Macrothink Institute™

International Journal of Linguistics
ISSN 1948-5425
2023, Vol. 15, No. 6

The Role of Motivation in Content-specific Learning in CLIL: Core vs Non-core Subjects (Note 1)

Fabiana Rosi (Corresponding author)

Dipartimento di Lettere e Filosofia, Università di Trento

Via T. Gar 14, 38122 Trento, Italy

E-mail: fabiana.rosi@unitn.it

Gioacchino Amato

Dipartimento di Lingue, Letterature e Culture Moderne, Università di Bologna

Via Cartoleria 5, 40124 Bologna, Italy

Alessandra Zappoli

Dipartimento di Formazione, Lingue, Intercultura, Letterature e Psicologia

Università di Firenze

Via Santa Reparata, 93-95, 50129 Firenze, Italy

Received: November 16, 2023 Accepted: December 19, 2023 Published: December 25, 2023

doi:10.5296/ijl.v15i6.21559 URL: https://doi.org/10.5296/ijl.v15i6.21559

Abstract

The paper investigates the effects of CLIL on the acquisition of content-specific competence in Italian high-school students. Two educational contexts are examined: Chemistry classes (CLIL and non-CLIL) in a science-oriented high school and Physics classes (CLIL and non-CLIL) in a humanities-oriented high school. The two subjects share many epistemological features but have a different status (core vs non-core subject) for students in the two educational contexts. Two types of content-specific competence are measured: receptive disciplinary knowledge (M1) and productive argumentation skill (M2). Findings show that CLIL students do not underperform non-CLIL peers in both disciplines, rather CLIL students of the non-core subject outperform the control group, especially for M2. Furthermore, a Multiple Factor Analysis illustrates that M1 and M2 results can be predicted

by considering the individual variability of students' motivation toward the discipline, CLIL, and Foreign Language Learning. The motivation toward the subject emerges as the trigger factor for CLIL benefit but plays an unexpected role: it is inversely proportional to content-specific learning. Implications for CLIL studies and educational policy are discussed.

Keywords: CLIL, Motivation, Affective factors, Content-specific competence

1. Introduction

In the last two decades, many studies have focused on Content and Language Integrated Learning (CLIL), the educational approach consisting in teaching non-linguistic disciplines in a language other than the students' first language, usually a foreign language (FL) (Marsh & Langé, 1999; Coyle *et al.*, 2010). The CLIL approach has received close attention from a theoretical perspective, as well as garnering a large consensus from policymakers, teachers, and families (Santipolo, 2021). Within the European context, one of the reasons for such a boost in theoretical and applied interest in the topic can be found in the recognition of the potential of CLIL to meet the European Commission and Council of Europe requirements of educating multicultural and multilingual citizens in all member countries (Eurydice, 2006). The key factor of CLIL effectiveness is the real, rather than simulated, integration of two educational focuses: language and content (Coyle, 2006). Students learn the content through the FL and, through this learning process, they improve language skills. Success in content acquisition is attained only when FL competences are well developed, and language acquisition occurs only when content learning is successful. Thus, the two components are intertwined in all teaching phases: teacher explanations, interactive discussions, student activities, corrections, at-home study, and assessment.

The literature examines CLIL from multiple points of view. Regarding the acquisition of FL, researchers converge on the positive effects of CLIL on student competence (for an overview, see Dalton-Puffer, 2011), as students experiment with a more extensive, variegated, and cognitively demanding use of the target language. In turn, the influence of CLIL on content-specific learning has received less attention for several reasons, including the lack of a standard for evaluating content acquisition in CLIL and the need for complex collaboration between experts of FL and non-linguistic subject(s). Nevertheless, it is not surprising that current reflections on steps forward for CLIL (Cenoz *et al.*, 2014; Nikula; 2017) call for more investigation on content-specific learning, as this component is crucial for students' achievements in CLIL and for a better understanding of the whole educational process.

The literature on CLIL has, to date, documented controversial findings about the acquisition of content in CLIL. Some studies point out that CLIL students obtain positive results (Serra, 2007; Van de Craen *et al.*, 2007; Gregorczyk, 2012; Canlas, 2016; Surmont *et al.*, 2016), while others show negative effects (Lim Falk, 2008; Dallinger *et al.*, 2016; Piesche *et al.*, 2016; Fernández-Sanjurjo *et al.*, 2019; Virdia, 2022), and still others show no significant differences between CLIL students and their traditionally-educated peers (Seikkula-Leino, 2007; Infante, 2010; Haagen-Schützenhöfer & Hopf, 2014; Madrid & Barrios, 2018). These results fall within many different subjects taught in CLIL: Mathematics (Seikkula-Leino, 2007; Serra, 2007; Van de Craen *et al.*, 2007; Murray, 2010; Surmont *et al.*, 2016),

Chemistry (Gregorczyk, 2012), Physics (Haagen-Schützenhöfer & Hopf, 2014; Canlas, 2016), Geography (Vollmer *et al.*, 2006), and History (Bauer-Marschallinger, 2016).

To the best of our knowledge, no study has yet concentrated on the comparison among students' outcomes in different subjects, particularly when the different subjects play a different role in the students' educational pathways, as in the case of core versus non-core subjects. Core subjects constitute the foundation disciplines for students' school careers, while non-core subjects serve a complementary role within students' education. Consequently, learners do not have the same degree of motivation toward a core and a non-core discipline. Indeed, in the Italian context students choose a high school depending on the core subjects they wish to study, so they are more motivated toward core disciplines than non-core ones.

The subdivision in core vs. non-core subjects is particularly relevant for CLIL in Italian education for a number of reasons. The Italian Ministry of Education (Law 53/2003 and implementing decrees of 2010) introduced CLIL as mandatory for students attending the last year of high school and defined that non-linguistic discipline teachers, instead of language teachers, are in charge of its delivery (Note 2). In each school, the principal, in agreement with the board of teachers of the single class, decides which disciplines are taught in CLIL and, consequently, which teachers can attend the ministerial training course for CLIL teachers (Note 3). The decision depends on the teacher's competence in the FL, which is English in most cases, and on the teacher's interest in carrying out a CLIL programme. When more than one teacher has the required competence in FL and expresses an interest in the CLIL programme, the principal usually grants priority to the teacher of a core subject, such as a social science subject within a humanities-oriented high school or a natural science subject within a science-oriented high school. Consequently, Italian students frequently attend CLIL classes in a subject toward which they have a high motivation, interest, and expectation.

The current study aims to ascertain whether the application of CLIL to a core subject rather than a non-core subject is conducive for the acquisition of disciplinary content. The study compares CLIL and non-CLIL students' content-specific learning in two different disciplines: Chemistry in a science-oriented high school with a Chemistry specialisation (core subject) and Physics in a humanities-oriented high school (non-core subject). Chemistry and Physics share many epistemological aspects, such as the use of mathematical symbols, an experimental approach, activities in the laboratory, and both are conceived as ideal subjects for CLIL because of the purported reduced relevance of language in content comprehension and acquisition, as compared with humanistic disciplines such as literature and philosophy.

However, these two subjects display a different status in the two investigated educational contexts. In the science-oriented high school, Chemistry is a leading subject, since it is crucial for students' careers; in the humanities-oriented high school, Physics is a non-core subject, as learners strive to study subjects such as Pedagogy and Psychology, not Physics. The impact of the status of the discipline taught through CLIL on students' attainment is specifically examined by adopting Linear Mixed-Effects Models and post-hoc comparisons of each group. In addition, the individual variability of the participants concerning their

attitude toward the content and the FL has been investigated constructing five clusters of students through the Multiple Factor Analysis (MFA). The MFA considered four affective variables: motivation toward the subject, toward English as a FL (EFL), toward CLIL, and students' self-evaluation about personal performance achieved in the CLIL teaching unit experienced. The student sample has been monitored in three elicitations and in-depth statistical analysis guarantees the validity of the data.

2. Motivation in CLIL

Motivation is a complex psychological construct generally considered one of the most triggering factors for learning (Paris & Turner, 1994; William & William, 2011). Motivation is crucial for learning as it leads students to pay attention, participate actively in lessons, ask for clarifications when they do not understand, and strive to study at home. Its effect is not direct: motivation influences behaviour which, in turn, influences study achievements.

A basic distinction arises between intrinsic and extrinsic motivation: intrinsic motivation relies on inherent personal interest and enjoyment; extrinsic motivation is nurtured by the desire to gain an external outcome (Ryan & Deci, 2000). For example, learners demonstrate intrinsic motivation when they study a subject because they like it and enjoy it and, by contrast, display extrinsic motivation when they want to reach a specific and concrete goal, such as a good evaluation.

Concerning FL learning, in Dörnyei's (2005) theory, motivation is a dynamic process that interacts with other individual affective variables, such as language aptitude, learning strategies, personality, effort, and anxiety. Its internal structure comprises several components (Csiźer & Dörnyei, 2005): integrativeness, instrumentality, cultural interest, vitality of the FL community, linguistic self-confidence, *milieu*. Integrativeness deals with the desire to become part of the culture of the FL, while instrumentality concerns the concrete benefits of FL proficiency, such as job prospects or enhanced educational opportunities. These two motivational factors are probably the most investigated in FL research. Cultural interest deals with the admiration for the FL culture and occurs especially in learning contexts where the contact with FL speakers is minimal. Vitality of the FL community refers to the perceived prestige of FL community, including its political, social and economic status, and the reputation of the FL-speaking countries. Linguistic self-confidence reflects self-perception of the potential for a learner's success in FL learning. Finally, *milieu* encompasses the family's or friends' attitude toward FL learning and their support to the learner.

In FL learning, motivation toward FL correlates, to a relevant extent, with language attainments, as pointed out by the large-scale meta-analysis in Masgoret and Gardner (2003). Furthermore, language learning, in comparison with learning more generally, is thought to be particularly affected by motivational factors given the relevance of language in the construction of identity (Gardner & Lambert, 1972; Ushioda & Dörnyei, 2009).

Motivation is also deemed to be one of the main benefits of CLIL (Coyle, 2006). Positive and direct correlation between motivation toward EFL and EFL achievements in CLIL is reported by Lasagabaster (2011) within a sample of 191 Basque secondary-school students. The study

reports that students develop better EFL competence and higher language-learning motivation within the CLIL programme rather than in traditional EFL contexts. Namely, CLIL students display a higher degree of effort, interest, instrumental motivation, and positive attitude toward EFL. Advantageous CLIL effects on student motivation toward EFL are confirmed by Huang (2011): Taiwan primary-school pupils are more active, willing to volunteer and produce output during CLIL lessons than during EFL classes. Positive CLIL effects on EFL motivation emerge also in Basque primary-school students: CLIL enhances students' motivation more than other teaching methodologies, such as the use of books and project work (Lasagabaster & López Beloqui, 2015). Furthermore, the large-scale study by De Smet et al. (2018), examining the language attitudes and motivation of 986 CLIL and non-CLIL students in Belgium primary and secondary schools, where CLIL is carried out either in English or in Dutch, shows that CLIL students have better language attitudes and higher motivation than non-CLIL students, especially CLIL students educated in English and attending secondary schools. It is worth underlining that the literature signals that CLIL enhances motivation not only in students, but also in teachers with respect to traditional learning (Coyle, 2006).

As for the causes of CLIL beneficial influence on motivation, several studies (Seikkula-Leino, 2007; Marsh *et al.*, 2008; Várkuti, 2010; Doiz *et al.*, 2014; Rumlich, 2014) report that CLIL fosters both intrinsic and instrumental motivation, because of the linguistic improvement perceived by the students and the prestige of English, the FL mainly involved in CLIL, for students and families. This could be linked to the enthusiastic consensus CLIL received from educational policymakers, language practitioners, and parents in many countries. Other studies (Coonan, 2007; 2012; Dalton-Puffer *et al*., 2009; Lasagabaster & Sierra, 2009; Ricci Garotti, 2017) underscore that motivation is supported by the environmental features CLIL entails: authentic and stimulating learning material, learner-centred approach, strong sensitivity of the teacher to students' needs.

Unlike motivation toward FL learning or CLIL, little attention has so far been paid to motivation toward disciplinary content (Rumlich, 2014). An exception is Lasagabaster and Doiz (2015) who include this aspect in their analysis of CLIL effects on different affective factors and motivation toward EFL learning. The authors analyse answers to questionnaires longitudinally collected from 304 CLIL and non-CLIL students in Spanish secondary schools over two or three years. The analysis reveals that CLIL students demonstrate a higher motivation toward content than their traditionally- educated peers and that in CLIL, students' motivation toward subject remains constant over time.

The present study aims to shed light on the influence of motivation toward core and non-core disciplines upon the acquisition of disciplinary content and on the interplay among motivation toward disciplines, CLIL, and Foreign Language in the learning process.

3. The Study

3.1 Research Questions

The study aims to answer the following research questions:

RQ1: Do CLIL students acquire content better than their traditionally-educated peers?

RQ2: Is there any difference between the acquisition of content in CLIL applied to core or non-core subjects?

RQ3: Does motivation toward the subject influence the acquisition of disciplinary competence?

3.2 Participants

Overall, 64 students of four third-year classes in Italian high schools took part in the study (Note 4). Specifically, 31 students (10 F, 21 M; mean age = 16.6) attended CLIL lessons and 33 students (23 F, 10 M; mean age = 16.5) constituted the control group and attended traditional lessons in the first language. For each group (CLIL and non-CLIL) two classes were examined in two different high schools in the Salerno region in the South of Italy: a science-oriented high school with Chemistry specialisation and a humanities-oriented high school. Within the science-oriented high school, 15 students attended CLIL lessons and 15 students attended traditional lessons in Chemistry, which is a core subject for this type of school; within the humanities-oriented high school, 16 students attended CLIL lessons and 18 students attended traditional lessons in Physics, which is a non-core subject in this context. In total four groups of students participated in the study: CLIL_core, CLIL_non-core, no_CLIL_core, no_CLIL_non-core. All learners were Italian-native speakers with an A2-B1 level of CEFR (Council of Europe 2018) in EFL.

3.3 Procedure

All students were tested through three written questionnaires in Italian, administered in three elicitation sessions: before the CLIL Teaching Unit (TU) began (T1), at the end of the CLIL TU (T2), and five weeks later (T3). The CLIL TUs in the core and non-core subjects were run in EFL by the regular teachers of Chemistry and Physics. Both teachers were women aged between 55 and 60, with similar teaching experience at school. Just before the research started, both teachers passed the C1 Cambridge Advanced English Examination and the methodological training course for CLIL teachers.

The TUs were, for both classes, students' first experience of CLIL. The CLIL TU lasted five weeks and focused on a topic that was part of the ministerial syllabus: UV-Vis Spectrophotometer for Chemistry and Newton's laws of dynamics for Physics. In both the experimental (CLIL) and control (non-CLIL) classes, the same TU was presented by means of teachers' frontal explanations, experiments in the laboratory, and interactive classroom activities, such as cooperative content review and exercises: the only difference in teaching in the CLIL and non-CLIL classes was the language of instruction (English vs Italian). As a matter of fact, for years, the teachers of the CLIL and non-CLIL classes had been working in parallel in their respective classes, sharing curricula, materials, and techniques, and they continued to do so in the monitored TUs. Indeed, the CLIL and non-CLIL classes showed analogous levels in Chemistry and Physics at T1 before the TU started. In all four classes, the TU was carried out in the regular timetable by the teacher who was normally in charge of the

subject: there were no differences, other than the language of instruction, with respect to the normal classroom environment.

The questionnaire handed out at T1 dealt with background on concepts of Chemistry or Physics, while at T2 and T3 two identical questionnaires about the TU topics were administered to investigate the short-term and long-term retention of the same content-specific competence (Note 5). The questionnaires, designed by the class-teachers, included two measurements of students' content-specific competence: multiple-choice questions (M1) and short explanations (maximum three lines) (M2) written by the students to justify the choice of their answer in M1. M1 scores measure receptive knowledge of disciplinary concepts taught in the TU; M2 scores gauge L1 productive argumentation skills related to the disciplinary topics. The questionnaires were composed of 10 multiple-choice questions (M1) followed immediately by 10 requests for explanation (M2). The 10 multiple-choice questions provided three answer options: a correct item, a foil one, and an incorrect one. The correct answers were assessed as 3, the foil as 2, the incorrect as 1, and 0 points were assigned in the case of no answer. The 10 explanations written by the students were assessed according to the following scale: 4 if correct and complete; 3 if correct but incomplete; 2 if not totally correct; 1 if incorrect; 0 if no answer was provided. The answers were blindly assessed by two external teachers, one for Chemistry, one for Physics. Researchers transcribed all data and forwarded them to the evaluators, who did not know whether they were assessing CLIL or non-CLIL students.

CLIL students were required to fill in two further questionnaires: the former was handed out at T1 and concerned their favourite school subject and their interest in the tested subject (Chemistry of Physics), in EFL and in CLIL approach; the latter was administered at T2 and regarded students' opinions about the CLIL TU experienced.

3.4 Data Treatment

To analyse the effects of CLIL on students' performance, Linear Mixed-Effects Models (LMM) (Bates *et al*., 2015) were run testing the relation between the questionnaires' scores and the following four predictors:

1) elicitation sessions (T1 vs T2 vs T3) considered as a numerical predictor
2) teaching approach (CLIL vs no_CLIL)
3) type of measurements (M1 vs M2)
4) status of core (for Chemistry) or non-core (for Physics) subject within the high-school programmes (core vs non-core).

The analysis considered a set of models going from the simplest one, including only random effects, to the most complex model considering the four predictors and their interaction. The best model was selected through the model selection procedure, which provides mathematical support for the choice of the simplest model (i.e. one that includes the fewest number of predictors according to the principle of parsimony) that best explains the data. Once the best

model was determined, post hoc multiple group comparisons were carried out, to highlight the significant differences between scores at the group level.

The second part of the analysis aimed to test the relation between CLIL students' scores, and the four categories of self-reported information collected through the surveys. Five clusters of participants sharing a homogenous profile (high, medium-high, neutral, medium-low, low) were constructed through MFA (Bécue-Bertaut & Pagès, 2008) by taking into consideration the following variables:

1) students' appreciation for the subject (subject_liking)
2) students' appreciation for EFL (English_liking)
3) students' appreciation for CLIL (CLIL_liking)
4) students' self-assessment of the benefits gained from the CLIL TU (self_assessment).

Finally, the analysis tested whether the students belonging to the five emerging clusters scored significantly differently when evaluated through M1 and M2, and which variable predicted higher scores.

4. Results

To guarantee the best approximation to the normal distribution, outliers exceeding the Interquartile range were excluded. The original dataset of 384 observations (192 in each measurement) resulted in 180 observations for M1 and 185 for M2, for a total of 365 observations.

4.1 Linear Mixed-Effects Models Analysis

A set of 11 models adopting the *lmer* function within the *lme4 package* (Bates *et al.*, 2015) on the *R* environment (R Core Team, 2020) were fitted. The predictors were individually tested as independent fixed effects and in interaction, by inserting first the elicitation sessions (T1, T2, T3), then the measurements (M1 or M2), the CLIL or non-CLIL teaching, and the status of core or non-core subject. The participant was inserted as a random effect in all models. The model comparison procedure indicates m11, the most complex model, in which all predictors interact with each other, as the best approximation to the real model describing our data (Residual Deviance = 2250.80; Degrees of Freedom (BIC index) = 18; BIC index = 2329.19; Chisq = 120.65; p. value < 0.001; $\eta^2 = 0.05$). Figure 1 shows the effect plot of m11 representing the interaction of the fixed predictors.

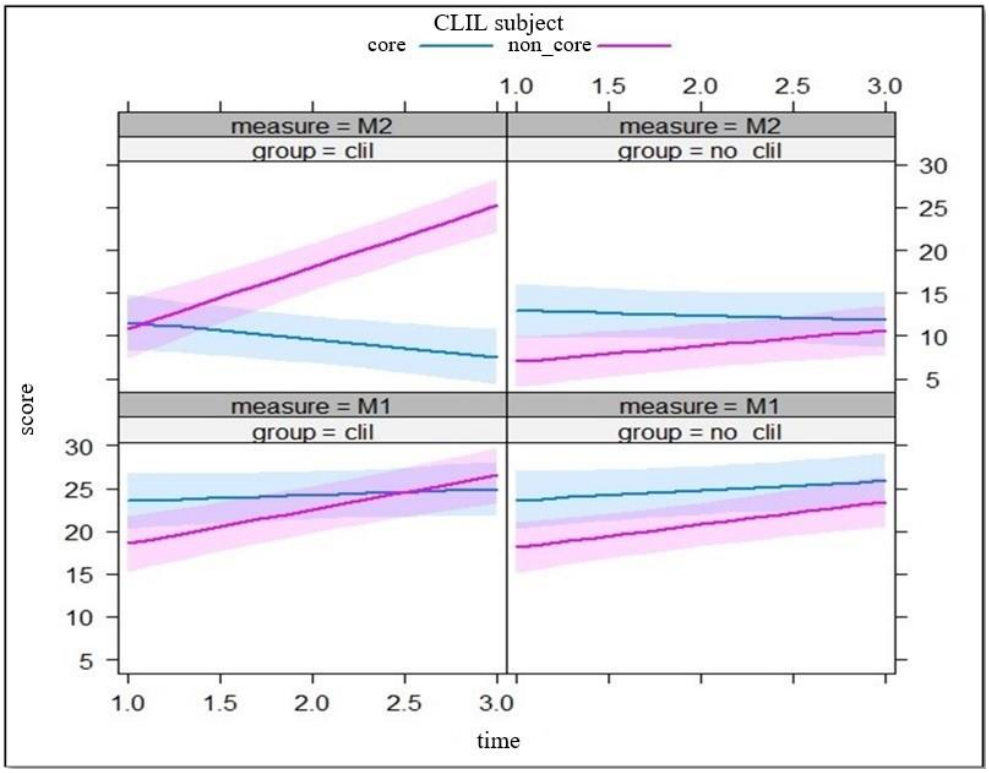

Figure 1. The scores of CLIL_core, CLIL_non-core, no_CLIL_core, no_CLIL_non-core groups at T1, T2 and T3 for either M1 or M2

Considering the M1 results, the post-hoc analysis performed on the m11 selected model, with False Discovery Rate (FDR) correction for multiple comparisons and using the *emmeans* package on R (Lenth *et al.*, 2018), reveals a significant improvement of CLIL_non-core students: their scores significantly increase from T1 to T2 ($Z = -4.63$; $p < 0.0001$), from T1 to T3 ($Z = -4.63$; $p < 0.0001$), and from T2 to T3 ($Z = -4.63$; $p < 0.0001$) with a comprehensive linear improvement in the scores over time. The significant difference of scores over time indicates that students improve their score because of the participation in the CLIL TU. The scores of the no-CLIL_non-core group also show a significant improvement, but only from T1 to T2 ($Z = -3.50$; $p = 0.001$) and from T1 to T3 ($Z = -3.50$; $p = 0,001$), while no significant difference emerged from T2 to T3. As for CLIL_core and no_CLIL_core students, the groups show constant scores over time as non-significant differences emerge for M1 in both groups between the three longitudinal tests (see the bottom part of Figure 1).

Regarding M2 results, the post-hoc analysis, with FDR correction for multiple comparisons, shows a significant linear trend of improvement of CLIL_non-core and no_CLIL_non-core. The CLIL_non-core group significantly improves M2 scores from T1 to T2 ($Z = -7.80$; $p < 0.0001$), from T1 to T3 ($Z = -7.80$; $p < 0.0001$), and from T2 to T3 ($Z = -7.80$; $p < 0.001$). The no_CLIL_non-core group significantly improves from T1 to T2 ($Z = -2.40$; $p = 0.021$), from T1 to T3 ($Z = -2.40$; $p = 0.021$), and from T2 and T3 ($Z = -2.40$; $p = 0.021$). In addition, the CLIL_non-core group scores significantly higher than the no_CLIL_non-core group at T2 ($Z = 4.87$ $p < 0.0001$) and at T3 ($Z = 6.70$; $p < 0.0001$). As for M2 scores in the core subject, the CLIL_core group shows a non-significant declining trend from T1 to T2 ($Z = 2.45$; $p =$

0.071), from T1 to T3 (Z = 2.45; p = 0.071), and from T2 to T3 (Z = 2.45; p = 0.071) (Note 6). These non-significant comparisons confirm that the performance is constant over time; however, the tendency of scores indicates that CLIL_core students risk lowering their M2 scores over time. In turn, M2 scores of no_CLIL_core group remain constant over time. Finally, considering only CLIL students, CLIL_non-core students scored significantly better than CLIL_core students at T2 (Z = -4.27; p < 0.0001) and at T3 (Z = -7.77; p < 0.0001).

4.2 Cluster Analysis of Student Motivation

The MFA and the cluster formation were run with the FactoMineR (Kassambara & Mundt, 2020) package on the R environment. The cluster formation is based on the four affective variables: subject_liking, English_liking, CLIL_liking, self_assessment (see section 3.4). Two participants were removed from the cluster formation procedure because of missing data; in total, the answers provided by 29 students attending the CLIL programme were considered. The first two generated dimensions (Dim1, Dim2), explaining 72 % of the variance, have been taken as input for Hierarchical Clustering on Principal Components (HCPC). The results of the HCPC provided the subdivision of the five clusters represented in Figure 2.

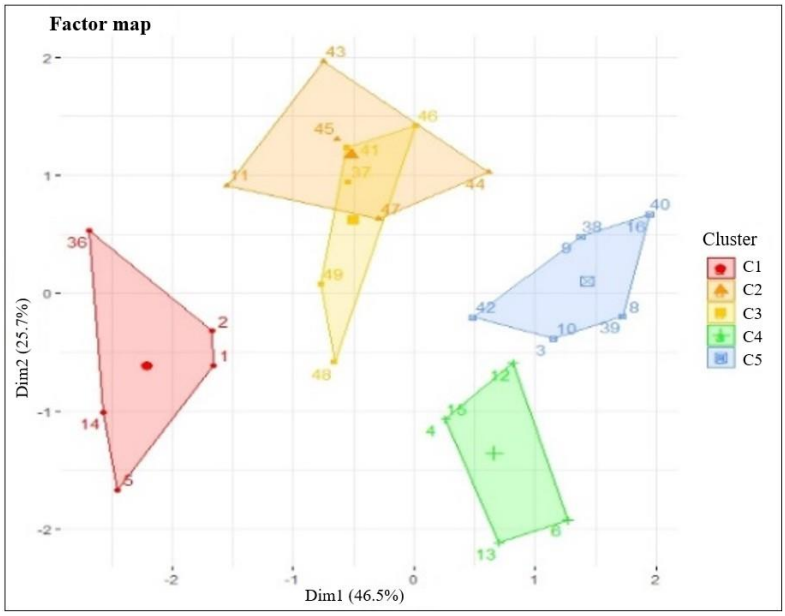

Figure 2. The five clusters of students generated by the HCPC analysis

Cluster 1 (C1) is composed of four CLIL_non-core students and one CLIL_core student with medium-low appreciation of all four variables. Cluster 2 (C2) consists of four CLIL_core students and one CLIL_non-core student expressing high appreciation for subject and EFL, low appreciation for CLIL and medium-low benefits from the CLIL TU. Cluster 3 (C3) comprises five CLIL_core students, with a high appreciation for the subject, a moderate appreciation for EFL, a medium-high appreciation for CLIL, and medium-low self-perceived benefits from the CLIL TU. Cluster 4 (C4) is made up of five CLIL_non-core students with medium-low appreciation for the subject but high appreciation for EFL, for CLIL and high self-perceived benefits from the CLIL TU. Finally, Cluster 5 (C5) is composed of four

CLIL_core students and five CLIL_non-core students with high appreciation for all the considered variables.

Table 1 shows the comparison among the five clusters for M1 and M2 aggregated.

Table 1. Mean score and standard deviation for the five clusters at T1, T2, T3 for M1+M2

Cluster	T1		T2		T3	
	mean	S.D.	mean	S.D.	mean	S.D.
C1	14.10	6.01	17.60	8.30	22.80	8.64
C2	14.00	10.13	13.90	10.83	15.00	12.00
C3	21.30	8.51	18.20	8.20	18.90	9.09
C4	13.40	4.35	17.90	3.31	27.10	5.09
C5	19.89	6.26	21.83	9.03	25.44	9.50

A set of four models adopting the *lmer* function within the *lme4 package* on the *R* environment is designed considering the participants as random factors and testing the relevance of the cluster predictor on the longitudinal score. Since only CLIL students are considered and subdivided into five clusters and that implies a reduced number of observations per cluster, the scores are analysed by aggregating CLIL_core and CLIL_non-core students' results as well as M1 and M2 results. The analysis examines the relevance of the motivation variables upon students' outcomes and experience, by profiling students who benefit more from CLIL and students with lower learning benefits and an unsatisfying experience.

The model comparison procedure highlights m3, which considers the interaction between belonging to a specific cluster and the predicted score over time, as the best approximation to the real model describing the data (Residual Deviance = 1207.60; Degrees of Freedom (BIC index) = 12; BIC index = 1240.34; Chisq = 14.42; p. value = 0.01; $\eta^2 = 0.01$).

Figure 3 shows the effect plot of m3, representing the different learning trajectories over time of students belonging to the five clusters.

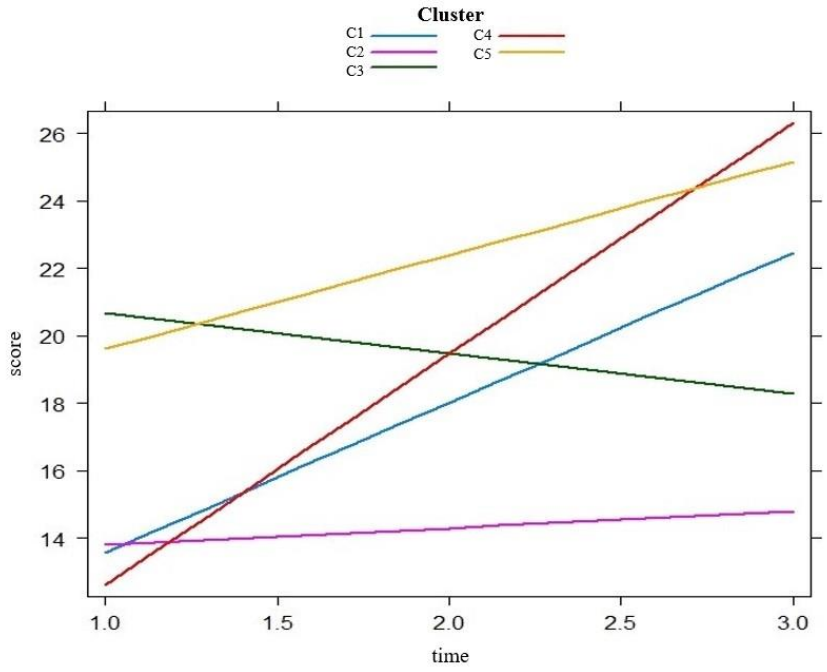

Figure 3. The interaction of the cluster and time predictor considering the overall score of CLIL students (M1+M2)

Post-hoc multiple comparisons, FDR corrected for multiple comparisons, show at T1 non-significant differences between the scores obtained from the students belonging to the five clusters. At T2, C5 students score significantly higher than C2 students ($Z = -3.057$; $p = 0.02$). At T3, C5 students score significantly higher than C2 students ($Z = -3.08$; $p = 0.01$) and show a non-significant tendency of scoring higher than C3 students ($Z = -2.052$; $p < 0.1$). In addition, at T3 C4 students score higher than C2 students ($Z = -3.02$; $p = 0.01$) and show a non-significant tendency of scoring higher than C3 students ($Z = -2.11$; $p < 0.1$). Finally, the multiple comparison analysis shows a non-significant tendency at T3 between C2 and C1 students ($Z = 2.03$; $p < 0.01$).

5. Discussion

The analysis shows that factors such as group (CLIL and non-CLIL), types of measurement (M1 and M2), and status of the subject (core and non-core) all interact with students' scores and play a role in the learning process. Before the CLIL TU, at T1, no significant differences between the CLIL and non-CLIL class occur, while at T2 and even more notably at T3, the CLIL classes perform differently from non-CLIL classes.

First, the effectiveness of CLIL emerges as different in the two types of content-specific competence examined in the study: receptive disciplinary knowledge and productive argumentation skill. The productive argumentation skill, necessary to write an explanation about disciplinary content (M2), is influenced by CLIL more than the receptive disciplinary knowledge needed to recognise the correct answer in multiple choice questions (M1). The students increase their scores after the CLIL TU, at T2 and T3, performing better in M2 than in M1. The two types of content-specific competence entail different cognitive operations and

linguistic abilities. M1 requires students to comprehend the text of the questions and of answer options, and to compare this with their own knowledge, in order to retrieve the relevant concepts and identify the answer option most suitable for them. M2 is more cognitively demanding than M1 since M2 involves deeper reasoning and conscious writing (Cisotto 1998). Indeed, M2 calls for a more complex process involving: a) the selection of the concepts most relevant for the given topic, based on students' knowledge; b) the reasoning about the connection among acquired concepts and the chosen answer option to be justified; c) the written explanation of their reasoning in students' own words. Thus, M2 questions are more difficult (Note 7) than M1 ones, as confirmed by the fact that both CLIL and non-CLIL students encounter more difficulties with M2 questions than with M1 ones and leave M2 questions blank more frequently than M1 ones, especially at T1.

The outcome that CLIL has a positive influence on M1 and that CLIL students respond adequately to the M2 task at T3 leads to a double insight: CLIL fosters in-depth content-specific proficiency and does not hinder students' mastery of the technical vocabulary and argumentative structure of the subject in their L1. Nonetheless, CLIL students of core and non-core disciplines display different results. In fact, according to an intergroup comparison, scores in the CLIL and non-CLIL classes are highly dependent on the distinction between the core subject (Chemistry in a science-oriented high school) and the non-core subject (Physics in a humanities-oriented high school). The two subjects share numerous features and are both part of the so-called hard sciences, but they play a crucially different role in science-oriented and humanities-oriented high schools. On the one hand, CLIL_core students display non-significant difference in comparison to no_CLIL_core students at T3, so the CLIL programme turns out to be neither disadvantageous nor beneficial, as some previous studies argued (Seikkula-Leino, 2007; Infante, 2010; Haagen-Schützenhöfer & Hopf, 2014; Madrid & Barrios, 2018). On the other hand, CLIL_non-core students show a linear improvement over time and outperform no_CLIL_non-core students, in accordance with studies which found a significant advantage of CLIL students over their peers attending regular classes (Serra, 2007; Van de Craen *et al.*, 2007; Gregorczyk, 2012; Canlas, 2016; Surmont *et al.*, 2016). CLIL emerges as a beneficial factor for the acquisition of content in the non-core subject, toward which students have low motivation and expectation, whereas CLIL shows no significant positive effects on the acquisition of content in the core subject, toward which students have high motivation and desire to gain solid expertise.

Students' motivation toward the subject appears to be the trigger factor for the effectiveness of CLIL, as also demonstrated in the MFA which profiles the sample in five clusters on the basis of four affective variables (subject liking, EFL liking, CLIL liking and self-evaluation of personal benefits gathered from the CLIL TU). According to the MFA, which aggregates M1 and M2 into comprehensive CLIL scores, the cluster with the best performance over time is C4, which corresponds to the only cluster entirely composed of students of the non-core subject. These students have medium-low appreciation for the subject, high appreciation for EFL and for CLIL, and high self-assessment of their own results in the CLIL TU experienced. C4 displays the lowest scores among all five clusters at T1, before the CLIL TU began, and

the highest scores at T3. So, C4 reaches the highest scores and improves the most after the CLIL TU.

These results underline once again the complexity and multifacetedness of the construct of motivation. Motivation toward the subject plays a significant role in determining the effectiveness of CLIL in acquiring the content. In particular, subject appreciation is inversely proportional to the effectiveness of CLIL: the lower the motivation toward the subject, the higher the appreciation of CLIL and the higher the students' achievements. C5, comprising students of the core and the non-core subject with high profile in all four affective variables, gains results similar to C4 at T3, although with a lower improvement rate than C4, as C5 obtains high scores also in T1 and T2. C1 students, who express medium-low appreciation for all the variables increase their performance over time as well. In turn, C2 students, who declare a high appreciation for the subject and EFL, but are not satisfied by CLIL and by the experience of the CLIL TU, are the students who show the lowest performance at T3. C3, composed of students who display high appreciation for the subject, for EFL and CLIL, but who declare low benefit from the CLIL TU, shows a decreasing pattern in performance over time, with a decline in scores from T1 to T2 and, even more so, from T2 to T3. What distinguishes students in C1, C4 and C5 from other clusters is high motivation toward EFL and CLIL. This indicates that high motivation toward the learning situation is crucial for the success of CLIL regarding acquisition of content, as has also been revealed for FL learning in CLIL (Huang 2011, Coonan 2012). Motivation toward the EFL and CLIL learning situation directly correlates with content-specific achievements: students with high appreciation of EFL and CLIL obtain high scores. Motivation toward the subject turns out, instead, to be inversely correlated with students' attainments: students improve disciplinary competence the most in a subject they do not like.

The feedback provided by the students in the questionnaire at T2, right after the CLIL TU, confirms that students of the non-core subject have a better attitude toward CLIL than students of the core subject. Students of the core subject acknowledge the advantages of learning Chemistry in English for their professional careers, but report difficulties in comprehension of concepts and technical lexicon and tend to consider CLIL as an obstacle to learning a subject in which they have a keen interest. Other surveys (Coonan, 2009; Di Martino, 2015) document Italian students' concerns about CLIL, caused by the experienced difficulties to understand disciplinary concepts and technical lexicon, as well as to explain complex issues in a FL, given the low linguistic proficiency they have in FLs (Note 8).

Conversely, students of the non-core subject declare that CLIL makes them more attentive, involves the entire class, introduces a new way of learning, makes lessons more captivating and allows for the study of two subjects at once, Physics and English (Author1, 2018). The high motivation toward the core subject can lead students to fear that CLIL hinders, at least partially, the acquisition of content, because of the difficulties introduced in the use of a FL to study content-specific concepts. In turn, as for the non-core subject, the elements of novelty inserted by CLIL can help arouse learners' motivation and curiosity toward a subject which has no appeal per se in the usual teaching environment.

Both quantitative analysis of students' achievements and affective variables and a qualitative review of students' opinions about the CLIL experience signal that motivation toward the core or the non-core subject involved in CLIL can be relevant to assess the effectiveness of the educational approach. With regards to disciplinary competence, motivation toward the subject is inversely proportional to the effectiveness of CLIL: students who take the CLIL TU in the non-core subject improve their content-specific competence more than their traditionally-educated peers, while students attending the CLIL TU in the core subject do not show any significant difference from their non-CLIL peers.

The better content-specific learning outcomes for CLIL_non-core students compared to CLIL_core students, which is related to a different student perception of CLIL in the two cases, puts into question the advantage of allocating core subjects for CLIL, the preferred practice in Italian high schools. Since CLIL increases students' motivation as well as teachers' motivation (Coyle 2006), there may be more than one professional within the same school who aspires to be trained as a CLIL teacher to deliver the CLIL programme (Note 9). In the case of availability of instructors for either core or non-core subjects who possess the required FL competence and willingness to carry out CLIL lessons, it is important to consider the high potential of CLIL in the non-core subjects in favouring the content-specific competences of learners. As motivation is pivotal for the success of learning, the appointment of CLIL teachers should consider in which subjects, core or non-core, students might be more motivated toward EFL and CLIL, and less anxious about difficulties in content learning.

6. Conclusions

With regard to RQ1 and RQ2, findings show significant effects of CLIL on content-specific learning with a difference between core and non-core subjects. At T1, differences among the two sets of CLIL and non-CLIL classes are statistically non-significant. This confirms that the groups have the same starting level in the content-specific competence in both disciplines and that subsequent differences between the groups are related to CLIL. The analyses reveal that students who improve the most from T1 to T2, and from T2 to T3, are the CLIL students of the non-core subject. They increase scores across the three elicitations for M1, and, to a greater extent, for M2. The scores of the non-CLIL students of the non-core discipline rise across T1, T2 and T3 as well, but to a lesser degree. In turn, CLIL and non-CLIL students of the core subject display a non-significant difference between the three elicitations for both M1 and M2.

The beneficial influence of CLIL upon content-specific learning emerges differently in the two analysed measurements. The CLIL_non-core students develop productive argumentation skills (M2) much more than receptive disciplinary knowledge (M1), although the former is more complex and difficult than the latter. This evidence suggests that CLIL's effects on the acquisition of the non-core subject are considerable. As studies on CLIL have so far not distinguished between these two levels of content-specific competence, this result calls attention to this issue, which is key to better comprehend the impact of CLIL on the acquisition of disciplinary content.

The MFA on CLIL students of core and non-core disciplines examining the interplay between scores in M1 and M2 and individual affective factors (motivation toward the subject, EFL and CLIL, and self-perceived benefits from the CLIL TU) specifically refers to RQ3. Students of the non-core discipline who declare themselves not particularly interested in the subject but appreciate EFL and CLIL derive the most advantage from CLIL, as their scores are the lowest at T1 and the highest at T3. At T3 they even surpass students with the best scores at T1, who are CLIL students of core and non-core subjects with a high appreciation for all listed variables.

The result that students expressing a low interest for the subject derive the most benefit from CLIL suggests that a CLIL approach can foster the acquisition of content in subjects with low appreciation. As demonstrated by student feedback in the questionnaire, CLIL makes lessons of the non-core discipline engaging and challenging, while CLIL applied to the core subject is perceived as an obstacle to an acquiring proper content-specific competence in a discipline crucial to students' careers. A key factor for the high motivation toward CLIL reported by students of the non-core discipline is the novelty of such an educational approach as compared with traditional learning. Since curiosity and motivation can decrease once the learning situation is no longer novel, as documented in studies on motivation toward EFL (Dalton-Puffer & Smit, 2013; Lasagabaster & Doiz, 2015), longitudinal testing to see whether motivation toward CLIL remains different in students of core and non-core subjects over time is needed.

From both analyses and students' feedback, motivation toward the subject, CLIL, and EFL turn out to be entwined: students with lower appreciation for the discipline involved in CLIL, attribute greater value to CLIL and EFL learning. Thus, they are all relevant variables for the acquisition of content, but in opposite directions: motivation toward CLIL and EFL are directly proportional to CLIL effectiveness, whereas motivation toward the subject is inversely correlated to CLIL benefits. More research is needed in CLIL programmes to investigate the effects of these diverse types of motivation on the acquisition of content, drawing on larger samples of students and disciplines, as well as on language-specific learning.

Albeit the small sample of students and disciplines investigated in the current study, the results may also have some implications for educational policy. In Italy, CLIL is generally applied to core disciplines instead of non-core ones, as school principals usually appoint instructors of core subjects to be trained as CLIL teachers and to deliver CLIL in the schools. On the contrary, the study provides the first-ever evidence for the beneficial effects of CLIL mainly in non-core subjects and suggests that CLIL can enhance content-specific learning as well as student motivation best when applied to non-core subjects. Further studies are called for to compare this first evidence with data collected from a larger sample of high-school students and from a larger set of disciplines, both humanities-oriented and science-oriented subjects, in order to test the validity of the presented analyses. Moreover, further research is necessary also to verify whether the benefit for non-core subjects remains constant over time or whether it is due to the novelty of the CLIL at the beginning of its employ within the lessons.

References

Aiello, J., Di Martino, E., & Di Sabato, B. (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. https://doi.org/10.1080/13670050.2015.1041873

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software, 67*(1), 1-48. https://doi.org/10.18637/jss.v067.i01

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

Bécue-Bertaut, M., & Pagès, J. (2008). Multiple Factor Analysis and Clustering of A Mixture Of Quantitative, Categorical and Frequency Data. *Computational Statistics & Data Analysis 52*(6), 3255-3268. https://doi.org/10.1016/j.csda.2007.09.023

Canlas, I. 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), 154-159.

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

Cisotto, L. (1998). *Scrittura e metacognizione. Linee teoriche e proposte operative*. Erickson.

Coonan, C. M. (2007). Insider Views of the CLIL Classroom Through Teacher Self-Observation-Introspection. *International Journal of Bilingual Education and Bilingualism, 10*(5), 625-646. https://doi.org/10.2167/beb463.0

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

Coonan, C. M. (2012). Affect and Motivation in CLIL. In D. Marsh & O. Meyer (Eds.), *Quality Interfaces: Examining Evidence & Exploring Solutions in CLIL* (pp. 53-66). Eichstaett: Eichstaett Academic Press.

Council of Europe. (2018). *Common European Framework of Reference for Languages: learning, teaching, assessment.* Companion volume. Council of Europe Publishing.

Coyle, D. (2006). Content and language integrated learning. Motivating learners and teachers. *Scottish Languages Review, 13*(5), 1-18

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

Csizér, K., & Dörnyei, Z. (2005). The Internal Structure of Language Learning Motivation and Its Relationship with Language Choice and Learning Effort. *The Modern Language Journal, 89*, 19-36. https://doi.org/10.1111/j.0026-7902.2005.00263.x

Dallinger, S., Jonkmann, K., Hollm, I., & Fiege, C. (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. https://doi.org/10.1016/j.learninstruc.2015.09.003

Dalton-Puffer, C. (2011). Content-and-Language Integrated Learning: From Practice to Principles?. *Annual Review of Applied Linguistics 31*, 182-204. https://doi.org/10.1017/S0267190511000092

Dalton-Puffer, C., & Smit, U. (2013). Content and Language Integrated Learning: A research agenda. *Language Teaching, 46*(4), 545-559. https://doi.org/10.1017/S0261444813000256

Dalton-Puffer, C., Hüttner, J., Schindelegger, V., & Smit, U. (2009). Technology-Geeks Speak Out: What Students Think About Vocational CLIL. *International CLIL Research Journal, 1*(2), 18-25.

De Smet, A., Mettewie, L., Galand, B., Hiligsmann, P., & Van Mensel, L. (2018). Classroom anxiety and enjoyment in CLIL and non-CLIL: Does the target language matter?. *Studies in Second Language Learning and Teaching, 8*(1), 47-71. https://doi.org/10.14746/ssllt.2018.8.1.3

Di Martino, E. (2015). Content and Language Integrated Learning: la voce dei più giovani. *Educazione linguistica – Language Education, 4*(1), 85-114. http://doi.org/10.14277/2280-6792/125

Doiz, A., Lasagabaster, D., & Sierra, M. C. (2014). CLIL and motivation: the effect of individual and contextual variables. *The Language Learning Journal, 42*(2), 209-224. https://doi.org/10.1080/09571736.2014.889508

Dörnyei, Z. (2005). *The Psychology of the Language Learner: Individual Differences in Second Language Acquisition.* Lawrence Erlbaum.

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

Fernández-Sanjurjo, J., Fernández-Costales, A., & Arias Blanco, J. M. (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. https://doi.org/10.1080/13670050.2017.1294142

Gardner, R. C., & Lambert, W. E. (1972). *Attitudes and motivation in second language learning.* Newbury House.

Gregorczyk, B. (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. https://doi.org/10.5294/laclil.2012.5.1.2

Haagen-Schützenhöfer, C., & Hopf, M. (2014). Content and Language Integrated Learning in Physics Teaching: Benefits, Risks, Requirements and Empirical Studies. In W. Kaminski & M. Michelini (Eds.), *Teaching and Learning Physics today: Challenges? Benefits?*. Lithostampa.

Huang, K. M. (2011). Motivating Lessons: A Classroom-Oriented Investigation of the Effects of Content-Based Instruction on EFL Young Learners' Motivated Behaviours and Classroom Verbal Interaction. *System, 39*(2), 186-201. https://doi.org/10.1016/j.system.2011.02.002

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

Kassambara, A., & Mundt, F. (2020). Factoextra: Extract and visualize the results of multivariate data analyses. *R Package Version,* 1.0.7.

Lasagabaster, D., & Doiz, A. (2015). A Longitudinal Study on the Impact of CLIL on Affective Factors. *Applied Linguistics, 38*(5), 688-712. https://doi: 10.1093/applin/amv059

Lasagabaster, D., & López Beloqui, R. (2015). The Impact of Type of Approach (CLIL Versus EFL) and Methodology (Book-Based Versus Project Work) on Motivation. *Porta Linguarum, 23*, 41-57. https://doi.org/10.30827/Digibug.53737

Lasagabaster, D., & Sierra, J. M. (2009). Language Attitudes in CLIL and Traditional EFL Classes. *The International CLIL Research Journal, 1*(2), 4-17.

Lasagabaster, D. (2011). English achievement and student motivation in CLIL and EFL settings. *Innovation in Language Learning and Teaching, 5*(1), 3-18. https://doi.org/10.1080/17501229.2010.519030

Lenth, R., Singmann, H., Love, J., Buerkner, P., & Herve, M. (2018). Emmeans: Estimated marginal means, aka least-squares means. *R package version, 1*(1), 3.

Lim Falk, M. (2008). *Svenska i engelskspr ̊akig skolmilj ö. Ämnesrelaterat spr ̊akbruk i tv ̊a gymnasieklasser [Swedish in an English classroom environment. Language use in two grammar school classes]*. Eddy.

Madrid D., & Barrios Espinosa, E. (2018). A Comparison of Students' Educational Achievement across Programmes and School Types with and without CLIL Provision. *Porta Linguarum, 29*, 29-50. https://doi: 10.30827/Digibug.54021

Marsh D., & Langé, G. (1999). *Implementing Content and Language Integrated Learning: A Research-driven*. TIE-CLIL Foundation Course Reader, Jyväskylä University.

Marsh, D., Zajac, M., & Gozdawa-Gołębiowska, H. (2008). *Profile Report Bilingual Education (English) in Poland.* National Centre for Teacher Training and Development.

Masgoret A. M., & Gardner, R. C. (2003). Attitudes, Motivation, and Second Language Learning: A Meta–Analysis of Studies Conducted by Gardner and Associates. *Language Learning, 53*(1), I23-163. https://doi.org/10.1111/1467-9922.00212

Murray, D. R. (2010). Irish-medium language immersion programs' effect on mathematics education. *Journal of Mathematics Education at Teachers College, 1*, 28-32. https://doi.org/10.7916/jmetc.v1i2.688

Nikula, T. (2017). Emerging themes, future research directions. In A. Llinares & T. Morton (Eds.), *Applied Linguistics Perspectives on CLIL*. Benjamins.

Pallotti, G. (2019). Assessing tasks: The case of interactional difficulty. *Applied Linguistics, 40*(1), 176-197.

Paris, S. G., & Turner, J. C. (1994). Situated motivation. In P. R. Pintrich, D. R. Brown, & C. E. Weinstein (Eds.), *Student motivation, cognition, and learning: Essays in honor of Wilbert J. McKeachie* (pp. 213-237). Lawrence Erlbaum Associates, Mahwah.

Piesche, N., Jonkman, K., Fiege, C., Keßler, J. (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*, 108-116.

R Core Team. (2020). R: *A Language and Environment for Statistical Computing. R Foundation for Statistical Computing*. Wien.

Ricci Garotti, F. (2017). Subject Knowledge in CLIL: Which Factors Contribute to Positive Results?. In C. M. Coonan, L. Favaro, & M. Menegale (Eds.), *A Journey through the Content and Language Integrated Learning Landscape* (pp. 55-68). Cambridge Scholars Publishing, Cambridge.

Rumlich, D. (2014). Prospective CLIL and non-CLIL students' interest in English (classes): A quasi-experimental study on German sixth-graders. In R. Breeze, C. Martínez Pasamar, C. Llamas Saíz, &C. Tabernero Sala (Eds.), *Integration of theory and practice in CLIL* (pp. 73-95). Rodopi, Amsterdam.

Ryan, R. M., & Deci, E. L. (2000). Intrinsic and Extrinsic Motivations: Classic Definitions and New Directions. *Contemporary Educational Psychology, 25*, 54-67.

Santipolo, M. (2021). The Content of CLIL and Beyond: Classifying Typologies of Microlanguages for Teaching and Learning Purposes. In A. Graziano, B. Turchetta, F. Benedetti, & L. Cinganotto (Eds.), *Pedagogical and Technological Innovations in (and through) Content and Language Integrated Learning* (pp. 18-28). Cambridge Scholars Publishing, Cambridge.

Seikkula-Leino, J. (2007). CLIL learning: achievement levels and affective factors. *Language and Education, 21*(4), 328-341. https://doi.org/10.2167/le635.0

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

Surmont, J., Struys, E., Van Den Noort, M., & Van De Craen, P. (2016). The effects of CLIL on mathematical content learning: A longitudinal study. *Studies in Second Language Learning and Teaching, 6*(2), 319-337.

Ushioda, E., & Dörnyei, Z. (2009). Motivation, language identities and the L2 self: A theoretical overview. In Z. Dörnyei, & E. Ushioda (Eds.), *Motivation, language identity and the L2 self* (pp. 1-8). Multilingual Matters, Bristol.

Van de Craen, P., Ceuleers, E., & Mondt, K. (2007). Cognitive development and bilingualism in primary schools: Teaching maths in a CLIL environment. In D. Marsh, & D. Wolff (Eds.), *Diverse contexts – converging goals: CLIL in Europe* (pp. 185-200) Peter Lang, Bern.

Várkuti, A. (2010). Linguistic Benefits of the CLIL Approach: Measuring Linguistic Competences. *International CLIL Research Journal, 1*(3), 67-79.

Virdia, S. (2022). The (heterogeneous) effect of CLIL on content-subject and cognitive acquisition in primary education: evidence from a counterfactual analysis in Italy. *International Journal of Bilingual Education and Bilingualism, 25*(5), 1877-1893. https://doi.org/10.1080/13670050.2020.1835805

Vollmer, H. 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.

Williams, K. C., & Williams, C. C. (2011). Five key ingredients for improving student motivation. *Research in Higher Education Journal, 11*, 1-23.

Notes

Note 1. The study is promoted by the Interuniversity Research Centre "LinE – Language in Education" (www.languageineducation.eu). For the requirements of Italian Academy, we declare that Fabiana Rosi is responsible for sections 1, 2, 3.1, 5 and 6; Gioacchino Amato for sections 3.2. and 3.3.; Alessandra Zappoli for sections 3.4. and 4.

Note 2. As documented by Dalton-Puffer et al. (2009), the majority of instructors involved in CLIL are teachers of non-linguistic disciplines without a specific competence in FL teaching.

Note 3. In Italy, the CLIL methodological training courses are aimed at in-service non-language high-school teachers. It is organized by regional scholastic institutions, selecting the teachers to be enrolled based on schools' indications, and provided by a set of university.

Note 4. The initial sample was composed of 89 students, but participants who were not present in all the 3 elicitation sessions were not included in the analysis.

Note 5. The repetition of the same questionnaire at T2 and T3 may produce an echo-effect at T3, but such effect, if any, occurs in all four groups and does not undermine the results.

Note 6. These results are relevant given the p. <0.1 and they could reach significance in case of larger samples.

Note 7. According to Pallotti (2019), a task is difficult when poses high cognitive and linguistic demands on the users.

Note 8. Italian students display among the lowest competence in FL in Europe (Aiello et al. 2017).

Note 9. The CLIL methodological training course is mandatory to be appointed as CLIL teacher in Italy. The Italian official prerequisites for a CLIL teacher are to hold a certification of C1 competence in the FL involved in CLIL and to pass the methodological course. Thus, the school principal's selection of the professional for the training determines which instructors are in charge of CLIL courses in every school.

Copyrights