EL.LE e-ISSN 2280-6792
Vol. 7 – Num. 1 – Marzo 2018

Content-Specific Learning in CLIL
The Case of Physics Teaching in Italy

Fabiana Rosi
(Università degli Studi di Salerno, Italia)

Abstract The research investigates the effects of two learning contexts, content and language integrated learning (CLIL) and traditional learning (non-CLIL), upon content-specific outcomes, an aspect less explored than language-specific achievements in CLIL. Specifically, the study provides an interdisciplinary analysis of English CLIL applied to Physics in Italian high school. Two different levels of student competence are examined: selecting answers for content-specific issues, and content-related argumentative skills, in order to measure how learners comprehend and discuss content. The comparison between CLIL and non-CLIL classes in pretest, posttest and delayed posttest accounts for content assimilation and retention. Findings show that CLIL students significantly outperform non-CLIL students in both levels of competence in posttest and even more so in delayed posttest, a difference which emerges in terms of mean and coefficient of variation. Furthermore, feedback questionnaires display enhanced motivation, a well recognised influential benefit of CLIL in language-specific learning. Such results have pedagogical implications and may contribute to a better understanding of the correlation between language, content and motivation in CLIL.

Keywords High school. Content and language integrated learning. Acquisition of disciplinary content. Argumentative competence. Motivation.

1 Introduction

Content and language integrated learning (CLIL) may be defined as a dual-focused educational approach that combines language and content by using an additional, usually foreign, language as medium of instruction for disciplinary content (Coyle, Hood, Marsh 2010).[1] In recent years CLIL programmes have burgeoned all around the world. In particular, many European school systems promote CLIL since it is considered a powerful

1 In the definition of Coyle, Hood and Marsh (2010, 1), the expression 'additional language' is used with the explicit purpose of also including in CLIL those programmes providing specific content instruction through a second language or minority language.

DOI 10.30687/ELLE/2280-6792/2018/01/002
Submitted: 2017-12-03 | Accepted: 2018-07-19

resource to meet the European Commission and Council of Europe requirements of educating multicultural and multilingual citizens in all member countries (Eurydice 2006). The recent diffusion of CLIL goes hand in hand with a growing interest from a scientific perspective as well. Research has so far shed light on benefits of CLIL for learners in language proficiency and motivation (Dalton-Puffer 2011; Dalton-Puffer, Nikula 2015; Roquet, Perez-Vidal 2017), while less attention has been paid to the impact of CLIL upon content-specific learning. As repeatedly affirmed (Dalton-Puffer 2008; Cenoz, Genesee, Gorter 2014; Nikula 2017), content-oriented research is crucial in order to grasp the integration of the two components comprising the dual focus of CLIL, i.e. language and content, with regard to the learning process and achievements. In particular, researchers, and stakeholders too, question whether learners enrolled in CLIL can acquire content to the same extent as peers schooled in their native language (L1). Do CLIL students master conceptual complexity and academic language at a comparable depth to that of non-CLIL students? Evidence for such questions is urgently needed, although inquiries on content-specific learning in CLIL face several theoretical and methodological difficulties. There is no standard evaluation of disciplinary competence analogous to the internationally validated instruments existing for language testing (Dalton-Puffer 2011). Nor has a standard for the assessment of disciplinary competence taught through CLIL been defined (Leone 2015). In addition, surveys on content-specific learning require an interdisciplinary research group to encompass all necessary expertise in both the discipline and language learning and teaching.

The paper aims to contribute to this open debate by comparing content-specific outcomes in students enrolled in CLIL and in non-CLIL instruction of Physics in Italian public high schools. The study draws on data from 34 students, all enrolled at the same high school in a small town in Southern Italy. Two classes were examined: one experienced a teaching unit on Newton's laws of motion in English; for the other, the same teaching unit was carried out in Italian. The research focuses on the effects of absolute initial immersion in a CLIL context, given that students were exposed to CLIL for the first time during the experiment. The fact that CLIL is a non-familiar methodology for learners may have a double impact. On the one hand, the use of a foreign language, fully mastered by neither teacher nor students, makes understanding and classroom interaction more complicated for pupils. For instance, Lo and Macaro (2015) notice that initial CLIL lessons are prevalently monological, since learners limit their questions and participation with respect to traditional subject-specific lessons, due to problems communicating in the foreign language. On the other hand, the change of teaching routine and the opening contact with an innovative methodology foster learners' attention and motivation. The latter position is supported by surveys on language-specific outcomes in

CLIL (Lasagabaster 2011; Doiz, Lasagabaster, Sierra 2014), reporting that students enrolled in CLIL settings are actually more motivated in their study of the foreign language than are their traditionally educated peers. The present research intends to verify whether CLIL may enhance student motivation for content-specific learning as well.

The choice of Physics as our investigated school subject is not accidental. Scientific subjects, and Physics above all, are most frequently selected to be taught through CLIL in Italy, since teachers tend to believe that the strict connection of disciplinary issues with concrete experiences and the use of non-verbal elements, such as symbols, images and graphs, make such disciplines bound up with language to a lesser extent than the Humanities and therefore more adaptable to CLIL. Nevertheless, in general, Italian students display significant difficulties in scientific school subjects,[2] which are perceived as complicated to understand and master. Thus, if the CLIL experiment with such a subject proves efficient, this would pave the way to proposing CLIL as fruitful methodology for a challenging subject matter and as a strategy to address a well documented weakness in Italian students' preparation. In this regard, it is worth noting that the data analysed come from a High School for the Humanities, where students are particularly unmotivated with regard to the study of Physics, since their focus is primarily on pedagogy and psychology.

Furthermore, in Italy interest in the effects of CLIL on content-specific achievements is especially lively, as disciplinary teachers are in charge of CLIL programmes instead of language teachers. This situation is mainly a consequence of the school reform established by the Italian Ministry of Education in 2010, which mandated the provision of CLIL instruction in the last year of high school, but provided not sufficient national training programme for teachers to be involved in CLIL.[3] Thus, school management had to implement the "CLIL revolution" relying only upon ordinary resources. What occurred was that reform affected in-service high school subject teachers, who frequently lacked adequate foreign language proficiency, prior CLIL experience, and language awareness[4] (Aiello, Di Martino, Di Sabato 2017).

2 OCSE-PISA Report 2015, http://www.oecd.org/pisa/ (2018-09-17).

3 From 2010, only two editions of national training programmes for in-service teachers were established by the Italian Ministry of Education (in 2014 and in 2018) and no specific educational programme for students interested in working as school teachers was instituted.

4 Language awareness is fundamental for teacher effectiveness in CLIL: it addresses not only the teacher's use of foreign language, but also possible student difficulties in using the foreign language to learn non-linguistic content (Eurydice 2006).

2 Content-Specific Learning in CLIL

Among studies on content-specific learning there is not yet an agreement regarding the impact of CLIL. Evidence for both positive (Van de Craen, Ceuleers, Mondt 2007; Serra 2007; Haagen-Schützenhöfer, Hopf 2014; Canlas 2016) and negative (Lim Falk 2008; Dallinger et al. 2016; Piesche et al. 2016; Fernández-Sanjurjo et al. 2017) impacts of CLIL have been reported. In some surveys no significant differences even emerged between CLIL and non-CLIL students with regard to knowledge of content (Seikkula-Leino 2007; Costa, Mariotti 2017a). Nevertheless, negative evidence has emerged mostly in studies examining content by means of an analysis of subject-language use, while evidence for the benefits of CLIL is more substantial in analyses of content-specific data.

An outperformance of students enrolled in CLIL programmes over traditionally schooled peers has been documented for different countries, languages, although the large predominance of the use English as language for CLIL, and disciplines, such as mathematics (Van de Craen, Ceuleers, Mondt 2007; Serra 2007; Murray 2010; Surmont, Van Den Noort, Van de Craen 2016), chemistry (Gregorczyk 2012), geography (Vollmer et al. 2006), and history (Bauer-Marschallinger 2016). The convergence of data from different disciplines lends support to the hypothesis that CLIL students' outperformance is not correlated to a discipline-specific effect, but rather to a cross-sectional factor shared by all investigated educational contexts. Vollmer et al. (2006) underline that the linguistic difficulties and frustration experienced in CLIL, far from leading students to abandon their studies, instead inspire them to work more persistently so as to develop higher cognitive strategies with which to construct knowledge. Indeed, CLIL learners show fine-grained analytical skills for accessing and understanding content, such as an advanced capacity to detect contradictions and to integrate details with overall meaning. Such students' persistent work and commitment are in line with motivation, a key factor of CLIL instruction which the present research intends to correlate with content-specific outcomes.

As for Physics, the discipline at issue in the present research, the literature points to the positive influence of CLIL upon several features of content learning. Jäppinen (2005) surveyed cognitional development, conceived as both concepts and conceptual structures, in students of Finnish schools (aged 7-15), who received science and mathematics instruction in English, French and Swedish, in one half of the sample, and in the other half, in L1 Finnish. Four content-specific tests, after four teaching units on individual topics, were taken in L1 Finnish by the entire sample over the school year. CLIL students mostly improved their cognitional skills at the same rate as their L1 taught peers. Namely, CLIL pupils aged 7-9 displayed some difficulties only in grasping the most abstract concepts,

which they did not experiment in every-day life; CLIL students aged 10-12 learned to the same extent and faster than their peers; in learners aged 13-15 the CLIL and non-CLIL group did not display differences, given the more limited CLIL hours in accordance with the Finnish school system. In particular, the author noticed that CLIL students outperformed their L1 taught peers in the capacity to compare different concepts; they explained this phenomenon as due to the familiarity in matching diverse languages and cultures thanks to their trilingual CLIL education. More recently, Haagen-Schützenhöfer and Hopf (2014) examined achievements in Physics and changes in motivation in Austrian high school pupils (Year 11). The Austrian learners studied were enrolled in a 4-month CLIL project encompassing both Physics lessons in English and English lessons dealing with issues related to Physics. They were pretested and posttested in their L1 about issues taught in English in some classes and in L1 in the others. The results indicated that CLIL contributed to a better content comprehension and to an increase in motivation, especially in students with a great interest in English. Besides, Canlas (2016) analysed the achievements of Kazakh high school pupils (Year 9) who studied Physics in English. The scores of content-specific tests, proposed in this case in English, showed that 84% of learners had mastered Physics issues addressed through CLIL, a confirmation of the positive effects of CLIL upon both conceptual understanding and critical thinking skills.

Within the Italian territoire, the learning of content has been investigated in CLIL and non-CLIL classes in both school and higher instruction. As for CLIL in school, the surveys on primary school carried out by Infante (2010) and on high schools accomplished by Ricci Garotti (2017) showed that CLIL students taught in English performed generally at the same level of their peers educated in Italian. In the survey about high school, some classes had experimented with the CLIL methodology before the survey, others not, however no differences emerged depending on the different familiarity of the students with CLIL. As for CLIL in higher instruction, Pigliapoco and Bogliolo (2009) found a slight advantage for students attending a CLIL course of Computer Science compared to peers involved in analogous course taught in L1, while Costa and Mariotti (2017b) documented a similar competence in CLIL and non-CLIL students of Economics (see also Costa, Mariotti 2017a) and International Relations and a slight advantage for non-CLIL students of Geometry and Physiopathology.

3 Methodology

3.1 Research Design

Our research questions are:

1. Does CLIL have an effect on short-term and long-term learning of content?
2. Does motivation have a role in content-specific learning in CLIL, in addition to that already documented for language-specific learning?

The data were elicited by means of a pretest, posttest and delayed posttest and compared between the CLIL class and the non-CLIL class. The two classes were in the same high school, so that students were comparable with respect to relevant factors, such as school setting, geographic provenance, cultural and socio-economic background, language repertoire (Italian and local dialect), and previous studies of English and Physics.

In the CLIL class, students started to experience CLIL instruction when the research began. They had never studied Physics in English before and CLIL was not a familiar methodology for them. The third year of high school was selected for the research precisely because students had not yet started to study any non-linguistic subject in a foreign language, a programme which takes place in the fifth year of high school in Italy.[5]

3.2 Participants

The initial sample was composed of 54 students, but only 34 students performed all three of the tests and were included in the analysis. Learners (aged 16) were students of the third year (Year 11) of the High School for Humanities in a public school, located in a small town (Nocera Inferiore) near Salerno, in Southern Italy. They were part of two classes: 16 students were in a CLIL class, 18 in a non-CLIL class. In both classes, there was one student with an L1 different from Italian, respectively German and Ukrainian. In the CLIL class, male (50%) and female (50%) pupils were equally balanced, while in the non-CLIL class the female students (89%) were much more numerous than their male counterparts (11%). At the beginning of the research, all pupils had studied English for 10 years and possessed an A2-B1 (low-intermediate) level. As for Physics, participants had studied it for 6 months, since the subject had only been introduced in the third school year. The two classes were homogenous with regard to

5 According to the school reform established by the Italian Ministry of Education in 2010, CLIL instruction starts from the third year only in language lyceums.

the starting level of competence in Physics, as revealed by pretest results. Given the school specialisation in Humanities, neither English nor Physics were the main subjects for students, who were more focused on pedagogy, psychology, sociology, and anthropology.

Two teachers took part in the experiment. Both teachers were women aged between 55 and 60. In the CLIL class, the teacher was a career Physics teacher with 25 years of teaching experience, in the control class, the Physics teacher also had comparable experience, with 27 years teaching. Indeed, the two teachers usually worked in parallel in their classes, sharing curricula, materials and techniques, as they did for the monitored teaching unit. The teacher in charge of the CLIL class had passed the C1 Cambridge Advanced English Examination just before the research began, as well as she passed two methodological courses on the CLIL approach: the teacher training course for teaching CLIL at the University of Salerno, and the blended course *Teaching your Subject in English* organised by Cambridge International Examinations.

3.3 Elicitation Procedure

Both classes carried out the same predetermined Physics teaching unit (TU) having as its topic Newton's laws of motion. The TU took 5 weeks, from March 4th to April 7th 2017, following the regular subject timetable (2 hours per week). Both teachers alternated frontal lessons, when explained the Newton's three laws and related concepts, with hours in a laboratory where students conducted experiments, and interactive classroom activities, such as collective content review and exercises. In the two classes, the same number of hours of Physics were provided: in the CLIL group, the TU was taught in English, in the other in Italian. After the experiment, the CLIL class returned to studying Physics in Italian.

In the two classes, three tests were carried out: a pretest before the beginning of the TU (March 4th), a posttest immediately after the conclusion of the TU (April 8th), and a delayed posttest after 5 weeks (May 13th). The pretest questions dealt with Physics issues preliminary to Newton's laws, such as the concepts of force, mass, and speed, in order to gauge the students' starting competence in Physics and to verify that the two classes were comparable. In the posttest, the questions addressed the topic of the TU. After the posttest, teachers passed to a subsequent topic, which entailed recall of TU issues. In the last session, the posttest was totally replicated, to measure retention after 5 weeks of the competence acquainted in TU. In both classes, the teacher immediately revised the posttest together with students. Consequently, all of the participants received feedback on correct answers 5 weeks before compiling the delayed posttest. Thus, both groups were analogously exposed to an echo effect in the delayed posttest.

The tests lasted 50 minutes and were performed by students during ordinary lesson hours under the supervision of the Physics teacher and a researcher, external to both classes. The tests were prepared by the two teachers together and were identical for the two classes. The language of the test was Italian, because the control group was taught in Italian. All three tests consisted of 10 multiple-choice questions and 10 open-ended questions: students read a question, selected one answer among the three options provided and wrote a reason for their choice (maximum three lines). The three answer options differed in accuracy: they included a correct answer, which a student who had fully mastered the topic could recognize; a trick answer, which would potentially confuse a student who had acquired partial knowledge of the topic, and an incorrect answer, which only a student who did not understand the topic would choose (see Table 1).

Table 1. Example of answer options for multiple-choice questions[6]

Question	L'inerzia di una bicicletta cambia se
Correct answer	leghiamo il nostro casco.
Trick answer	pedaliamo su una strada in salita.
Incorrect answer	proviamo a spingerla.

As emerges from Table 1, the three answer options were homogenous for complexity, number of words and use of technical terms. Throughout the questions, the order of the three types of answer was counterbalanced.

The prompt for the open-ended question was: "Why?". Unlike the multiple-choice questions, students did not receive any cues and were asked to verbalise answers in their own words regarding content-specific issues they had learnt. The answers collected differed in both accuracy and complexity. Learners produced correct, partially correct and incorrect argumentations. Table 2 reports three examples of reasoning provided for the question showed in Table 1: a correct argumentation, appropriately presenting the correlation between inertia and mass; a partially correct answer, focusing on the increase of force, but not clarifying that the increase of force is required by an increment of mass; and an incorrect answer, displaying the learner's confusion regarding the concepts of force and mass.

6 Question "The inertia of a bicycle changes if", correct answer "we tie our helmet", trick answer "we are pedaling on an uphill road", incorrect answer "we try to push it".

Table 2. Examples of answer for open-ended questions[7]

Correct answer	Aggiungendo il casco alla bici aumenta la massa quindi l'inerzia cambia.
Partially correct answer	Perché bisogna applicare una forza maggiore e in questo caso l'inerzia cambia.
Incorrect answer	In questo modo aggiungiamo un'altra forza.

In the pretest session, students also completed a background questionnaire that contained questions about their age, mother tongue and other known languages, number of years of study of English, a self-evaluation of language proficiency in English, their personal opinion about English and Physics and their favourite school subject. Meanwhile, teachers completed a background questionnaire containing questions about their age, number of years of teaching experience, teaching methods and techniques, their personal opinion about CLIL, and a TU description. In the delayed post-test session, teachers and students completed a feedback questionnaire on the concluded TU. Finally, teachers provided detailed written lesson plans for the TU.

3.4 Transcription, Scoring and Analysis

Tests performed by students who were absent in at least one of the three sessions were excluded from our data. The total amount of answers was 2.040 (20 answers × 3 sessions × 34 participants): 1.020 answers to multiple-choice questions and 1.020 answers to open-ended questions. The two types of answers provide two measurements of students' competence in Physics: the ability to select the appropriate issues for the given context (M1) and content-related argumentative skills (M2).

The answers were transcribed and scored using a dichotomous scale. The answers to multiple-choice questions were scored 1 if correct, 0 if trick, incorrect or left blank, in accordance with the answer key provided by the teachers when they prepared the test. The answers to open-ended questions were scored 1 if correct, 0 if partially correct, incorrect or left blank. For the second type of answers, the accuracy of the students' productions was rated by a high school Physics teacher external to the classes, in order to avoid any bias in rating due to personal knowledge of the participants, as in the case of teachers of the two classes. The evaluator

7 Correct answer "By adding the helmet to the bicycle the mass increases so the inertia changes", partially correct answer "Because a stronger force is needed and in this case the inertia changes", incorrect answer "In this way we add a further force".

was familiar with the topic and the methods used to teach it at the level of the investigated classes but was external to the school and independent in evaluation.

As for the analysis, the sum score for every student was calculated from each test. These scores were used to calculate the mean of each class, test by test. The differences between the mean of the two classes, based on the sum scores of the students' tests, were analysed with a T-test when $p < 0.05$ and $p < 0.01$, in order to verify the reliability of comparisons between means. A between-groups analysis and a within-group analysis were performed using a T-test to validate comparison of results between the two classes in each test as well as comparison of a single class' scores among the three elicitations. After having calculated the means and standard deviation, the coefficient of variation was calculated so that data dispersion within classes would be taken into account.

Together with a quantitative analysis of test scores, a qualitative analysis of answers on the feedback questionnaire from both teacher and students was performed.

4 Analysis

4.1 The Selection of Appropriate Content-Specific Issues (M1)

Table 3 reports the differences between the CLIL class and the non-CLIL class in the three tests for multiple-choice questions, which measure the ability to select the appropriate Physics-specific issues for the given context.

Table 3. Mean and coefficient of variation (CV) for answers to multiple-choice questions

	Pretest	Posttest	Delayed posttest
CLIL class (mean)	4.06	5.5	9.12
Non-CLIL class (mean)	3.83	3.83	5.95
CLIL class (CV)	0.39	0.29	0.14
Non-CLIL class (CV)	0.41	0.40	0.29

On the pretest, the mean of test scores is very similar between CLIL and non-CLIL contexts and no statistical significance emerges between the two classes. Therefore, the starting point in Physics competence of the experimental and control groups is aligned. On the posttest, after the TU taught either through CLIL or not, a divergence between the two classes arises and emerges as statistically significant ($t = 3.0583$, df = 32, $p < 0.05$). By the time of the delayed posttest, the difference between the two class-

es is much more marked and more statistically significant (t = 5.9029, df = 32, $p < 0.01$) than on the posttest. The difference consists in a higher mean score for CLIL students compared to non-CLIL students. The outperformance of the experimental group thus increases over time: on the posttest, CLIL students outperform non-CLIL students by 1.67 points and on the delayed posttest that figure is 3.17 points. Moreover, for the CLIL class, the mean on the delayed posttest is more than the double the pretest mean (9.12 vs. 4.06), while the progress of the non-CLIL class is not so remarkable (5.95 vs. 3.83). It is worth noticing that the experimental group score varies from the pretest to the posttest, whereas in the non-CLIL class no change occurs. An increase of non-CLIL student scores arises only with the delayed posttest, probably due to the echo effect of the replicated test in the second and third elicitation sessions. In parallel with these results, the coefficient of variation decreases in both classes, but in the CLIL group it decreases to a stronger extent. The comparison among the three tests is corroborated by the results of within-group analysis: for the CLIL class, the difference first between the pretest and posttest and second between the posttest and delayed posttest are significant (respectively: t = 2.5065, df = 30, $p < 0.05$; t = 6.8227, df = 30, $p < 0.01$); for the non-CLIL class, instead, the difference between the pretest and posttest is not significant, since no relevant variation occurs in the teaching unit on Newton's law with respect to previous teaching, only the difference between the posttest and the delayed posttest turns out to be significant (t = 3.8618, df = 34, $p < 0.01$), possibly due to the replication of the same test. As for the mean, the coefficient of variation changes from the pretest to the posttest only in the CLIL class and with regard to the delayed posttest the decrease of this value is more notable in the CLIL group (0.25 points) than in the non-CLIL group (0.11 points). In other words, the disparity of student competence within the class is analogous between the two groups on the pretest, diminishes only in the CLIL class after CLIL intervention and lowers in both groups, but to a greater degree in the experimental one, when the posttest is replicated in the third elicitation.

To sum up, our overall findings for M1 point to a positive effect of CLIL on content-specific learning: after the CLIL experience, students increased their capacity to provide correct answers and became more homogenous in answering, in posttest and even more in delayed posttest, while non-CLIL learners improved their scores and became more homogenous only when they completed the posttest for the second time in the delayed posttest session.

4.2 Content-Related Argumentative Skills (M2)

Content-related argumentative skills were measured on the basis of answers to open-ended questions, which asked for a defense of previous answer to content-based multiple-choice questions. Students appeared to have attained a lower level of this type of content-specific competence in both classes in all the three tests, as evident in Table 4, where the maximum score is 3.81/10 while in Table 3 the highest score is 9.12/10.[8]

Table 4. Mean and coefficient of variation (CV) for answers to open-ended questions

	Pretest	Posttest	Delayed posttest
CLIL class (mean)	0.25	2.12	3.81
Non-CLIL class (mean)	0.33	0.88	1.38
CLIL class (CV)	4.00	0.98	0.58
Non-CLIL class (CV)	2.30	1.15	0.99

On the pretest, no significant difference arises among the two classes. Students' argumentative competence does not diverge in the experimental and control groups. In turn, on the posttest the two classes appear to be significantly different ($t = 2.2268$, df = 32, $p < 0.05$), a divergence which increases further on the delayed posttest ($t = 3.8626$, df = 32, $p < 0.01$). As with the results for M1, for M2 the difference between the two classes consists in the outperformance of CLIL students over non-CLIL students, which starts immediately after the CLIL experience (1.24 points of divergence between the two groups in posttest) and rises in the long run (2.43 points in delayed posttest). Results of within-group analysis are consistent: for the CLIL class the difference between the pretest and the posttest and between the posttest and the delayed posttest are significant ($t = 3.2325$, df = 30, $p < 0.01$; $t = 2.2080$, df = 30, $p < 0.05$); for the non-CLIL class differences among the tests are not significant. The diminution of the coefficient of variation from pretest to delayed posttest shows that the classes become more homogeneous.

Thus, findings for M1 and M2 display the same tendencies: test scores increase more in the CLIL class than in the non-CLIL class, and disparity among students diminishes more in the experimental group than in the control one. Both tendencies occur to a larger extent on the delayed posttest than on the immediate posttest.

8 It is interesting to notice that in both measurements the highest score is attained by the CLIL class on the delayed posttest.

4.3 Teacher and Student Voices

Feedback questionnaires display the perceptions of the strengths and weakness of CLIL instruction as experienced by the teacher and students. In this paper, participant voices were qualitatively analysed to identify claims about motivation, in order to answer the second research question. On one hand, the teacher underlined the importance of introducing CLIL starting from the first moment she meets classes in the third year of high school, as in the experiment, two years before the ministerial obligation. On the other, learners mostly declared that they enjoyed CLIL instruction and explained their appreciation on the basis of several features: CLIL leads students to be more attentive, involves the entire class, introduces a new way of learning, makes lessons more alluring and allows for the study of two subjects at once, as reported in Table 5.

Table 5. Example of students' opinions about CLIL[9]

Student1	Non capisco molto bene l'inglese, ma siccome la prof parlava in inglese stavo più attento e cercavo di capire.
Student2	Sì diciamo che è stata un'esperienza nuova che ha catturato la mia attenzione.
Student3	Rende la lezione più interessante e coinvolge tutti.
Student4	È stato un modo di apprendere fisica diverso dal solito.
Student5	Studi due cose contemporaneamente la Fisica e la lingua Inglese.

The first claim is particularly interesting, as it demonstrates the fact that the difficulty in CLIL caused by using a foreign language may actually have positive consequences, in that students are stimulated to commit themselves more intently to the lesson in order to avoid missing content-specific information or risk lagging behind their peers.

On the contrary, students underestimated their content achievements and emphasised language problems they encountered in CLIL instruction. The teacher also underestimated the students' performances, believing the best students to have gained a greater advantage than the others had. These evaluations are in contrast with actual class outcomes (Table 3 and 4).

9 Student1 "I do not understand English very well, but since teacher spoke in English I was more attentive and sought to understand", Student2 " Yes, let's say that it was a new experience which caught my attention", Student3 "It makes lessons more interesting and it involves everybody", Student4 " It was a different way of learning Physics than usual", Student5 "You study two subjects at the same time, Physics and English".

5 Discussion

The first research question addresses whether CLIL has an effect upon content-specific competence, as operationalised in two separate measurements: the ability to select the appropriate Physics-specific issues for the given context (M1) and content-related argumentative skills (M2). M1 deals with a receptive competence, in that students are required to identify the correct solution selecting among three given options. In order to limit the likelihood of identifying the correct answer thanks only to a vague representation of concepts, a trick option was inserted among the answers for every question (see Table 1) and scored 0, as well as incorrect or blank answers were scored 0. This enabled researchers to disentangle participants' incomplete acquisition of the topic from their mastery of the topic. Differently, M2 implies a more autonomous and complex competence: students first have to retrieve that content-specific knowledge they possess which is pertinent to the topic at issue and then they have to verbalise it in their own words, without any cues. Moreover, content-specific issues are communicated with the intentional purpose of defending the answer given to a question, so a student's claim is not an explanation but rather an argumentation, attempting to provide a compelling reason for the assertion previously chosen. The argumentation has to be clear, complete, characterised by relevant information and precise terminology. This level of competence necessarily involves a deep comprehension of content and autonomous critical thinking, since learners can provide a correct reason only if they elaborate on content they studied. It is no wonder that the results for M1 are better than those for M2. In both classes, test scores are higher for multiple-choice questions (M1) than for open-ended questions (M2) in all three elicitations (see Table 3 vs. Table 4).

What was not so easily predictable is the outperformance of CLIL students over non-CLIL students for both measurements. Indeed, both M1 and M2 analyses show that after the CLIL experience, content-specific competence increases for students in the CLIL class more than that in the non-CLIL class and that the experimental group becomes more homogeneous in its abilities than the control one. The reduction of disparity among students in the CLIL class is not oriented towards low scores but rather towards high ones. These findings account for a positive role of CLIL[10] upon content-specific learning at both levels of student competence. It is worth

10 CLIL appears to be the unique variable responsible for the outperformance of the experimental group over the control one. The general teaching style of the two different teachers may not account for results, as demonstrated by the alignment of means of the two classes for both M1 and M2 in the pretest, before the beginning of the CLIL experience in the experimental context. Moreover, in both classes the teacher followed the same predetermined lesson planning for the teaching unit under investigation.

noting that before any CLIL education, in pretest, the two groups show the same degree of both levels of competence, and that subsequently for the non-CLIL class no significant difference appears between the pretest and the posttest – and this is the case for both M1 and M2. As the students were allowed to use their L1 in the tests, so the causes of differences in answers between the two classes do not depend on language problems in understanding the questions or in producing the answers, but rather on the degree of comprehension and elaboration of content taught, either through CLIL or not.

Moreover, an increasing trend emerges in our data: CLIL students outperform non-CLIL students even more on the delayed posttest than on the immediate posttest. This tendency might be related to a growing familiarity with task structure and questions content for participants. The delayed posttest was identical to the immediate posttest and the teacher revised the test after the posttest elicitation, 5 weeks before the delayed posttest. Nevertheless, if an echo effect occurs, it is equally distributed across the two learning contexts, and so cannot be the cause of any increase in difference between the two classes in the delayed posttest. Thus, the enhanced competence in the delayed posttest may be correlated to effects of CLIL, which arise in short-term learning and, to a stronger extent, in long-term learning. The best outperformance of CLIL students over non-CLIL students in the delayed posttest for both measurements is indicative of a deeper learning of content in the class where such content was taught in English. A possible explanation relies on the role played by richer classroom interaction. Indeed, in the collected feedback questionnaire, students enrolled in CLIL underlined the importance of the teacher's interactional strategies, such as scaffolding, the abundance of examples and the repetition of difficult concepts in both the foreign language and the mother tongue.[11] These reformulation and clarification efforts testify to the teacher's attentiveness to the linguistic difficulties students may encounter in a foreign language and an increased language awareness on the part of the non-language teacher (Sisti 2017). This is an undeniably positive consequence of CLIL. As Cummins (1979) pointed out, all teachers teach also language, since they employ academic language and linguistic structures not shared with common language to convey disciplinary content, but they are usually not aware of it. For instance, in Airey's (2012) interviews with Swedish Physics teachers, they state that they correct students' math-

11 The repetition of contents in mother tongue is a controversial issue in CLIL teaching, since, on one hand, it supports the comprehension of topics not understood in the foreign language, on the other, it reduces the room of the foreign language and student motivation to communicate in the foreign language. According to the CLIL teacher involved in the experiment, students' low proficiency in the foreign language makes necessary to repeat the most difficult concepts in Italian.

ematical mistakes but not linguistic ones, since they do not perceive them as part of their own duty. In turn, CLIL puts disciplinary teachers face to face with language difficulties and asks them to find solutions, which have a positive impact on disciplinary teaching as a whole, not only on students' linguistic competence. Learners interviewed in the current research revealed that they found CLIL lessons alluring, challenging, and advantageous since they were useful for learning two subjects at the same time. They recognised that difficulty in understanding most complicated concepts had led them to be more attentive in classroom, to ask the teacher for more clarification, to study more in their textbooks at home than usual (see Table 5). These claims agree with data in Vollmer et al. (2006), who underline how difficulties in foreign language, far from leading students to abandon their studies, instead cause them to work more persistently, strengthening higher cognitive ability. All these behaviours account for an increased motivation for the study of Physics, although beforehand the CLIL students were not particularly motivated for this subject, nor for English, as self-declared on the background questionnaire. Indeed, the literature highlights how CLIL fosters motivation, since it introduces a change in teaching methodology and, often, a "democratisation of classroom practices" (Banegas 2012, 42). The interactional hierarchy between teacher and pupils is frequently renegotiated, in that students are asked to actively contribute to co-construct knowledge through presentations to the class, group activities and self-evaluation. The reconfiguration of interactional roles around content nurtures students' attention and participation, so that a more dynamic interaction takes place in classroom. It is not a coincidence that another beneficial effect of CLIL is the increase in learner autonomy, as they become more collaborative and context-responsive. This is a consequence of the learner-centred teaching approach usually adopted in CLIL lessons (Gerena 2012). The renegotiation of teacher and student roles is also encouraged by participants' perception of a greater tolerance than usual for language and content mistakes, due to frequent linguistic difficulties experienced by students as well as by the teacher (Costa 2012). The findings of the current research show that such an increased willingness of learners to participate leads them both to better assimilate and retain content-specific knowledge, not only at an individual level but also at a class level. This means that not only clever pupils, but indeed most of the students in the class engage in Physics more in CLIL than in traditional teaching. Motivation, which is a well recognised influential benefit of CLIL upon language learning, appears to positively affect content learning too.

In the questionnaire, together with students' statements testifying to a rise in motivation, significant concern about content-specific outcomes emerges from pupils, who put emphasis on the difficulties they encounter in the use of a foreign language as the medium of instruction for Physics. In particular, learners label as problematic the comprehension of new spe-

cialised terminology, strictly intertwined with the new concepts to which it refers, according to the results of previous qualitative analysis of Italian students' voices (Coonan 2009). Prior surveys on student evaluation of CLIL also indicate that learners tend to underestimate their capacity to study content-specific issues in a foreign language and perceive themselves as less competent in both language and content than traditionally schooled peers (Seikkula-Leino 2007; Mezek et al. 2015). This underestimation may be determined by difficulties and failures experienced in CLIL environments and does not coincide with actual achievements of CLIL students, who tend to outperform non-CLIL students. As a matter of fact, while adolescents of the same geographical area of the present research identify CLIL as a menace for successful learning of school subjects, because of a lack of teacher and student language proficiency (Di Martino 2015), our findings provide significant evidence for good achievements in Physics of students experiencing CLIL in a situation very close to that of investigated adolescents in Di Martino (2015). A possible objection to our results is that simplification of content was realised in the investigated CLIL class. This objection is fairly rejected by the complete equivalence of the three tests given to both the experimental and control groups. The outperformance of CLIL students in M1 and M2 entails that in the CLIL class the teacher did not simplify content nor reduce demands on students. Moreover, outperformance in M2, encompassing learners' productions in Italian, suggests that CLIL students did not possess a limited mastery of Physics-specific terminology and academic language in L1. They exploited both technical terms acquired in Italian before the CLIL TU and new terminology acquired during CLIL lessons in English, often provided by the teacher in Italian too. In this regard, it is worth pointing out that students were enrolled in CLIL only for 5 weeks and exclusively with regard to one content-specific topic, whereas prior Physics knowledge was built on through L1. Thus, our results need to be compared with findings of research on longer CLIL experience, so as to verify if the same trend occurs in educational contexts where CLIL is systematically adopted, such as in the last year of high school.

6 Concluding Remarks

The research intended to add a piece to the puzzle of understanding overall learning processes taking place in CLIL, by analysing the effects of CLIL upon content-specific outcomes and the role played by the factor of motivation. The assessment of content-specific achievements in Physics was measured through tests specifically developed for the study by disciplinary teachers involved in the research. Findings show positive effects of CLIL,

in that CLIL students outperform peers instructed in L1 for both measurements taken into consideration: the selection of appropriate specific issues for the given context (M1) and content-related argumentative skills (M2). The two classes started at the same level of content-specific competence, as documented by the pretest, but immediately after CLIL instruction, the experimental group scored better than the control one and even better 5 weeks later. These findings account for good short-term achievements and even better long-term attainments in CLIL learning. A further benefit of CLIL which emerged in analyses is the major reduction of disparity among individual students in the experimental group immediately after the CLIL instruction and, to a larger extent, in the long run. Given such positive results, in line with previous research on Physic-specific learning (Jäppinen 2005; Haagen-Schützenhöfer, Hopf 2014; Canlas 2016) as well as on other disciplines (Vollmer et al. 2006; Murray 2010; Gregorczyk 2012; Bauer-Marschallinger 2016), CLIL appears not only to be an effective language teaching methodology, but a "genuine tool for educational innovation" as well (Van de Craen, Surmont 2017, 25).

Students in the investigated class (Year 11) are two years younger than students for whom in Italy enrolment is compulsory in a CLIL programme (Year 13). If this last aspect is regarded and combined with findings of the present study, pedagogical implications seem straightforward: the sooner a CLIL programme is implemented the better. The undeniable difficulties in studying a non-linguistic subject in a foreign language appear to be compensated for by increased motivation, entailed by the change of ordinary teaching methodology towards more interactive and stimulating classroom practice. As Italian students believe, the most relevant conditions for CLIL success are "on the one hand, the individual teacher's ability to impact his/her students' learning, on the other the student's own effort to progress" (Di Martino 2015, 85). In our findings, on the one hand, teacher effectiveness is raised by her enhanced awareness of language as medium of instruction and motivation, on the other students are more inclined to commit extra time to studying in order to understand content, as shown in results (M1 and M2) and spontaneously declared on a feedback questionnaire (M3). Therefore, motivation turns out to be a central benefit of CLIL also upon content-specific learning, as revealed for language-specific learning (Coyle, Hood, Marsh 2010; Lasagabaster 2011; Banegas 2012) even in disadvantaged educational settings (Gerena 2012). It is noteworthy here that the investigated school is not in a privileged educational context. The school is located in a small town in Southern Italy, characterised by a not particularly wealthy socio-economic-cultural context, and the students were focused on Humanities more than on English or Physics. As a matter of fact, English proficiency was not high in students of either class, as self-declared on the background questionnaire and confirmed by the teacher, nor was competence in Physics, as demonstrated by overall test scores

(see Tables 3 and 4). Given the situation, these positive effects of CLIL documented with not well predisposed students in a non-favoured area are important and call for more space for CLIL in Italian educational policies.

Since the present study focuses on the differences between CLIL and non-CLIL classes, only group analysis was carried out and individual differences in students' performance were not investigated. Further analyses of learner characteristics, such as individual proficiency and motivation for either English or Physics, are currently in progress, in order to examine which learners take advantage of CLIL to the greatest extent. In parallel with deepening analyses of collected data, broadening the sample is recommended. Findings of the present research allow to draw attention to assimilation and retention of content-specific competence in one CLIL class compared to one traditional class with respect to one subject. In order to generalise the results, it is crucial to gather further data, by involving different types of high schools and different disciplines.[12] In addition, the possibility to compare more than one set of CLIL and non-CLIL classes may reduce the risk that results are affected by a different personal style of teaching. In the present study, the personal pedagogical style of the two teachers did not lead the two classes to different levels of competence at the pretest before the beginning of CLIL experiment, rather differences between the classes emerged only after the CLIL intervention. Thus, it can be reasonably assumed that the influence of personal style on content-specific learning in the present case is negligible. However, the comparison of several pairs of CLIL and non-CLIL classes with different teachers may shed more light on the role of teachers personal style of teaching. Finally, future studies are called for to comprehend another aspect as well. The positive effect revealed for CLIL may be due to the fact that the experimental group started to experience CLIL as the research began. Since the increment of motivation could be stimulated by the novelty of a different teaching methodology, longer CLIL experiences have to be investigated to verify whether the motivational benefit of CLIL upon content-specific learning is retained over time. Such future steps are profitable in order to bridge the gap between the comprehension of language-specific and content-specific learning in the present-day educational challenge of CLIL.

12 It would be particularly important to compare sciences and humanities, since they are characterised by different knowledge construction models, according with Bernstein (1999) who distinguishes between hierarchical and horizontal disciplines. The former, such as natural sciences, construct knowledge by integrating new issues to previously acquainted issues in a higher-ranking level of abstraction. The latter, such as Humanities, progress mainly by means of the accumulation of new information.

Acknowledgments

I wish to thank Professors Rosa Labadia and Maria Buonocore of the "Alberto Galizia" High School in Nocera Inferiore (Salerno), for preparing tests and teaching unit and for carrying out the experiment in their classes; Professor Alessio De Giorgi for evaluating the students' answers to the tests; and Gioacchino Amato for support in data collection and transcription. I'm also grateful to the two anonymous reviewers for their comments and suggestions.

References

Aiello, Jacqueline; Di Martino, Emilia; Di Sabato, Bruna (2017). "Preparing Teachers in Italy for CLIL: Reflections on Assessment, Language Proficiency and Willingness to Communicate". *International Journal of Bilingual Education and Bilingualism*, 20(1), 69-83.

Airey, John (2012). ""I Don't Teach Language": The Linguistic Attitudes of Physics Lectures in Sweden". *AILA Review*, 25, 64-79.

Banegas, Darío L. (2012). "Motivation and Autonomy through CLIL. A Collaborative Undertaking". Anglada, Liliana B.; Banegas, Darío L. (eds.), *Views on Motivation and Autonomy in ELT = Selected Papers from the XXXVII FAAPI Conference* [e-book]. Bariloche: APIZALS, 39-45.

Bauer-Marschallinger, Silvia (2016). "The Acquisition of Historical Competences in the CLIL History Classroom". *ÖGSD Tagungsberichte*, 1, 3-6.

Bernstein, Basil (1999). "Vertical and Horizontal Discourse: an Essay". *British Journal of Sociology Education*, 20, 157-73.

Canlas, Ian P. (2016). "The Use Of Content And English Language Integrated Learning (CELIL) Methodologies In Teaching Selected Concepts In Physics". *International Journal of Scientific and Technology Research*, 5(10), 154-9.

Cenoz, Jasone; Genesee, Fred; Gorter, Durk (2014). "Critical Analysis of CLIL: Taking Stock and Looking Forward". *Applied Linguistics*, 35(3), 243-62.

Coonan, Carmel M. (2009). "Opportunità di usare la LS nella funzione CLIL: importanza, problemi, soluzioni". *Studi di Glottodidattica*, 2, 20-34.

Costa, Francesca (2012). "Focus on Form in ICLHE Lectures in Italy: Evidence from English-Medium Science Lectures by Native Speakers of Italian". *AILA Review*, 25, 30-47.

Costa, Francesca; Mariotti, Cristina (2017a). " Differences in Content Presentation and Learning Outcomes in English-Medium Instruction (EMI) vs. Italian-Medium Instruction (IMI) Contexts". Valcke, Jennifer; Wilkinson, Robert (eds.), *Integrating Content and Language in Higher*

Education: Perspectives on Professional Practice. Frankfurt am Main: Peter Lang, 187-204.

Costa, Francesca; Mariotti, Cristina (2017b). "Students' Outcomes in English-Medium Instruction: Is There Any Difference Related to Discipline?". *L'analisi linguistica e letteraria*, 25(2), 361-71.

Coyle, Do; Hood, Philip; Marsh, David (2010). *CLIL Content and Language Integrated Learning*. Cambridge: Cambridge University Press.

Cummins, Jim (1979). "Cognitive/Academic Language Proficiency, Linguistic Interdependence, the Optimum Age Question and Some Other Matters". *Working Papers on Bilingualism*, 19, 121-9.

Dallinger, Sara; Jonkmann, Kathrin; Hollm, Jan; Fiege, Christiane (2016). "The Effect of Content and Language Integrated Learning on Students' English and History Competences – Killing Two Birds with One Stone?". *Learning and Instruction*, 41, 23-31.

Dalton-Puffer, Christiane (2008). "Outcomes and Processes in Content and Language Integrated Learning (CLIL): Current Research from Europe". Delanoy, Werner; Volkmann, Laurenz (eds.), *Future Perspectives for English Language Teaching*. Heidelberg: Carl Winter, 139-57.

Dalton-Puffer, Christiane (2011). "Content-and-Language Integrated Learning: From Practice to Principles?". *Annual Review of Applied Linguistics*, 31, 182-204.

Dalton-Puffer, Christiane; Nikula, Tarja (2015). "Content and Language Integrated Learning", Special Issue of *Language Learning*, 43(3).

Di Martino, Emilia (2015). "Content and Language Integrated Learning: la voce dei più giovani. Resoconto di un'indagine sulle preoccupazioni e le attese degli studenti". *EL.LE*, 4(1), 85-114. DOI 10.14277/2280-6792/125.

Doiz, Aintzane; Lasagabaster, David; Sierra, Juan M. (2014). "CLIL and Motivation: The Effect of Individual and Contextual Variables". *The Language Learning Journal*, 42(2), 209-24.

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

Fernández-Sanjurjo, Javier; Fernández-Costales, Alberto; Arias Blanco, José M. (2017). "Analysing Students' Content Learning in Science in CLIL vs. non-CLIL Programmes: Empirical Evidence from Spain". *International Journal of Bilingual Education and Biligualism*, 1-14. DOI 10.1080/13670050.2017.1294142.

Gerena, Linda (2012). "Empowering English Language Learners to Build Autonomy and Motivation in Urban Secondary Math and Science Classes: the Peer Enabled Restructured Classroom (PERC)". Anglada, Liliana B.; Banegas, Darío L. (eds.), *Views on Motivation and Autonomy in ELT = Selected Papers from the XXXVII FAAPI Conference* [e-book]. Bariloche: APIZALS, 27-38.

Gregorczyk, Beata (2012). "An Empirical Study on the Acquisition of Content in a CLIL-Based Chemistry Course: a Preliminary Report". *Latin American Journal of Content & Language Integrated Learning*, 5(1), 9-32.

Haagen-Schützenhöfer, Claudia; Hopf, Martin (2014). "Content and Language Integrated Learning in Physics Teaching: Benefits, Risks, Requirements and Empirical Studies". Kaminski, Wanda; Michelini, Marisa (eds.), *Teaching and Learning Physics Today: Challenges? Benefits?* Pasian di Prato (Udine): Lithostampa, 1-8.

Infante, Debora (2010). *Il Content and Language Integrated Learning (CLIL) in Italia, Modelli didattici e sperimentazioni nella scuola primaria*. Roma: Editrice Nuova Cultura.

Jäppinen, Aini 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), 147-68.

Lasagabaster, David (2011). "English Achievement and Student Motivation in CLIL and EFL Settings". *Innovation in Language Learning and Teaching*, 5(1), 3-18.

Lim Falk, Maria (2008). *Svenska i engelskspr°akig skolmilj¨o. ¨Amnesrelaterat spr°akbruk i tv°a gymnasieklasser* (Swedish in an English Classroom Environment. Language Use in Two Grammar School Classes). Stockholm, Sweden: Eddy.

Leone, Andrea R. (2015). "Outlooks in Italy: CLIL as Language Education Policy". *Working Papers in Educational Linguistics*, 30(1), 43-63.

Lo, Yuen Y.; Macaro, Ernesto (2015). "Getting Used to Content and Language Integrated Learning: What can Classroom Interaction Reveal?". *The Language Learning Journal*, 43(3), 239-55.

Mezek, Spela; Pecorari, Diane; Shaw, Philip; Irvine, Aileen; Malmström, Hans (2015). "Learning Subject-Specific L2 Terminology: the Effect of Medium and Order of Exposure". *English for Specific Purposes*, 38, 57-69.

Murray, Diane R. (2010). "Irish-Medium Language Immersion Programs' Effect on Mathematics Education". *Journal of Mathematics Education at Teachers College*, 1, 28-32.

Nikula, Tarja (2017). "Emerging Themes, Future Research Directions". Llinares, Ana; Morton, Tom (eds.), *Applied Linguistics Perspectives on CLIL*. Amsterdam: Benjamins, 307-12.

Piesche, Nicole; Jonkmann, Kathrin; Fiege, Christiane; Keßler, Jörg U. (2016). "CLIL for All? A Randomised Controlled Field Experiment with Sixth-grade Students on the Effects of Content and Language Integrated Science Learning". *Learning and Instruction*, 44, 108-16.

Pigliapoco, Erika; Bogliolo, Alessandro (2009). "CLIL Process Monitoring in Higher Education: Instruments and Results". Sisti, Flora (ed.), *CLIL Methodology in University Instruction: Online and in the Classroom. An Emerging Framework*. Perugia: Guerra, 79-88.

Ricci Garotti, Federica (2017). "Subject Knowledge in CLIL: Which Factors Contribute to Positive Results?". Coonan, Carmel M.; Favaro, Luciana; Menegale, Marcella (eds.), *A Journey Through the Clil Landscape: Problems, Prospects*. Cambridge: Cambridge Scholars Publishing, 55-68.

Roquet, Helena; Perez-Vidal, Carmen (2017). "Do Productive Skills Improve in Content and Language Integrated Learning Contexts? The Case of Writing". *Applied Linguistics*, 38(4), 489-511.

Seikkula-Leino, Jaana (2007). "CLIL Learning: Achievement Levels and Affective Factors". *Language and Education*, 21(4), 328-41.

Serra, Cecilia (2007). "Assessing CLIL at Primary School: a Longitudinal Study". *The International Journal of Bilingual Education and Bilingualism*, 10(5), 582-602.

Sisti, Flora (2017). "CLIL in Higher Education: What if…?". Coonan, Carmel M.; Favaro, Luciana; Menegale, Marcella (eds.), *A Journey Through the Clil Landscape: Problems, Prospects*. Cambridge: Cambridge Scholars Publishing, 175-90.

Surmont, Jill; Struys, Esli; Van Den Noort, Maurits; Van de Craen, Piet (2016). "The Effects of CLIL on Mathematical Content Learning: a Longitudinal Study". *Studies in Second Language Learning and Teaching*, 6(2), 319-37.

Van de Craen, Piet; Ceuleers, Evy; Mondt, Katrien (2007). "Cognitive Development and Bilingualism in Primary Schools: Teaching Maths in a CLIL Environment". Marsh, David; Wolff, Dieter (eds.), *Diverse Contexts – Converging Goals: CLIL in Europe*. Frankfurt am Main: Peter Lang, 185-200.

Van de Craen, Piet; Surmont, Jill (2017). "Innovative Education and CLIL". *Research Papers in Language Teaching and Learning*, 8(1), 22-33.

Vollmer, Helmut J.; Heine, L.; Troschke, R.; Coetzee, D.; Kuttel, V. (2006). "Subject Specific Competence and Language Use of CLIL Learners: the Case of Geography in Grade 10 of Secondary Schools in Germany". Paper presented at the ESSE8 Conference, London.