International Journal of Bilingual Education and Bilingualism

ISSN: 1367-0050 (Print) 1747-7522 (Online) Journal homepage: https://www.tandfonline.com/loi/rbeb20

Analysing students' content-learning in science in CLIL vs. non-CLIL programmes: empirical evidence from Spain

Javier Fernández-Sanjurjo, Alberto Fernández-Costales & José Miguel Arias Blanco

To cite this article: Javier Fernández-Sanjurjo, Alberto Fernández-Costales & José Miguel Arias Blanco (2019) Analysing students' content-learning in science in CLIL vs. non-CLIL programmes: empirical evidence from Spain, International Journal of Bilingual Education and Bilingualism, 22:6, 661-674, DOI: 10.1080/13670050.2017.1294142

To link to this article: https://doi.org/10.1080/13670050.2017.1294142

INTERNATIONAL JOURNAL OF BILINGUAL EDUCATION AND BILINGUALISM
2019, VOL. 22, NO. 6, 661–674
https://doi.org/10.1080/13670050.2017.1294142

Analysing students' content-learning in science in CLIL vs. non-CLIL programmes: empirical evidence from Spain

Javier Fernández-Sanjurjo, Alberto Fernández-Costales and José Miguel Arias Blanco

Department of Educational Sciences, University of Oviedo, Oviedo, Spain

ABSTRACT

This paper investigates students' performance in content-subjects in the framework of CLIL programmes in Spain. So far, CLIL research has focused primarily on language attainment in the L2 and the L1, but students' achievements as regards content-subjects have been largely ignored. Competence in Science in the L1 is analysed by comparing pupils enrolled in mainstream schools with students in the so-called 'bilingual streams' offering CLIL-based approaches. The main objective is to assess if students learning Science through the L2 (English) outperform their counterparts studying in the L1 (Spanish) as regards content acquisition. The paper analyses a sample of 709 6th grade Primary Education students enrolled in public schools in the Principality of Asturias (Spain). A test to assess students' knowledge in Science and a context questionnaire (measuring participants' socio-economic status) were designed. The main finding is that students learning contents in their L1 perform slightly better than those studying Science through the L2. Furthermore, participants with lower socio-economic status obtain lower scores than those coming from more privileged backgrounds.

ARTICLE HISTORY
Received 29 September 2016
Accepted 7 February 2017

KEYWORDS
Content and language integrated learning (CLIL); science; subject-matter learning; primary education; socio-economic status (SES)

Introduction

The implementation of Content and Language Integrated Learning (CLIL) in Europe has been a major trend since the concept was coined in the 1990s and promoted by European institutions (Cenoz 2015; Eurydice 2006; Marsh 2002; Roquet and Pérez-Vidal 2015). Aiming to promote citizens' command of languages, CLIL has emerged as a 'timely solution to European plurilingual education' (Pérez-Cañado 2012, 315). In particular in Spain, CLIL schools – especially in Primary Education – have proliferated in the last two decades on the basis of the language, cultural, and cognitive benefits reported by research (Casal and Moore 2009; Cenoz 2015; Coyle 2002; Dafouz and Guerrini 2009; Escobar Urmeneta 2001; Escobar Urmeneta and Sánchez Sola 2009; Gajo 2007; Grandinetti, Langellotti, and Ting 2013; Halbach 2008; Lasagabaster and Ruiz de Zarobe 2010; Madrid and Hughes 2011).

More often than not, the implementation of CLIL follows reactive reasons (contexts where there is poor competence in the foreign language) and proactive responses (promoting situations that reinforce plurilingualism) as pinpointed by Coyle, Hood, and Marsh (2010, 7). The spread of CLIL in Spain is no surprise if we account for two elements: firstly, Spain is a multilingual territory with bilingual regions where CLIL approaches may be particularly suitable for the promotion of trilingualism (e.g. the Basque Country – see Cenoz 2009; Lasagabaster 2008 – and Catalonia – see Navés and Victori 2010; Pérez-Vidal and Juan-Garau 2010); secondly, there is a social need to improve competence in foreign languages, after being identified as one of the weaknesses in the Spanish education system: according to the results of the *Eurobarometre 386* (European Commission 2012a), or

CONTACT Alberto Fernández-Costales fernandezcalberto@uniovi.es

Eurydice's key data on teaching languages at school (Eurydice 2012), Spain consistently lags behind other European member states in the command of foreign languages.

In this framework, CLIL provisions have been implemented with multifaceted and heterogeneous approaches throughout the country in order to meet the demands of the new global society (Fernández Fontecha 2009; Fortanet Gómez and Ruiz-Garrido 2009; Lasagabaster and Ruiz de Zarobe 2010, 11; Pérez-Cañado 2012) and face the delivery gap between foreign language education and outcomes in terms of learner performance (Marsh 2002, 9). Nowadays, all European countries – except Denmark, Greece, Iceland, and Turkey – offer CLIL provisions (Eurydice 2012, 39). CLIL is an extended practice in Spain, which outnumbers most European countries regarding the number of schools offering CLIL provisions (Eurydice 2012, 155); in 'bilingual streams' in Spain, students are normally taught 5 h of English per week (instead of the 3 h of regular groups) besides taking (at least) 2 subjects taught through English (Science, Mathematics, Arts, Physical Education, or Music).

The relevance and social impact of CLIL has drawn international attention (Dalton-Puffer and Nikula 2006; Ruiz de Zarobe and Cenoz 2015), and research has been devoted to analyse students' language competence in the L2, the transfer to the L1, and other language-related issues such as code switching (Cenoz and Ruiz de Zarobe 2015; Dalrymple-Smith, Karagiannakis, and Papadopoulos 2012; Dalton-Puffer 2008; Escobar Urmeneta 2001; Escobar Urmeneta and Sánchez Sola 2009; Gajo 2007; Grandinetti, Langellotti, and Ting 2013; Ianos et al. 2016; Lasagabaster 2008; Ruiz de Zarobe 2008). So far, results have confirmed that CLIL effectively benefits the development of language competence. However, further research is needed, as some areas have been neglected, and fundamental issues about the effectiveness of CLIL have been ignored (Alejo and Piquer-Píriz 2016; Cenoz 2015; Coyle, Hood, and Marsh 2010, 149; Lasagabaster and Doiz 2016; Pérez-Cañado 2012). In particular, we feel that there is a dearth of research as regards the investigation of the effectiveness of CLIL in the development of contents in Primary Education, as little is known about subject-matter learning in CLIL (Cenoz, Genesee, and Gorter 2014; Lasagabaster and Ruiz de Zarobe 2010). Indeed, as claimed by Paran (2013, 323), 'CLIL research is conducted by language educators rather than subject specialists, and therefore focuses almost exclusively on language, with content knowledge rarely examined or measured'.

We believe this study makes a contribution to CLIL research, and provides new insights about content learning in Primary Education. Investigation in subject-learning in CLIL is scarce (Cenoz, Genesee, and Gorter 2014) and clearly deserves further attention; in addition, there is a shortage of research in CLIL in monolingual contexts in Spain (Fernández Fontecha 2009, 15).

Research in content learning in CLIL

CLIL has received considerable attention in the last decade due to the demand for increased language command in Europe. Although CLIL has its roots in Canadian immersion and bilingual programmes in North America, it also suits the specific features and needs of the European setting (Lasagabaster 2008; Lasagabaster and Doiz 2016; Lorenzo 2007; Pérez-Cañado 2012).

Extensive research has been conducted in the last decade, including in the relation between CLIL, foreign language teaching, content-based instruction and immersion programmes in North America (Cenoz 2013, 2015; Lasagabaster and Sierra 2010; Pérez-Cañado 2012; Ruiz de Zarobe 2008; Ruiz de Zarobe and Cenoz 2015), and pedagogy, teacher training and effective teaching performance (de Graaff et al. 2007; Escobar Urmeneta 2013; Llinares 2015; Pavón Vázquez and Rubio 2010). However, it can be argued that previous research was focused primarily on the language impact of CLIL on language performance (Admiraal, Westhoff, and de Bot 2006; Dalton-Puffer 2008; Lasagabaster and Ruiz de Zarobe 2010; Roquet and Pérez-Vidal 2015).

In the case of Primary Education, results show that CLIL students develop better proficiency in the L2 than those in monolingual educational settings (Pérez-Vidal 2011; Ruiz de Zarobe 2008; Ruiz de Zarobe and Jiménez Catalán 2009). In particular, most researchers agree when identifying listening

skills and vocabulary acquisition as the principal areas where language competence is favoured (Dalton-Puffer 2007; Jiménez Catalán and Ruiz de Zarobe 2007; Navés and Victori 2010).

Although the positive effects of CLIL are widely accepted, opposing views are also found on the possible drawbacks and the need for more solid research designs (Bruton 2011, 2013, 2015), suggesting that the initial enthusiasm for CLIL should not neglect challenges of this approach (Paran 2013; Pérez-Cañado 2016). Some scholars have also noted the ambiguity of CLIL and demand more critical research beyond analysing language gains (Cenoz 2013; Cenoz, Genesee, and Gorter 2014). In this vein, recent studies have called for more attention on the social milieu of CLIL (Alejo and Piquer-Píriz 2016) and the socio-economic status (SES) of the parents of CLIL students (Pérez-Cañado 2012).

Furthermore, with regards to analysing content learning, Cenoz, Genesee, and Gorter (2014, 10) claim that 'even were one to accept that there is a greater focus on language than on academic achievement in Canadian immersion research, the same can be said of research on CLIL where research on content is extremely limited', an argument seconded by Pérez-Cañado (2012, 315), who posits 'there is still a well-documented paucity of research in this area'. In our view, this paucity of research is visible in the case of content learning and CLIL.

Although this paper focuses on content-acquisition in non-linguistic subjects, it should be noted that the expansion of academic language is fundamental, as subject-specific literacies are related to student performance: the capacity of young learners to process academic language in content-subjects does influence the learning outcomes in bilingual provisions (Mohan, Leung, and Slater 2010). The study by Meyer et al. (2015) underlines that learning contents through an additional language is linked to subject-specific literacies and discursive features. In fact, this research concludes that the academic language of content-subjects should be a priority for teachers (who, more often than not, consider specific vocabulary as non-relevant). Cummins' (2008) classic distinction between Cognitive Academic Language Proficiency (CALP) and Basic Interpersonal Communicative Skills (BICS) still deserves consideration in the case of CLIL and bilingual provisions.

Results in the field of students' performance in subject-knowledge generally report positive outputs. In the case of North American bilingual programmes, it is commonly accepted that students learning through an additional language assimilate contents at the same pace as their counterparts in monolingual systems (see Pérez-Cañado 2012). Looking into European contexts, this tendency is backed by studies analysing European schools (Baetens Beardsmore 1990; Baetens Beardsmore and Swain 1985), where positive outcomes are identified in terms of L1 and L2 acquisition, meta-linguistic skills, and language awareness. Notwithstanding the value of the findings in European schools, it should be noted that the nature of these educational establishments differs from regular bilingual provisions (i.e. public schools) in Europe.

Focusing on CLIL schools, Ullmann (1999) explored the performance of students in the United Kingdom and found that pupils assimilating subject-contents through French showed enhanced subject-matter learning among other language-related attainments. Similar results are reported by Wode (1999), whose study with students of Secondary Education in Germany concludes that pupils in bilingual provisions perform better in Geography and History than those studying in German; Jäppinen (2005) analysed the thinking processes of 600 young learners in a longitudinal study in Finland and concluded that – indirectly – content-subject learning might be promoted by CLIL as a result of the stimulation of cognition processes. Similarly, Bergroth (2006) analysed CLIL programmes in Sweden to conclude that students learning Mathematics in Swedish (L2) and English (L3) did not have lower results than pupils studying through Finnish (L1) when finishing Secondary Education.

The general tendency observed in the aforementioned studies is supported by research conducted in the Netherlands by Admiraal, Westhoff, and de Bot (2006). Their longitudinal study takes a sample of 1305 students of Secondary Education enrolled in CLIL programmes with English as the tuition language. Among the main findings, the research team underlined that no negative impact is observed in content-knowledge in students taking subjects in the L2; however, we need to note that the sample for content-subjects (History and Geography) was limited to 57 participants.

A neutral situation regarding gains in subject-content is also reported by Stehler (2006), who studied a heterogeneous sample of CLIL students enrolled in several grades in public and private establishments in Switzerland and whose second languages were French and German.

Two other studies in Europe found significant differences when analysing content-learning in CLIL: Xanthou (2011) conducted 2 small-scale experiments with Science students of Primary Education in Cyprus, and concluded that assimilating contents through English was beneficial for participants. The longitudinal study by Serra (2007) with students of Primary Education (grades 1–6) in three public schools in Switzerland reported better results in Mathematics in the CLIL group. For the purpose of our study, it is worth mentioning that Serra's research focuses on CLIL classrooms where code-switching was a regular practice and that students learned half of the subjects in the L1 and half in the L2.

In the Spanish context, extensive research has been published in the last 5 years, although there is a lack of studies addressing subject-content in CLIL specifically. However, within the overall agreement on the language and cognitive benefits of CLIL, the few studies available report that no detrimental effects are observed when analysing the content knowledge of students in bilingual streams (Alonso, Grisaleña, and Campo 2008). As regards Primary Education, Halbach (2008) identifies two main challenges that might prevent an optimal output in learners' academic achievement: the lack of a specific methodology for teaching contents and language, and the syllabus design in bilingual streams.

We suggest that this paper fills a gap in the research literature, the first study with a representative sample of participants to focus on the acquisition of contents in Science in Primary Education (most studies investigated Mathematics and/or Secondary Education). In addition, we take into account the socio-economic status of students, aiming to assess the possible impact on performance, as this variable has proven to be consistently influential (Alejo and Piquer-Píriz 2016; Alonso, Grisaleña, and Campo 2008; Lorenzo, Casal, and Moore 2010; Pérez-Cañado 2012). Following Roquet and Pérez-Vidal (2015), the study provides evidence from real CLIL programmes that are not pilot experiences, and where learners have not been selected by the schools to access bilingual education.

The present study

The current research was conducted in the Principality of Asturias, a region with a similar curriculum to most regions in Spain. Participating schools were public establishments offering bilingual streams[1] (where students learn Science through English, the L2) and mainstreams schools (where all subjects are taught through Spanish, the L1).

The research hypothesis is that a decrease in students' academic performance is not expected in the case of CLIL schools. The rationale for this hypothesis is that, although students must develop key competences through the L2, the curriculum is similar in CLIL and non-CLIL establishments. In addition, so far, no negative effects of CLIL approaches on content-subjects have been documented to refute this hypothesis. The research objectives are the following:

(1) Assess the level of CLIL and non-CLIL students in Natural Science in the L1 when they finish Primary Education (6th grade).
(2) Analyse if there are statistically significant differences between students in CLIL and non-CLIL schools as regards Science knowledge. If differences are found, determine if a relation can be established between students' performance and the type of school (CLIL vs. non-CLIL) or socio-economic status.

Methodology

Due to the characteristics of the sample and the nature of the study, the research relies on quantitative methodology. After looking for validated tests to assess students' competence in Science in

Table 1. Dimensions of the socio-economic questionnaire.

	Initiation to scientific activity	The human being and health	Living beings	Matter and energy	Technology, objects and machines
Science content test					
Number of close questions	1	4	10	5	4
Number of open questions	4	1	1	6	6
Social context questionnaire					
Personal dimension			Gender, age, languages spoken at home.		
Academic dimension			Opinion about their school, attitudes and perceptions about learning process and the subjects.		
Socio-economic and cultural dimension			Family material goods, personal goods, cultural goods at home, favourite leisure activities, family cultural activities.		

Primary Education, the authors concluded that specific research tools were needed. Therefore, two surveys were designed for this project:

(1) Student socio-economic survey: this questionnaire enquired about participants' social and economic background. The survey included questions related to students' material and economic well-being, the type of cultural activities they are involved in, and the presence of cultural elements in the family environment (e.g. number of books at home). The survey design followed similar tools used in international evaluations, such as PISA, TIMSS, and PIRLS reports[2] and assessed students' socio-economic status (SES) by enquiring participants about the three dimensions shown in Table 1.
(2) Evaluation test for science contents (scored between 0 and 30): test designed to assess the degree of development of curricular contents in Natural Science for 6th year students of Primary Education in Spain. This test was built according to the curricular contents of the area of Science in the Spanish and Asturian curricula,[3] and included closed and open questions enquiring about contents and topics of the Science curriculum in participants' L1.

Both tools were validated by a group of experts composed of teachers of Primary Education (teaching Science in the 6th year) and university lecturers working in educational research methods. After the experts' approval, a final validation process of the tests was carried out through a pilot study with a group of students with similar characteristics to the research sample. The results obtained were satisfactory and the tools were accepted as valid for the current research.

To assess the reliability of the Science test, Cronbach's Alpha (0.790) was estimated, showing a good level of homogeneity in the items of the survey. The pilot stage was supported by the Kolmogorov-Smirnov test, which suggests that the sample presents a normal distribution.

Research sample

The research sample is composed of students of 6th year of Primary Education in public schools in Asturias. To guarantee a representative sample, the official records of the Regional Ministry for Education, Culture and Sports were checked: in 2015, the number of students enrolled in 6th grade of Primary Education in Asturias was 7651, 70% in public schools. Accordingly, the estimation for a representative sample was 400 students, though aiming to have the most representative sample possible, data were collected from 709 participants: 49.6% in CLIL schools (offering CLIL streams where students learn 2 subjects through the L2 and have two additional hours of English language per week) and 50.4% in non-CLIL establishments.

The selection of the participating schools followed an intentional sampling process, aimed to guarantee quality and balance: rural schools were discarded, as they account for less than 20% of the school population; in addition, in the case of Asturias, rural schools usually group students

belonging to several academic years in one single class in Grouped Rural Schools (*Colegios Rurales Agrupados*). Therefore, the sample includes schools located only in more urban areas, which comprises 80% of the population (1,042,370 citizens in 2016[4]). This sample includes most urban and semi-urban schools, and most of the infant population. All CLIL schools had developed their bilingual programmes for more than 6 years, and participants have been in the CLIL programme since they entered Primary Education. Science classrooms in CLIL provisions follow a similar pattern: contents are delivered in English, and Spanish is used only for rephrasing or reformulating concepts (sometimes, using code-switching); Textbooks used in bilingual streams follow the same structure, and include similar contents and type of activities; in the CLIL schools participating in the study, foreign language teachers deliver the Science lessons in the L2 with the support of general teachers.

To achieve a balanced sample, the social and economic context of schools was also considered. Hence, the project includes 18 schools (9 CLIL and 9 non-CLIL); aiming to avoid possible distortions caused by the social, economic, and cultural background of the schools, the sample includes CLIL and non-CLIL establishments in all urban and semi-urban areas.

Procedure

After selecting the sample, the research team visited the schools to deliver the surveys. Both tests were answered in Spanish in order to check students' command of Science in their native tongue.[5] The process lasted for 6 weeks until all 18 schools were visited. Data were analysed with SPSS v.22.

Results

First, the results of students' performance by type of school (CLIL vs. non-CLIL) are presented; then, we report on results according to participants' socio-economic status; finally, a bi-variate analysis with both variables (type of school and students' socio-economic status) is expounded.

Results according to the type of school

Table 2 shows the descriptive statistics for participants according to type of school attended:

As can be observed, there are differences between the means of the two groups, with students attending non-CLIL schools showing higher scores. In addition, students attending CLIL schools render a slightly higher standard deviation.

With regards to the variance homogeneity of the sample, Levene's test (3.732, Sig = .054) reports that the variance is homogeneous in the two groups. After observing the difference in the mean scores, we analysed if they are statistically significant through a one-way ANOVA.

The ANOVA shows that there are statistically significant differences between CLIL and non-CLIL schools ($F_{(1701)} = 37.185$, $p < 0.000$, $\eta^2 = 0.05$), with students learning contents through the L1 showing higher scores in Science than their counterparts studying through English.

Results according to students' socio-economic status

The descriptive statistics and the corrected score of students according to socio-economic status are shown in Table 3:

The variance homogeneity of the sample was analysed again using Levene's test, which rendered 7.453 (Sig = .001). As the test resulted in a significance of below 0.05, the hypothesis of the

Table 2. Results according to type of school.

	N	Mean	Standard deviation	Standard error
CLIL	352	14.99	5.50	.29319
Non-CLIL	357	17.41	5.07	.26842

Table 3. Results according to students' SES.

	N	Mean	Standard deviation	Standard error
High	191	16.65	5.19	.37588
Medium	306	17.02	4.98	.28472
Low	135	15.04	6.17	.53104

Table 4. Grouped socio-economic levels of the sample.

			Alpha subset = 0.05	
	Socioeconomic status	*N*	1	2
HSD Tukey	Low	135	15.04	
	High	191		16.65
	Medium	306		17.02
	Sig.		1.000	.775
Range Ryan-Einot-Gabriel-Welsch	Low	135	15.04	
	High	191		16.65
	Medium	306		17.02
	Sig.		1.000	.494

homogeneous variances in the groups was rejected, perhaps caused by the higher variance in students with low socio-economic status, with a bigger standard deviation regarding the other two groups (Table 3). As such, we conclude that the group of students with low social and economic status is more heterogeneous than the other two.

After analysing descriptive statistics on the differences in performance according to socio-economic status, we used a one-way ANOVA to check for statistically significant differences in the mean scores.

Results confirm that students of high and medium status achieve similar scores, while those belonging to low socio-economic status lag behind the first two groups. Statistically significant differences have been found when assessing this variable ($F_{(2629)} = 6.614$, $p < 0.001$, $\eta^2 = 0.021$). To determine in which groups differences can be found, the mean differences were analysed according to socio-economic status (Table 4):

The results confirm the trend identified with the descriptive statistics (Table 3), as students with high and medium socio-economic status are grouped together without significant differences between them; however, there are statistically significant differences between these two groups and that of students from a low socio-economic background.

Results of students according to the type of school and socio-economic status

Next, we analysed students' performance according to the two variables of this study: CLIL vs. non-CLIL, and the socio-economic status of participants. Table 5 shows the descriptive statistics:

Figure 1 shows that the results of students learning Science in non-CLIL contexts are higher than their counterparts in CLIL for every socio-economic group. It is also worth mentioning that students' performance is similar between students belonging to medium and high socio-economic status

Table 5. CLIL/Non-CLIL and SES mean.

CLIL/Non-CLIL	Socio-economic status	Mean	Standard deviation
CLIL	High	14.79	5.34
	Medium	16.09	5.05
	Low	14.04	6.01
	Total	15.25	5.42
Non-CLIL	High	18.37	4.41
	Medium	17.94	4.73
	Low	16.14	6.19
	Total	17.71	5.02

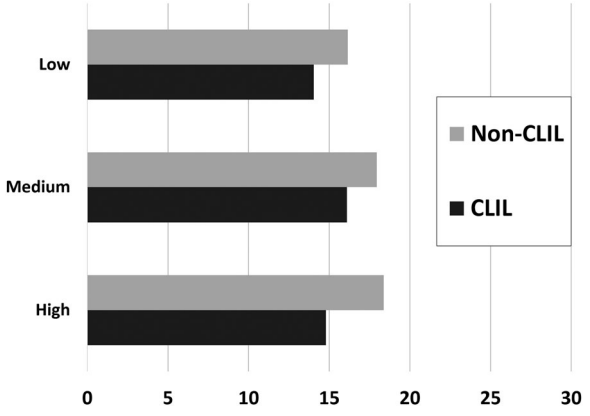

Figure 1. Comparison CLIL/Non-CLIL and SES mean.

Table 6. Inter-subject analysis.

Origin	Type 3 sum of square	gl	Root mean square	*F*	Sig.
Corrected model	1399.226	5	279.845	10.447	.000
Intercept	148,795.601	1	148,795.601	5554.512	.000
CLIL/Non-CLIL	893.135	1	893.135	33.341	.000
SES	349.335	2	174.667	6.520	.002
CLIL/Non-CLIL * SES	92.118	2	46.059	1.719	.180

enrolled in non-CLIL schools. Likewise, there are differences in the performance of CLIL and non-CLIL students belonging to low socio-economic and cultural status as regards their counterparts from higher social backgrounds. An inter-subjects test was performed to analyse the results further (Table 6):

The results of our analysis suggest there is a gap between students from lower socio-economic background and those participants with more favoured status, which can be grouped in one single layer. Students enrolled in CLIL schools and coming from less privileged backgrounds in our study outperform the other two categories.

Discussion

According to the results of the study, our research hypothesis has to be rejected, since the analysis shows that CLIL students perform slightly below non-CLIL pupils when evaluating their knowledge of Science in the L1. The initial hypothesis was that CLIL would not have any detrimental effect on content learning, since the implementation of any innovation in education (studying through the L2, in this case) should never negatively impact other areas.

Our second objective was to determine the possible effect of the type of school and students' SES on performance. From the results, we see that the type of school can be a determining factor in students' performance in Science, as students in bilingual provisions underperform their counterparts in mainstream schools. As for students' SES, the analysis shows that the distribution of the sample is similar in CLIL and non-CLIL establishments, with students from high, medium, and low backgrounds distributed in a similar way in schools. However, when observing students' performance according to status, statistically significant differences were found between students with lower SES and the first group composed of participants from medium and high backgrounds.

Arguably, the most relevant and original finding of the study is that students in CLIL provisions do not compare to those studying through the L1 when learning a content-subject. At this point, it is important to underline that CLIL programmes in Spain provide the same curriculum in Science as

mainstream schools and the conditions (teaching hours per week, etc.) are exactly the same. Also, we need to stress that both groups considered in this research meet the minimum standards in Science established in the Spanish curriculum of Primary Education. In any case, the results show that students in mainstream schools show better competence of Science contents in Spanish than those enrolled in CLIL provisions.

The main results of this study, however, do not tally with prior research, where no statistically significant differences were found between CLIL and non-CLIL groups (Admiraal, Westhoff, and de Bot 2006; Bergroth 2006; Stehler 2006), or studies that report higher attainment in subject-matters by students in bilingual systems (Serra 2007; Wode 1999). A possible cause for the poorer results of CLIL students in Science is, precisely, that they learn contents in the L2 in a region where exposure to the foreign language is limited (Muñoz 2007). Moreover, Science lessons are entirely delivered in English, as opposed to contexts where language alternation has been adopted as in Switzerland as noted by Serra (2007), who suggests that parents supported code-switching in the classroom to be able to follow their children's progress and help them with their assignments in the L1. This is becoming an issue in Spain, as many parents do not have sufficient language competence in English to help their children; therefore, this lack of support at home can also account for underperforming CLIL students. Another possible reason for the lower competence of CLIL students may be teacher training: in Spain, many teachers have not received specific training on CLIL methodology; frequently, foreign language teachers are in charge of subject-matters in bilingual streams as in other European settings (Eurydice 2012, 87; de Graaff et al. 2007); the fact that non-specialists deliver Science lessons might have an impact on students' learning of contents. Additionally, it should be noted that following a Science lesson in a foreign language is an additional challenge for students (Grandinetti, Langellotti, and Ting 2013). Also, the materials being used in Science in bilingual streams should be assessed in comparison with the textbooks utilised in L1 lessons with the aim to analyse the appropriateness and the suitability of textbooks in English regarding the contents specified in the curriculum, the length of the units, the number of examples, audiovisual materials, and so on.

Furthermore, what type of CLIL is being implemented in Primary Education should be assessed to check if 'weak CLIL' – where language is the learning objective and contents are highlighted but not usually assessed – is the main tendency as opposed to 'strong CLIL' – which is content-driven but where there is still a language focus – (see Paran 2013, 321). Finally, we also need to note that Spain does not have admission criteria for CLIL in public education, as in other countries which do require students' subject knowledge – e.g. the Czech Republic –, language competence in the L2 – e.g. France –, or both – e.g. the Netherlands – (Eurydice 2012, 42; Lorenzo, Casal, and Moore 2010; Pérez-Cañado 2012).

It is also important to underline that the evaluation of Science contents in our study was done in Spanish in CLIL and non-CLIL schools. The use of the L1 for the test is supported by the belief that there should be no differences in the knowledge of contents in a student's mother tongue, as CLIL is intended to promote both languages and not only the L2. In the event students command terminology in the L2, but do not master the specific vocabulary in the L1, CLIL would be failing to meet the objectives envisaged by the European Union and could be contributing to English monolingualism (Phillipson 2003; Seidlhofer 2004). The fact that students develop a strong register in the L2 at the expense of their mother tongue should be regarded as a failure in any bilingual provision, as suggested by Dalrymple-Smith, Karagiannakis, and Papadopoulos (2012). In any case, it would be interesting to know if CLIL students would have obtained better results if they had answered the Science test in English.

Conclusions

This paper has addressed the output of CLIL programmes in the development of students' competence in Science in the L1 by comparing students learning through the L1 and the L2. The

investigation aimed to evaluate students' academic performance in Science, as this is lacking in previous research, which has largely concentrated on exploring the linguistic benefits of CLIL.

The most relevant contribution of this paper is that students learning Science through an L2 in Primary Education underperform those who acquire contents in their mother tongue. To our best knowledge, this is the first study with a representative sample of the population of students in a Spanish Autonomous Community to evaluate academic achievements in public schools of Primary Education. Furthermore, the analysis of results allows us to conclude that CLIL might not be rendering expected results in content learning. In this framework, education authorities should take appropriate measures to prevent CLIL students from having better language competence and meta-linguistic skills than those in mainstream schools, yet show weaker knowledge in key areas such as Science. Also, we find that better teacher training in CLIL and the organisation of bilingual streams is needed; although the data of the current paper do not identify what causes the poorer results by CLIL students, we feel that the structure of bilingual programmes may need revisiting, as the poorer performance of the participants in the study may indicate that the integration of content and language is not being fully achieved: CLIL may be used in some schools and education settings as a synonym for language immersion (Lasagabaster and Sierra 2010; Pérez-Vidal and Juan-Garau 2010; Ruiz de Zarobe and Cenoz 2015); in addition, the integration between teaching and learning of language and content may require 'considerable training for both teachers and pupils' as concluded by Gajo (2007, 578). Second, the results suggest a need for a specific curriculum (or an adapted version) to suit the challenges of CLIL schools; finally, more support to teachers working in CLIL may be required (as suggested by Halbach 2008), especially, if we note that, frequently, it is language specialists who are teaching Science in Spain. Furthermore, our findings do not concur with prior research, as our results underline that CLIL students slightly underperform in the acquisition of Science contents in the L1 by the end of Primary Education. As such, we should enquire if new measures are required to optimise CLIL results as regards content acquisition. In our view, some possibilities may include: (a) review the curriculum of CLIL or design specific curricula for content-subjects in CLIL in Primary Education to guarantee that students in bilingual streams do not lag behind those in mainstream schools; (b) increase the number of teaching hours of content-subjects in bilingual streams; (c) improve teacher training (including language and methodological training).

The results of the present study should be taken with caution: first, the research has been carried out with students of Primary Education and the investigation does not generalise to other education stages (although researchers are invited to explore this line); second, the study has been carried out in a monolingual region with limited exposure to the L2. Furthermore, although this study might provide insights into bilingual education in several settings, we must note that CLIL is context-dependent and results cannot always be easily extrapolated from one country to another (Lasagabaster 2008; Pérez-Cañado 2012). Due to the diversity in the implementation of CLIL, results should not be generalised, as a comprehensive vision of CLIL requires from meta-analyses which can be conducted on the basis of studies similar to the present one.

To be sure, the current paper clearly underlines the need for further research in CLIL, investigating the development of contents and pedagogical issues related to the acquisition of non-language contents in Primary Education; in addition, this study could be replicated in other subjects such as Arts or Physical Education. Longitudinal studies and research into CLIL evaluation, classroom materials, students with special needs, and investigations that include students' socio-economic status are also welcome to analyse the effectiveness of CLIL in European education.

Notes

1. Although participating schools offer bilingual streams and are not 'completely CLIL establishments', for the sake of simplicity, the terms CLIL and non-CLIL schools will be used.

2. The PISA report is developed by the OECD: https://www.oecd.org/pisa/. TIMMS and PIRLS are published by the IEA: http://timss.bc.edu.
3. Ministry of Education, Culture and Sports of the Principality of Asturias. Asturian curriculum: https://sede.asturias.es/bopa/2014/08/30/2014-14753.pdf.
4. Institute of Economic Development of Asturias. http://www.idepa.es/sites/web/idepaweb/productos/cifras/demografia/index.jsp?section=3.
5. All students were delivered the questionnaire in order not to discriminate any pupil. However, students with special needs were omitted from the sample. In all the cases, students were excluded after consulting the teaching board of the schools and only in order to avoid possible distortions in the analysis. Students and schools were informed about the nature of this project and all ethical considerations were taken into account when writing the paper.

Disclosure statement

No potential conflict of interest was reported by the authors.

Funding

This work was supported by the Ministerio de Economía y Competitividad of the Government of Spain [grant number EDU2014-54093-R] and Government of Catalonia – Agència de Gestió d'ajuts Universitaris i de Recerca (AGAUR) [grant number 2014 SGR 208].

Notes on contributors

Javier Fernández-Sanjurjo is a researcher at the Department of Educational Sciences of the University of Oviedo, Spain.

Alberto Fernández-Costales is a Lecturer in Applied Linguistics at the University of Oviedo, Spain.

José Miguel Arias Blanco is Associate Professor in Research Methods and Diagnosis in Education at the University of Oviedo, Spain.

References

Admiraal, W., G. Westhoff, and K. de Bot. 2006. "Evaluation of Bilingual Secondary Education in the Netherlands: Students' Language Proficiency in English." *Educational Research and Evaluation* 12 (1): 75–93.

Alejo, R., and A. Piquer-Píriz. 2016. "Urban vs. Rural CLIL: An Analysis of Input-related Variables, Motivation and Language Attainment." *Language, Culture and Curriculum* 8318: 1–18. doi:10.1080/07908318.2016.1154068.

Alonso, E., J. Grisaleña, and A. Campo. 2008. "Plurilingual Education in Secondary Schools: Analysis of Results." *International CLIL Research Journal* 1 (1): 36–49.

Baetens Beardsmore, H. 1990. "Multilingual Education in Europe: Theory and Practice." In *Korean Language Education in China*, 107–129. Seoul: The Korean Society of Bilingualism.

Baetens Beardsmore, H., and M. Swain. 1985. "Designing Bilingual Education: Aspects of Immersion and European School Models." *Journal of Multilingual and Multicultural Development* 6: 1–15. doi:10.1080/01434632.1985.9994181.

Bergroth, M. 2006. "Immersion Students in the Matriculation Examination Three Years After Immersion." In *Exploring Dual-focussed Education. Integrating Language and Content for Individual and Societal Needs*, edited by S. Björklund, K. Mard-Miettinen, M. Bergström, and M. Södergard, 123–134. Vaasa: University of Vaasa, Centre for Immersion and Multilingualism. http://www.uwasa.fi/materiaali/pdf/isbn_952-476-149-1.pdf.

Bruton, A. 2011. "Is CLIL so Beneficial, or Just Selective? Re-evaluating Some of the Research." *System* 39 (4): 523–532. doi:10.1016/j.system.2011.08.002.

Bruton, A. 2013. "CLIL: Some of the Reasons Why … and why not." *System* 41: 587–597. doi:10.1016/j.system.2013.07.001.

Bruton, A. 2015. "CLIL: Detail Matters in the Whole Picture. More than a Reply to J. Hüttner and U. Smit (2014)." *System* 53: 119–128. doi:10.1016/j.system.2015.07.005.

Casal, S., and P. Moore. 2009. "The Andalusian Bilingual Sections Scheme: Evaluation and Consultancy." *International CLIL Research Journal* 1 (2): 36–46.

Cenoz, J. 2009. *Towards Multilingual Education: Basque Educational Research in International Perspective*. Bristol: Multilingual Matters.

Cenoz, J. 2013. "Discussion: Towards an Educational Perspective in CLIL Language Policy and Pedagogical Practice." *International Journal of Bilingual Education and Bilingualism* 16: 389–394. doi:10.1080/13670050.2013.777392.

Cenoz, J. 2015. "Content-based Instruction and Content and Language Integrated Learning: The Same or Different?" *Language, Culture & Curriculum* 28 (1): 8–24. doi:10.1080/07908318.2014.1000922.

Cenoz, J., F. Genesee, and D. Gorter. 2014. "Critical Analysis of CLIL: Taking Stock and Looking Forward." *Applied Linguistics* 35 (3): 243–262. doi:10.1093/applin/amt011.

Cenoz, J., and Y. Ruiz de Zarobe. 2015. "Learning through a Second or Additional Language: Content-based Instruction and CLIL in the Twenty-first Century." *Language, Culture and Curriculum* 28 (1): 1–7. doi:10.1080/07908318.2014.1000921.

Coyle, D. 2002. "Relevance of CLIL to the European Commission's Language Learning Objectives." In *CLIL / EMILE - The European Dimension: Actions, Trends, and Foresight Potential*, edited by David Marsh, 27–28. DG Education & Culture, European Commision. University of Jyväskylä.

Coyle, D., P. Hood, and D. Marsh. 2010. *CLIL*. Cambridge: Cambridge University Press.

Cummins, J. 2008. "BICS and CALP: Empirical and Theoretical Status of the Distinction." In *Encyclopedia of Language and Education: Vol. 2. Literacy*, edited by B. Street, and N. H. Hornberger, 71–83. New York: Springer.

Dafouz, E., and M. Guerrini, eds. 2009. *CLIL Across Educational Levels. Experiences from Primary, Secondary and Tertiary Contexts*. Madrid: Santillana.

Dalrymple-Smith, A., V. Karagiannakis, and K. Papadopoulos. 2012. "Does International Education Risk the Development of Students' Mother Tongues? An Investigation into Language use in Mathematics Classrooms." Masters' Thesis. Utrecht University. http://dspace.library.uu.nl/handle/1874/281151.

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

Dalton-Puffer, C. 2008. "Outcomes and Processes in Content and Language Integrated Learning (CLIL): Current Research from Europe." *Future Perspectives for English Language Teaching* May: 139–157. http://www.univie.ac.at/Anglistik/Dalton/SEW07/CLILresearch overview article.pdf.

Dalton-Puffer, C., and T. Nikula. 2006. "Introduction." *Vienna English Working Papers* 15 (3): 4–7.

de Graaff, R., G. J. Koopman, Y. Anikina, and G. Westhoff. 2007. "An Observation Tool for Effective L2 Pedagogy in Content and Language Integrated Learning (CLIL)." *International Journal of Bilingual Education and Bilingualism* 10 (5): 603–624. doi:10.2167/beb462.0.

Escobar Urmeneta, C. 2001. "Los Programas de Aprendizaje Integrado de Contenidos Curriculares y Lengua [Content and Language Integrated Programmes]." In *Didáctica de las Lenguas Extranjeras en la Educación Secundaria Obligatoria* [Foreign Language Teaching in Compulsory Secondary Education], edited by L. Nussbaum and M. Bernaus, 229–232. Madrid: Síntesis.

Escobar Urmeneta, C. 2013. "Learning to Become a CLIL Teacher: Teaching, Reflection and Professional Development." *International Journal of Bilingual Education and Bilingualism* 16 (3): 334–353. doi:10.1080/13670050.2013.777389.

Escobar Urmeneta, C., and A. Sánchez Sola. 2009. "Language Learning through Tasks in a Content and Language Integrated Learning (CLIL) Science Classroom." *Porta Linguarum* 11: 65–83. http://www.ugr.es/~portalin/articulos/PL_numero11/4%20C%20Escobar_A%20Sanchez.pdf.

European Commission. 2012a. "Europeans and their Languages. Special Eurobarometre 386." Accessed September 12, 2016. http://ec.europa.eu/public_opinion/archives/ebs/ebs_386_en.pdf.

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

Eurydice. 2012. "Key Data on Teaching Languages at School in Europe." Accessed September 12, 2016. http://eacea.ec.europa.eu/education/eurydice/documents/key_data_series/143EN_HI.pdf.

Fernández Fontecha, A. 2009. "Spanish CLIL: Research and Official Actions." In *Content and Language Integrated Learning. Evidence from Research in Europe*, edited by Y. Ruiz de Zarobe, and R. M. Jiménez Catalán, 3–21. Bristol: Multilingual Matters.

Fortanet Gómez, I., and M. F. Ruiz-Garrido. 2009. "Sharing CLIL in Europe." In *Content and Language Integrated Learning: Cultural Diversity*, edited by M. L. Carrió-Pastor, 47–75. Frankfurt-am-Main: Peter Lang.

Gajo, L. 2007. "Linguistic Knowledge and Subject Knowledge: How does Bilingualism Contribute to Subject Development?" *International Journal of Bilingual Education and Bilingualism* 10 (5): 563–581. doi:10.2167/beb460.0.

Grandinetti, M., M. Langellotti, and Y. L. Ting. 2013. "How CLIL can Provide a Pragmatic Means to Renovate Science Education Even in a Sub-optimally Bilingual Context?" *International Journal of Bilingual Education and Bilingualism* 16 (3): 354–374. doi:10.1080/13670050.2013.777390.

Halbach, A. 2008. "Bilingual Methodology in Primary Schools." *Revista de Educación* 346: 455–466. http://dialnet.unirioja.es/servlet/articulo?codigo=2591989&info=resumen&idioma=ENG.

Ianos, A., Huguet, A., Janés, J., and C. Lapresta. 2016. "Can Language Attitudes be Improved? A Longitudinal Study of Immigrant Students in Catalonia (Spain)." *International Journal of Bilingual Education and Bilingualism*. doi:10.1080/13670050.2015.1051508.

Jäppinen, A. K. 2005. "Thinking and Content Learning of Mathematics and Science as Cognitional Development in Content and Language Integrated Learning (CLIL): Teaching through a Foreign Language in Finland." *Language and Education* 19 (2): 147–168. doi:10.1080/09500780508668671.

Jiménez Catalán, R. M., and Y. Ruiz de Zarobe. 2007. "*Does the Type of Instruction Have any Bearing on EFL Learners' Receptive Vocabulary?*" Paper presented at the ELIA 10 conference. Issues in teaching, learning, and using vocabulary in an L2, University of Seville, Spain.

Lasagabaster, D. 2008. "Foreign Language Competence in Content and Language Integrated Courses." *The Open Applied Linguistics Journal* 1: 30–41. doi:10.2174/1874913500801010030.

Lasagabaster, D., and A. Doiz. 2016. "CLIL Students' Perceptions of their Language Learning Process: Delving into Self-perceived Improvement and Instructional Preferences." *Language Awareness* 25 (1–2): 110–126. doi:10.1080/09658416.2015.1122019.

Lasagabaster, D., and Y. Ruiz de Zarobe, eds. 2010. *CLIL in Spain. Implementation, Results and Teacher Training*. Newcastle: Cambridge Scholars.

Lasagabaster, D., and J. M. Sierra. 2010. "Immersion and CLIL in English: More Differences than Similarities." *ELT Journal* 64 (4): 367–375.

Llinares, A. 2015. "Integration in CLIL: A Proposal to Inform Research and Successful Pedagogy." *Language, Culture and Curriculum* 28 (1): 58–73. doi:10.1080/07908318.2014.1000925.

Lorenzo, F. 2007. "The Sociolinguistics of CLIL: Language Planning and Language Change in 21st Century Europe." *Revista Española de Lingüística Aplicada (RESLA)* 1: 27–38.

Lorenzo, F., S. Casal, and P. Moore. 2010. "The Effects of Content and Language Integrated Learning in European Education: Key Findings from the Andalusian Bilingual Sections Evaluation Project." *Applied Linguistics* 31 (3): 418–442. doi:10.1093/applin/amp041.

Madrid, D., and S. Hughes, eds. 2011. *Studies in Bilingual Education*. Bern: Peter Lang.

Marsh, D. 2002. *CLIL/EMILE-the European Dimension: Actions, Trends and Foresight Potential*. Strasbourg: European Commission.

Meyer, O., D. Coyle, A. Halbach, K. Schuck, and T. Ting. 2015. "A Pluriliteracies Approach to Content and Language Integrated Learning – Mapping Learner Progressions in Knowledge Construction and Meaning-making." *Language, Culture and Curriculum* 28 (1): 41–57.

Mohan, B., C. Leung, and T. Slater. 2010. "Assessing Language and Content: A Functional Perspective." In *Testing the Untestable in Language Education*, edited by A. Paran, and S. Lies, 217–240. Clevedon: Multilingual Matters.

Muñoz, C. 2007. "CLIL: Some Thoughts on its Psycholinguistic Principles." *Revista Española de Lingüística Aplicada (RESLA)* 1: 17–26.

Navés, T., and M. Victori. 2010. "CLIL in Catalonia: An Overview of Research Studies." In *CLIL in Spain. Implementation, Results and Teacher Training*, edited by D. Lasagabaster, and Y. Ruiz de Zarobe, 30–54. Newcastle upon Tyne: Cambridge Scholars Publishing.

Paran, A. 2013. "Content and Language Integrated Learning: Panacea or Policy Borrowing Myth?" *Applied Linguistics Review* 4 (2): 317–342. doi:10.1515/applirev-2013-0014.

Pavón Vázquez, V., and F. Rubio. 2010. "Teachers' Concerns and Uncertainties about the Introduction of CLIL Programmes." *Porta Linguarum: Revista Internacional de Didáctica de Las Lenguas Extranjeras* 14: 45–58. http://dialnet.unirioja.es/servlet/articulo?codigo=3253054.

Pérez-Cañado, M. L. 2012. "CLIL Research in Europe: Past, Present, and Future." *International Journal of Bilingual Education and Bilingualism* 15 (3): 315–341. doi:10.1080/13670050.2011.630064.

Pérez-Cañado, M. L. 2016. "From the CLIL Craze to the CLIL Conundrum: Addressing the Current CLIL Controversy." *Bellaterra Journal of Teaching Language and Literature* 9 (1): 9–31. doi:10.5565/rev/jtl3.667.

Pérez-Vidal, C. 2011. "Language Acquisition in Three Different Contexts of Learning: Formal Instruction, Stay Abroad, and Semi-immersion (CLIL)." In *Content and Foreign Language Integrated Learning. Contributions to Multilingualism in European Context*, edited by Y. Ruiz de Zarobe, J. M. Sierra, and F. Gallardo del Puerto, 103–127. Bern: Peter Lang.

Pérez-Vidal, C., and M. Juan-Garau. 2010. "To CLIL or not to CLIL? From Bilingualism to Multilingualism in Catalan/Spanish Communities in Spain." In *CLIL in Spain: Implementation, Results and Teacher Training*, edited by D. Lasagabaster, and Y. Ruiz de Zarobe, 115–138. Newcastle: Cambridge Scholars Publishing.

Phillipson, R. 2003. *English-only Europe? Challenging Language Policy*. London: Routledge.

Roquet, H., and C. Pérez-Vidal. 2015. "Do Productive Skills Improve in Content and Language Integrated Learning Contexts? The Case of Writing." *Applied Linguistics* 2015: 1–24. doi:10.1093/applin/amv050.

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. http://www.icrj.eu/11/article5.html.

Ruiz de Zarobe, Y., and J. Cenoz. 2015. "Way Forward in the Twenty-first Century in Content-based Instruction: Moving Towards Integration." *Language, Culture and Curriculum* 28 (1): 90–96. doi:10.1080/07908318.2014.1000927.

Ruiz de Zarobe, Y., and R. M. Jiménez Catalán, eds. 2009. *Content and Language Integrated Learning. Evidence from Research in Europe*. Bristol: Multilingual Matters.

Seidlhofer, B. 2004. "Research Perspectives on Teaching English as a Lingua Franca." *Annual Review of Applied Linguistics* 24: 209–239.

Serra, C. 2007. "Assessing CLIL at Primary School: A Longitudinal Study." *International Journal of Bilingual Education and Bilingualism* 10 (5): 582–602. doi:10.2167/beb461.0.

Stehler, U. 2006. "The Acquisition of Knowledge in Bilingual Learning: An Empirical Study on the Role of Language in Content Learning." *Vienna English Working Papers* 15 (3): 41–46.

Ullmann, M. 1999. "History and Geography through French: CLIL in a UK Secondary School." In *Learning through a Foreign Language. Models, Methods and Outcomes*, edited by J. Masih, 96–105. London: Centre for Information on Language Teaching and Research.

Wode, H. 1999. "Language Learning in European Immersion Classes." In *Learning through a Foreign Language. Models, Methods and Outcomes*, edited by J. Masih, 16–25. London: Centre for Information on Language Teaching and Research.

Xanthou, M. 2011. "The Impact of CLIL on L2 Vocabulary Development and Content Knowledge." *English Teaching: Practique and Critique* 10 (4): 116–126.