a measure of lexical richness which is derived by computing a set of type/token ratios for each text) (2014: 4-5). The results of the study showed that global scores of writing showed no change over time. The only significant improvement participants in the current study showed was in their fluency (measured via text length). That is, they could write longer texts in the time allowed. There were no observed gains in accuracy, syntactic and lexical complexity (2014: 8-10).

Knoch *et al.* (2015) examined undergraduate students' L2 (ESL) writing proficiency following a three-year degree study in an L2-medium university. The study used a test-retest design which required participants to write an argumentative essay on the same topic at the commencement and at the end of their degree program. A range of measures was used to assess writing, including global and discourse measures (accuracy, fluency, complexity). Accuracy was measured via the percentage of error-free T-units and clauses. Fluency was measured by counting the number of words for each essay, by the number of T-units and T-unit length, i.e., the average number of words per T-unit. Grammatical complexity was measured via the average numbers of words per clause, clauses per T-unit, and the ratio of dependent clauses to all clauses. For lexical complexity, three different measures were used which included percentage of words from the Academic Word List, lexical sophistication, and average word length (2015: 44). Consistent with Knoch *et al.* (2014) global scores of writing did not improve significantly over the three years of degree study. In terms of the discourse measures, also consistent with Knoch *et al.*, (2014), fluency (measured via word count) increased significantly over three years of degree study, suggesting that participants were able to produce more words within the same allotted time, whereas accuracy, grammatical and lexical complexity did not change over time (2015: 50).

Navés, Torras, and Celaya (2003) and Godfrey, Treacy, and Tarone (2014) also report longitudinal studies, but comparing the performance of different groups. Navés, Torras and Celaya (2003) analysed the development of the written production of six groups of primary and secondary education learners. Among the main findings of this study is that accuracy, fluency, syntactic and lexical complexity do not develop in tandem, but correlate differently depending on the learners' age group (2003: 123-124).

For their part, Godfrey and Treacy (2014) examined the writing of eight university learners of French—four during study abroad and four in on-campus courses—over the course of a semester. This study applied measures focused on the complexity, accuracy, fluency, and form-function relationships of writing samples collected at the beginning and end of the semester. The measure of fluency was the total number of words per essay. Accuracy was measured by counting the percentage of correct instances in which a student had to make a decision about gender. Lexical complexity was analysed using such measures as number of different words, type-



token ratio (TTR), and the index of Guiraud. Syntactic complexity was analysed with a clause/T-unit analysis (2014: 52-54).

Results (2014: 56) suggest that progress toward more advanced academic L2 writing occurred for both groups of students, although in different ways. Students in both groups improved their fluency in writing, as measured by length of their essays, but the domestic group seemed to increase essay length more than the study abroad group did. On a measure of accuracy, the study abroad group increased both their use of French gendered nouns and their accuracy in gender marking more than the on-campus group did. A T-unit analysis showed that, while both groups increased the syntactic complexity in their writing, the domestic group improved more than the study abroad group did. Both groups' use of linguistic forms and expressions to make supported claims and use appropriate discourse markers improved, while the on-campus group increased their hedging of such claims more than the study abroad group.

3. AIMS OF THE STUDY

As we have seen in the previous section, general results regarding the effect of CLIL on written competence seem to be positive although mixed results are also found in the literature. Moreover, studies on written competence in both CLIL and Formal Instruction contexts show significant differences among course levels and reveal improvement over time. However, some research works show that this development is limited to certain writing measures used to assess such competence. Against this mixed-findings backdrop, the present paper aims to examine the written competence of two groups of secondary education students: one enrolled on a CLIL programme and another group enrolled on a non-CLIL programme, including grade in the analysis. Written competence is characterised, as stated above by three dimensions of language proficiency: fluency, accuracy and complexity. We assume that foreign language writers will write more fluently, or write more in the same amount of time, write more accurately, or produce fewer errors in their writing, and write more grammatically and lexically complex sentences as they become more proficient. The following research questions are the focus of the study:

- Research Question 1: Is there a significant difference in written competence between the CLIL and the non-CLIL groups?
- Research Question 2: Is there a significant difference in written competence between the first and fourth year groups?
- Research Question 3: What will the relationship be between the measures in the areas of fluency, accuracy, lexical complexity and grammatical complexity and accuracy?



3.1. METHOD

3.1.1. *Participants*

The participants were 400 secondary education students enrolled in different state schools in Asturias, all of whom had started learning English at the age of five. The sample was divided into four groups: Two groups of students enrolled on a CLIL program, consisting of a first group made up of 100 students in the first year of compulsory education (CSE) and a second group made up of 100 students in the fourth year of CSE. Two groups of students who did not follow a CLIL programme, which consisted of a group of 100 students in the first year of CSE, and a group of 100 students in the fourth year of CSE.

Data collection took place at the end of the 2014-15 academic year. As we can see in Table 1, CLIL and non-CLIL students' average age was 13 years at grade 1 and 16 at grade 4. First and fourth-year CLIL students had received four hours a week of EFL instruction plus three hours a week of CLIL instruction in geography and history. For their part, the first and fourth grade non-CLIL students had received exclusively four hours a week of Formal Instruction in EFL.

TABLE I. DESIGN OF THE STUDY

PROGRAMME	GRADE	AGE	EFL HOURS/ ACADEMIC YEAR	CLIL HOURS/ ACADEMIC YEAR	TOTAL HOURS/ ACADEMIC YEAR
CLIL	First CSE	13	144 h.	108 h.	252 h.
	Fourth CSE	16	144 h.	108 h.	252 h.
NON-CLIL	First CSE	13	144 h.	-	144 h.
	Fourth CSE	16	144 h.	-	144 h.

Regarding the methodology used by the two instruction programmes, both contexts followed a communicative approach, but the Formal Instruction programme put a greater emphasis on language forms and grammar. Writing activities were part of the English as a Foreign Language curriculum in both programmes although students in the bilingual setting received writing instruction and writing assignments in their content subject classes in addition to their English classes.

In order to ensure that the non-CLIL group was comparable to the experimental group at the onset of the study, all participants answered a profile questionnaire that enabled us to rule out any important differences in extra-school exposure to the target language. By so doing, we managed to have two groups of learners whose performance in the domains examined did not differ statistically at the outset of the study. A chi-square test of independence was carried out to determine if the percentage of children attending extracurricular English classes differed by type of instructional setting. The results showed that the slight difference observed was not significant, $\chi^2(1, N=400) = 0.4233, p = 0.5516$.

In addition, to rule out possible differences between CLIL and non-CLIL students being due to greater exposure to the target language in favour of the former,



a standardized English Test was administered to the students prior to the experiment. This test was administered by the teachers in their own classes. This is a commercialised test designed by Oxford, which consists of a total of 200 items including listening comprehension, grammar and Use of English items. For our study, we selected those students that obtained a score between 90 and 104 in year 1 and those students with a score between 105 and 119 in year 4 (in the test, this corresponds to levels A1 and A2 respectively of the Common European Framework of Reference for Languages). We excluded all students with a score below 90 or higher than 119.

3.1.2. Procedure

For the present study, the data come from a written composition activity, which was administered to participants in their own classroom. For the written activity, students had to write on the topic ‘What are your favourite TV programmes?’ All the participants were given 30 minutes for the writing activity. In this way, both time and topic constraints were controlled in order to make results comparable (Wolfe-Quintero *et al.* 1998).

The writing measures used to analyse the written production of the four groups of learners have been classified into four areas following Wolfe-Quintero *et al.* (1998): fluency, accuracy, grammatical complexity and lexical complexity. To measure fluency we counted the total number of words. In addition, we used sentence length (total number of words divided by total number of sentences) as a measure of the fluency of writing. For accuracy, the measures used were error-free sentence ratio (total number of error-free sentences divided by total number of sentences) and errors per word ratio (total number of errors divided by total number of words). Regarding the grammatical complexity measures, we used the sentence complexity ratio (total number of clauses divided by total number of sentences). As an additional measure, we also counted the total number of connectors in the writing samples. Finally, for lexical complexity we used the ratio of the number of word types to the square root of two times the word tokens.

Regarding the scoring procedure of the writing test, following Bulté and Housen (2014), the rating score was a combination of rating on five 5-point scales: Content, Organization, Language Use, Vocabulary and Mechanics. A global grade was assigned to each essay. This grade resulted from adding up the points obtained in each scale and dividing them by five. All essays were analysed, annotated, and counted by two researchers. Inter-coder agreement initially varied from 85% to nearly 100%. All disagreements were discussed until agreement was reached.

4. RESULTS

A statistical analysis was carried out with the program R Development Core Team 2012, version 2.15. In what follows, we will see the results obtained for each research question formulated.



Research Question 1: Is there a significant difference in written competence between the CLIL programme and the non-CLIL programme groups?

Results revealed significant differences between the CLIL and non-CLIL programme groups. The CLIL programme group significantly outperformed the non-CLIL programme group in the general quality of the composition, and in three areas of writing (accuracy, fluency and grammatical and lexical complexity) examined. As we can see in Table II, the CLIL group obtained a higher mean for the composition score ($M=5.85$, Welch test, $p<0.001$). Moreover, CLIL programme students' fluency in writing as measured by the total number of words is significantly higher than that of non-CLIL programme students ($M=101.25$, Welch test, $p<0.001$). In addition, CLIL programme students also significantly outperform non-CLIL program students in sentence length measured as the total number of words per sentence ($M=20.65$, Welch test, $p<0.001$). CLIL programme students' accuracy in writing as measured by the percentage of error-free sentences and by errors per word ratio is significantly higher than that of non-CLIL programme students ($M=0.25$, Welch test, $p<0.001$; $M=0.07$, Welch test, $p<0.001$). CLIL programme students significantly outperform non-CLIL programme students in both grammatical complexity indicators, i.e., the total number of connectors ($M=3.68$, Welch test, $p<0.001$), and the sentence complexity ratio ($M=0.56$, Welch test, $p<0.001$). Finally, the CLIL group significantly outperforms the non-CLIL group in the lexical complexity measure ($M=1.36$, Student's t-test $p<0.001$).

TABLE II. CLIL AND NON-CLIL PROGRAMME STUDENTS' WRITTEN COMPETENCE

	GROUP	MEAN	SD	P
Composition score	CLIL	5.85	2.31	0.001
	Non- CLIL	3.68	2.73	
Total Number of words	CLIL	101.25	50.17	0.001
	Non- CLIL	64.38	60.59	
Sentence length	CLIL	20.65	6.38	0.001
	Non- CLIL	15.81	7.18	
Error-free sentence ratio	CLIL	0.25	0.15	0.001
	Non- CLIL	0.19	0.17	
Errors per word ratio	CLIL	0.07	0.06	0.001
	Non- CLIL	0.18	0.16	
Sentence complexity ratio	CLIL	0.56	0.15	0.001
	Non- CLIL	0.55	0.21	
Total Number of connectors	CLIL	3.68	2.38	0.001
	Non- CLIL	2.07		1.81
Word variation	CLIL	1.36	0.49	0.001
	Non- CLIL	1.18	0.44	



Research Question 2: Is there a significant difference in written competence between first and fourth grade students?

The second major goal of this study was to find the development of writing measures within the groups. The results show a trend of development in written competence from first to fourth grade in both groups. Table III indicates that, in the CLIL group, fourth graders significantly outperform first graders in the general quality of the composition ($M=6.27$, Welch test, $p<0.001$), and in the two measures of fluency in writing, i.e., total number of words ($M=115.50$, Student's t-test, $p<0.001$), and sentence length ($M=20.73$, Student's t-test, $p<0.05$). This trend is also observed in the two measures of accuracy in writing, i.e. percentage of error-free sentences ($M=0.32$, Student's t-test, $p<0.001$) and errors per word ratio ($M= 0.05$, Student's t-test, $p<0.001$). With respect to grammatical complexity, fourth graders outperform first graders in the total number of connectors ($M=4.11$, Student's t-test, $p<0.05$) as well as in the sentence complexity ratio ($M=0.49$, Student's t-test $p<0.001$). Finally, fourth graders also outperform first graders in lexical complexity ($M=1.60$, Student's t-test $p<0.001$)

In the non-CLIL group, we observe that fourth graders significantly outperform first graders in the general quality of the composition ($M=4.69$, Welch test, $p<0.001$), in the two measures of fluency in writing, i.e., total number of words ($M=93.01$, Student's t-test, $p<0.001$), and sentence length ($M=18.30$, Student's t-test, $p<0.001$). Fourth graders also outperform first grade students in accuracy in writing as measured by the percentage of error-free sentences ($M=0.22$, Student's t-test, $p<0.05$) and as measured by errors per word ratio ($M= 0.12$, Student's t-test, $p<0.001$), and in lexical complexity ($M=1.30$, Student's t-test $p<0.001$). Regarding grammatical complexity, fourth graders outperform first graders in the total number of connectors ($M=2.68$, Student's t-test, $p<0.001$) and in the sentence complexity ratio ($M=0.44$, Student's t-test $p<0.001$).

TABLE III. CLIL AND NON-CLIL PROGRAMME FIRST AND FOURTH GRADE STUDENTS' WRITTEN COMPETENCE

	YEAR	MEAN	SD	P
Composition score	CLIL			
	Fourth	6.27	1.88	0.01
	First		5.50	2.63
	Non- CLIL			
	Fourth	4.69	2.66	0.001
	First	2.75	2.49	
Total Num. words	CLIL			
	Fourth	115.40	51.63	0.001
	First		88.37	46.89
	Non- CLIL			
	Fourth	93.01	67.38	0.001
	First	88.36	39.37	

	YEAR	MEAN	SD	P
Sentence length	CLIL			
	Fourth	20.73	6.75	0.03
	First	18.71	5.93	
	Non- CLIL			
	Fourth	18.30	6.55	0.001
	First	12.78	5.41	
Error-free sent. ratio	CLIL			
	Fourth	0.32	0.27	0.001
	First	0.27	0.32	
	Non- CLIL			
	Fourth	0.22	0.18	0.03
	First	0.14	0.16	
Errors per word ratio	CLIL			
	Fourth	0.05	0.05	0.001
	First	0.07	0.06	
	Non- CLIL			
	Fourth	0.12	0.10	0.001
	First	0.23	0.18	
Sentence comp. ratio	CLIL			
	Fourth	0.49	0.14	0.001
	First	0.34	0.16	
	Non- CLIL			
	Fourth	0.44	0.17	0.001
	First	0.32	0.24	
Num. connectors	CLIL			
	Fourth	4.11	2.19	0.04
	First	3.33	2.54	
	Non- CLIL			
	Fourth	2.68	1.83	0.001
	First	1.31	1.31	
Word variation	CLIL			
	Fourth	1.60	0.47	0.001
	First	1.32	0.48	
	Non- CLIL			
	Fourth	1.30	0.29	0.001
	First	1.08	0.36	



We also compared the CLIL first grade students and the first and fourth grade non-CLIL students. We found that the students enrolled on first grade CLIL programmes scored higher than the non-CLIL fourth graders in most measures. CLIL first graders significantly outperform non-CLIL first and fourth graders in the general quality of the composition ($M=5.50$, Kruskal-Wallis test, $p<0.001$), in fluency in writing as measured by sentence length ($M=18.71$, Kruskal-Wallis test, $p<0.001$), in accuracy in writing both as measured by the percentage of error-free sentences ($M=0.27$, Kruskal-Wallis test, $p<0.001$) and as measured by errors per word ratio ($M=0.07$, Kruskal-Wallis test, $p<0.001$), and in lexical complexity ($M=1.32$, Kruskal-Wallis test, $p<0.001$). With respect to grammatical complexity CLIL first graders significantly outperform non-CLIL first and fourth graders in the total number of connectors ($M=3.33$, Kruskal-Wallis test, $p<0.001$) and outscore first non-CLIL graders in sentence complexity ratio ($M=0.34$, Kruskal-Wallis test, $p<0.001$).

Research Question 3: What will the relationship be between the measures in the areas of fluency, accuracy, lexical complexity and grammatical complexity and accuracy?

A further aim of this study was to find out how the different measures of writing ability correlated with each other. All the variables turned out to be associated significantly with the global score, except for sentence length in the CLIL group (see Table IV below). The relationship between total number of words, sentence length, error-free sentence ratio, sentence complexity ratio, number of connectors and lexical complexity with the global composition score was positive (when one increases, the score increases) in both groups. On the other hand, the relationship between errors per word ratio with the composition score is negative (when one of them increases the score decreases), in both groups as well.

TABLE IV. CORRELATION BETWEEN GLOBAL SCORES AND WRITING MEASURES

	CLIL		Non-CLIL	
	<i>p</i>	p value	<i>p</i>	p value
Composition score –Number of words	0.72	<0.001	0.74	<0.001
Composition score –Sentence length	0.12	0.1026	0.64	<0.001
Composition score –Error-free sentence ratio	0.69	<0.001	0.58	<0.001
Composition score –Errors per word ratio	-0.76	<0.001	-0.68	<0.001
Composition score –Sentence complexity ratio	0.27	<0.001	0.54	<0.001
Composition score –Number of connectors	0.58	<0.001	0.76	<0.001
Composition score –Word variation	0.61	<0.001	0.59	<0.001

With respect to the measures of written competence, we obtain significant relationships between all the pairs studied. The relationships were positive except for total number of words and errors per word ratio, sentence length and errors per word ratio, errors per word ratio and lexical complexity, errors per word ratio and



sentence complexity ratio. These were negative (when one of them increases the other decreases).

Fluency-Accuracy: We find a significant correlation between total number of words and error-free sentence ratio ($p=0.44$, $p<0.001$) and between total number of words and errors per word ratio ($p=-0.60$, $p<0.001$). Sentence length is significantly correlated with error-free-sentence ratio ($p=0.16$, $p<0.001$) and with errors per word ratio ($p=-0.55$, $p<0.001$). In other words, the more fluent learners are, the more accurate their writing may also be.

Accuracy and Lexical Complexity: Error-free sentences and errors per word ratio significantly correlate with lexical complexity ($p=0.45$, $p<0.001$; $p=-0.51$, $p<0.001$). We can then say that accurate writers write more lexically complex texts.

Fluency and Lexical Complexity: Total number of words and sentence length significantly correlate with lexical complexity ($p=0.56$, $p<0.001$; $p=0.31$, $p<0.001$). The longer the compositions, the more lexically complex they are.

Accuracy and grammatical complexity: Error-free sentence ratio significantly correlates with sentence complexity ratio ($p=0.15$, $p<0.002$) and total number of connectors ($p=0.43$, $p<0.001$). In other words, the larger the number of sentences without errors the larger the number of clauses and the larger the number of connectors. Errors per word ratio significantly correlates with sentence complexity ratio ($p=-0.52$, $p<0.001$) and total number of connectors ($p=-0.52$, $p<0.001$). The larger the number of errors, the less grammatically complex the compositions were.

Fluency and grammatical complexity: Total number of words significantly correlates with sentence complexity ratio ($p=0.47$, $p<0.001$) and total number of connectors ($p=0.70$, $p<0.001$). Sentence length significantly correlates with sentence complexity ratio ($p=0.75$, $p<0.001$) and with total number of connectors ($p=0.37$, $p<0.001$). The longer the compositions, the more grammatically complex they were.

Lexical complexity and grammatical complexity: Lexical complexity significantly correlates with sentence complexity ratio ($p=0.38$, $p<0.001$) and total number of connectors ($p=0.54$, $p<0.001$). The more lexically complex compositions have larger the number of clauses and larger number of connectors.

5. DISCUSSION

In this study, we investigated and compared the written competence of a group of students enrolled on a CLIL program and another group enrolled on a non-CLIL program. What follows is a discussion of the main results.

1. We can affirm that there is a significant difference in written competence between the CLIL and the non-CLIL programme groups. The CLIL group outperforms the non-CLIL group in the general quality of the compositions, in all the fluency, accuracy, grammatical and lexical complexity measures of writing used. This finding supports previous results (e.g. Jexenflicker and Dalton-Puffer, 2010; Lasagabaster, 2008; Lorenzo and Ridríguez, 2014; Merisuo-Storm and Soininen,



2014; Navés and Victori, 2010; Ruiz de Zarobe, 2010) and points to the effectiveness of bilingual settings to develop written competence.

2. The comparison between first and fourth graders allows us to identify how the four indicators of writing proficiency (fluency, accuracy, grammatical and lexical complexity) develop within a school setting. We observe that both the CLIL and the non-CLIL program fourth grade students outperform first graders in the general quality of the compositions as well as in the fluency, accuracy, grammatical and lexical complexity measures of writing used. These results are in line with those of Knoch *et al.* (2014, 2015) and Godfrey *et al.* (2014) that also show this tendency towards improvement in written competence. The results obtained seem to confirm the assumption of much second language writing research that fluency, accuracy, syntactic and lexical complexity progress in parallel (Wolfe-Quintero 1998). We observe a development for each of these measures, with the elder groups of learners being better in fluency and accuracy indicators, in grammatical and lexical measures.

Moreover, not only did the CLIL programme first graders outperform non-CLIL programme first graders, but the former even outscored non-CLIL fourth graders in the general quality of the composition, in sentence length, in the percentage of error-free sentences, in errors per word ratio, in the total number of connectors and in lexical complexity. These results are in line with those of Lasagabaster (2008) and Navés, and Victori (2010) that seem to indicate that the positive effects of CLIL education are evident even when the bilingual programme students are younger than the control students are.

3. With respect to the correlations between the writing measures used, all the variables turned out to be associated significantly with the general quality of the composition, in both groups, except for sentence length in the bilingual group. With respect to the measures of written competence, we obtain significant relationships between all the pairs studied, which supports previous studies (e.g. Navés *et al.* 2003). From the results obtained, we can say that accuracy induces longer compositions, more lexically complex with a larger number of clauses and connectors. Longer compositions are also more lexically and grammatically complex compositions. More lexically complex compositions are also more grammatically complex.

6. CONCLUSIONS

We can conclude that bilingual programmes exert a positive influence on all the language aspects measured. There are two main reasons that may explain why these programmes offer sound benefits in written competence to students. On the one hand, students in CLIL programmes are more frequently exposed to the English language. On the other hand, bilingual settings, which involve integrating both content and language goals, seem to provide suitable contexts in which to develop written discourse. CLIL programmes share many aspects of Communicative Language Teaching, while emphasising academic content as the substance of the



communication. This is supposed to make this communication more relevant and purposeful, which may offer the necessary conditions for effective learning to take place and for written competence to develop.

A trend of development in written competence from first to fourth grade is observed in both groups indicating that the measures of fluency, accuracy, grammatical and lexical complexity progress at the same rate and they are significantly correlated.

Finally, we acknowledge some limitations of the present study. Although the groups of participants selected for the study were as homogeneous as possible, some variables such as out-of-school exposure to English or socio-cultural family background could not be controlled. New studies will have to be carried out in the future taking these variables into account, so as to confirm the results obtained in the present study.

RECIBIDO: junio de 2016; ACEPTADO: septiembre de 2016.

REFERENCES

- BABBA, Kyoko and NITTA, Ryo (2014): «Phase transitions in development of writing fluency from a complex dynamic systems perspective», *Language Learning* 64: 1-35.
- BULTÉ, B. and HOUSEN, A. (2014): «Conceptualizing and measuring short-term changes in L2 writing complexity», *Journal of Second Language Writing* 26: 42–65.
- COYLE, Do (2008): «CLIL—A Pedagogical Approach from the European Perspective», in N. Van Deusen-Sholl and N.H. Hornberger (eds.), *Encyclopedia of Language and Education. Second and Foreign Language Education*. New York: Springer Science+Business Media LLC, 97-111.
- COYLE, Do, HOOD, P., and MARSH, David (2010): *CLIL: Content and Language Integrated Learning*. Cambridge: Cambridge University Press.
- GODFREY, Lee Ann and TREACY, Corbin (2014): «Change in French second language writing in study abroad and domestic contexts», *Foreign Language Annals* 47(1): 48-65.
- GENÉ-GIL, María, JUAN-GARAU, María and SALAZAR-NOGUERA, Joana (2014): «Writing development under CLIL provision», in María Juan Garau and Joana Salazar (eds.), *Content-based Language Learning in Multilingual Educational Environments*. Heidelberg New York Dordrecht London: Springer, 139-161.
- HERAS, Arantxa and LASAGABASTER, David (2015): «The impact of CLIL on affective factors and vocabulary learning», *Language Teaching Research* 19(1): 70-88.
- JEXENFLICKER, Silvia and DALTON-PUFFER, Christiane (2010): «The CLIL Differential», in Christiane Dalton-Puffer, Tarja Nikula and Ute Smit (eds.), *Language Use and Language Learning in CLIL Classrooms*. The Netherlands: John Benjamins Publishing, 169-190.



- KNOCH, Ute, ROUHSAD, Amir and STORCH, Neomy (2014): «Does the writing of undergraduate ESL students develop after one year of study in an English-medium university? », *Assessing Writing* 21: 1-17.
- KNOCH, Ute, ROUHSAD, Amir, PING OON, Sue and STORCH, Neomy (2015): «What happens to ESL students' writing after three years of study at an English medium university? », *Journal of Second Language Writing* 28: 39-52.
- LARSEN-FREEMAN, Diane (2006): «The emergence of complexity, fluency, and accuracy in the oral and written production of five Chinese learners of English», *Applied Linguistics* 27(4): 590-619.
- LASAGABASTER, David (2008): «Foreign language competence in content and language integrated courses», *The Open Applied Linguistics Journal* 1: 31-42
- LEKI, Ilona (2000): «Writing, literacy, and applied linguistics», *Annual Review of Applied Linguistics* 20: 99-115.
- LLINARES, Ana and WHITTAKER, Rachel (2006): «Linguistic analysis of secondary school students' oral and written production in CLIL contexts: studying social science in English», *Vienna English Working Papers* (Special Issue: Current Research on CLIL) 15(3): 28-32.
- LORENZO, Francisco and RODRÍGUEZ, Leticia (2014): «Onset and expansion of L2 cognitive academic language proficiency in bilingual settings: CALP in CLIL», *System* 47: 64-72.
- MARSH, David (2008): «Language Awareness and CLIL», in Jasone Cenoz and Nancy Hornberger (eds.), *Encyclopedia of Language and Education. Knowledge about Language*. New York: Springer Science+Business Media LLC, 233-246.
- MERISUO-STORM, T. (2007): «Pupils' attitudes towards foreign-language learning and the development of literacy skills in bilingual education», *Teaching Teacher Education* 23: 226-235.
- NAVÉS, Teresa and VICTORI, Mia (2010): «CLIL in Catalonia: An Overview of Research Studies», in David Lasagabaster and Yolanda Ruiz de Zarobe (eds.), *CLIL in Spain: Implementation, Results and Teacher Training*. Newcastle: Cambridge Scholars Publishing, 30-54.
- NAVÉS, Teresa, TORRAS, María Rosa and CELAYA, María Luz (2003): «Long-term effects of an earlier start. An analysis of EFL written production», *EUROSLA Yearbook* 3: 103-129.
- PÉREZ-VIDAL, Carmen and ROQUET, Helena (2015): «The linguistic impact of a CLIL Science programme: An analysis measuring relative gains», *System* 54: 80-90.
- PLADEVALL-BALLESTER, Elisabet and VALLBONA, Anna (2016): «CLIL in minimal input contexts: A longitudinal study of primary school learners' receptive skills», *System* 58: 37-48.
- ROQUET, Helena (2011): A study of the acquisition of English as a Foreign Language: integrating content and language in mainstream education in Barcelona. Unpublished PhD diss., Universitat Pompeu Fabra.
- RUIZ DE ZAROBÉ, Yolanda (2010): «Written production and CLIL: an empirical study», in Catherine Dalton Puffer, T. Nikula, and U. Smit (eds.), *Language Use in CLIL. Berlin: John Benjamin*, 191-210.
- WHITTAKER, Rachel, LLINARES, Ana and MCCABE, Anne (2011): «Written discourse development in CLIL at secondary school», *Language Teaching Research* 15(3): 343-362.
- WOLFE-QUINTERO, Kate, INAGAKI, Shunji and KIM, Hae-Young (1998): *Second Language Development in Writing: Measures of Fluency, Accuracy, and Complexity*. Hawai'i: University of Hawai'i at Manoa.

