

Are CLIL Students More Motivated? An Analysis of Affective Factors and their Relation to Language Attainment

MACARENA NAVARRO PABLO

EDUARDO GARCÍA JIMÉNEZ

University of Seville

Received: 16 May 2017 / Accepted: 2 October 2017

ISSN: 1697-7467

ABSTRACT: This paper addresses the purported belief that Content and Language Integrated Learning positively influences students' affective stance. It compares the motivation of CLIL and non-CLIL learners in seven state schools and one charter school in the province of Seville, both at Primary (n=194) and Compulsory Secondary Education (n=158) level. Affective aspects pertaining to motivation and anxiety are grouped around four clusters of factors: (i) desire to work and self-esteem (containing 10 items); (ii) anxiety in the face of exams (with a negative-inhibitory content and made up of 9 elements); (iii) lack of interest in studying (comprising 9 items); and (iv) realistic personal self-demand (consisting of 7 elements). The interaction of motivation and language attainment (considering use of English, vocabulary, listening, speaking, and reading) is also measured in order to confirm or refute prior findings which tend to assign higher levels of motivation to CLIL strands.

Keywords: CLIL, motivation and language attainment.

¿Están los alumnos AICLE más motivados? Análisis de los factores afectivos y los resultados de aprendizaje lingüístico

RESUMEN: Este artículo aborda la creencia de que el Aprendizaje Integrado de Contenidos y Lengua tiene una influencia positiva sobre la situación afectiva del alumnado. En este trabajo se compara la motivación de estudiantes que cursan programas bilingües "AICLE" frente a los que no. La muestra pertenece a siete colegios públicos de la provincia de Sevilla en niveles de Educación Primaria (n=194) y Secundaria Obligatoria (n=158). Los factores afectivos relacionados con la motivación y la ansiedad se han agrupado en cuatro apartados: (i) el deseo de trabajar y la autoestima (con 10 ítems); (ii) la ansiedad frente a los exámenes (con un efecto negativo sobre el aprendizaje de contenido y que cuenta con 9 ítems); (iii) falta de interés por el estudio (con 9 ítems); y (iv) auto-exigencia personal (con 7 ítems). También se ha analizado la interacción entre la motivación y los resultados de aprendizaje lingüístico (considerando uso del inglés, vocabulario, destrezas de comprensión y expresión oral y lectura) con el objetivo de confirmar o refutar resultados de investigaciones previas que parecen asignar niveles altos de motivación asociados al AICLE.

Palabras clave: AICLE, motivación y resultados de aprendizaje lingüístico

1. INTRODUCTION

Content and Language Integrated Learning (CLIL) has been steadily gaining ground across Europe over the last two decades. A significant body of research has explored and defined CLIL characterization, implementation and research in very diverse educational settings (Graaff and Westhoff, 2007; Dalton-Puffer, 2011; Breidbach and Viebrock, 2012; Pérez-Cañado, 2012; Wegner, 2012; Dalton-Puffer and Smit, 2013a, 2013b; Coyle, 2013; Kerstin, 2013; Pérez-Cañado, 2014). Many studies have focused on the positive effects of CLIL and the cognitive benefits of bilingualism (Casal & Moore, 2009; Lorenzo, Casal & Moore, 2009; Nieto, 2016).

Nevertheless, some authors have also indicated a number of unsolved problems concerning the implementation of this methodology (Pavón & Rubio, 2010; Coyle, 2013; Pavón, 2013; Pérez Cañado, 2016). For example, it has been noted that one of its main downsides in Europe is the lack of a systematic and standardized approach to its implementation throughout the continent, where teaching experiences and learning environments differ not only among countries but also among regions within the same country and even among schools in a single town due to different social, cultural, economic and political contexts (Arribas, 2016; Alejo and Piquer, 2016). Arribas (2016: 271) explains these differences by pointing to a series of factors: “the lack of regulation or official guidelines, and more particularly a) teachers’ linguistic competence, b) teachers’ training, c) language level to be achieved by learners, and d) distribution of CLIL hours”. This author provides examples from France, Poland and Hungary where “CLIL students are selected based on their performance on entrance exams”, whereas in Spain, Germany, Finland and Sweden “CLIL is normally open to every single student”.

Against this backdrop, we deem it necessary to analyse the effectiveness of CLIL methodology in a specific setting with robust empirical data drawn from seven state schools and one charter school in the province of Seville, both at Primary and Compulsory Second Education level. The purpose of this study is twofold. Firstly, to examine the impact of the CLIL methodological approach on pupils’ language attainment in use of English, vocabulary, listening, speaking and reading. Secondly, to consider the interaction of affective aspects and achievement levels. With regard to these affective aspects, four factors are regarded as important: self-esteem, self-demand, anxiety and lack of interest.

2. CONCEPTUALIZING MOTIVATION IN L2 LEARNING

There is no consensus regarding the definition and nature of motivation. In fact, Dörnyei (2001a: 7) states that “researchers disagree strongly on virtually everything concerning the concept” and calls it an “umbrella term,” “an abstract, hypothetical concept,” “a multi-faceted construct” (Dörnyei, 2001a: 7; 2001b: 1; 2009: 117). Similarly, Coyle (2013) concurs that a considerable deadlock has been reached regarding the nature of motivation. Also, Lasagabaster (2011:10) refers to motivation as having an “abstract and multidimensional character.”

The reason why motivation is conceptualized in such abstract terms is probably due to the fact that there are different kinds of motivation that are likely to affect students’ commitment and outcomes in second language learning differently. Finally, Dörnyei (2001a: 8) points to the *direction* and *magnitude* of human behaviour as central concerns regarding

motivation. These refer to “the choice of a particular action,” and “the effort expended on it.” In other words, motivation is responsible for: “why people decide to do something,” “how long they are willing to sustain the activity” and “how hard they are going to pursue it.” In this way, Dörnyei (2009: 118) describes motivation, not as a function of stimuli and reinforcement, but as a process focused on the individual’s thoughts and beliefs that are transformed into action.

The concept of motivation has developed from different theories. Dörnyei (2001a) provides a detailed overview of theories in motivational and social psychology. The following table (Table 1) summarises some of these theories:

Table 1. Leading motivation theories in Psychology

Expectancy value framework: Achievement motivation theory (Atkinson, 1974)		
Expectancies of success	Attribution theory (Weiner, 1992)	
	Self-efficacy theory (Bandura, 1993)	
	Self-worth theory (Covington, 1992)	
Incentives values	Attainment value	
	Intrinsic values	
	Extrinsic utility values	
	Cost	
Need for achievement		
Fear of failure		
Goal theories (Maslow, 1970): focused on basic human needs		
Five classes of needs: Physiological Safety Love Esteem Self-actualization	The concept need has been replaced by the more specific construct of goal	Goal-setting theory
		Goal-orientation theory
Intrinsic vs. extrinsic motivation (Deci and Ryan, 1985) Self-determination theory		
Intrinsic motivation	<i>To learn</i>	Engaging in an activity for the pleasure and satisfaction of understanding something new, <u>satisfying one’s curiosity and exploring the world.</u>
	<i>Towards achievement</i>	Engaging in an activity for the satisfaction of surpassing oneself, coping with challenges and <u>accomplishing or creating something.</u>
	<i>To experience stimulation</i>	Engaging in an activity to experience pleasant sensations.
Extrinsic motivation	<i>External regulation</i>	Refers to the least self-determined form of extrinsic motivation, coming entirely from external sources such as rewards or threats.
	<i>Introjected regulation</i>	Involves externally imposed rules that the student accepts as norms to be followed in order not to feel <u>guilty.</u>
	<i>Identified regulation</i>	Occurs when the person engages in an activity because he or she highly values and identifies with the behaviour, and sees its usefulness.
	<i>Integrated regulation</i>	Involves choice behaviour that is fully assimilated with the individual’s other values, needs and identity.
Amotivation	Lack of any regulation	

Two theories can clearly be discerned as the most influential ones in the theoretical studies on motivation and L2 learning. The first one is Gardner's motivation theory. Within this theory, the socio-psychological model, developed by Gardner and Lambert (1972:1) helped to answer a rather simple question: "How is it that some people can learn a foreign language quickly and expertly while others, given the same opportunities to learn, are utter failures?" Gardner's model was based on Goal Theories and it was included within the so-called "goal-orientation theory." It focused on an integrative and instrumental motivation paradigm. This model, as MacIntyre, Mackinnon and Clément (2009) point out, is linked to a "high-quality measurement tool," the Attitude and Motivation Test Battery (AMTB Gardner, 1985; Gardner & MacIntyre, 1991). This instrument, designed to assess motivational aspects affecting English-speaking students' learning of French as a second language in elementary and secondary schools in Canada (Gardner, 1985), is described by Dörnyei (2009:52) as a "multicomponent motivation test" made up of over 130 items with a list of 13 constituent scales: attitudes towards French Canadians, interest in foreign languages, attitudes towards European French people, attitudes towards learning French, integrative orientation, instrumental orientation, French class anxiety, parental encouragement, motivational intensity, desire to learn French, orientation index, evaluation of French teacher, and evaluation of French courses.

In the 1990s, as Dörnyei (2001: 104) claims, there was a significant "educational shift in L2 motivation". Questions were raised about the gap between general motivational theories and L2 motivational ones. For this reason, over the next decade, there was a considerable increase in studies on L2 motivation and the influence of affective factors on students' achievement. In this context, Gardner revised the model in the light of recent cognitive motivational theories (Tremblay and Gardner's revised model, 1995) so that new elements from expectancy-value and goal theories were incorporated. In this model, they included goal salience, valence and self-efficacy.

The second influential theory to analyse motivation was Dörnyei's L2 Motivational Self System. This model drew on previous motivational theories, especially Gardner's, Clément's, and Crookes and Schmidt's theories. As described by Dörnyei (2001), L2 motivation is conceptualized within a framework of three distinct levels represented in Table 2.

Table 2. Framework of L2 motivation

Dörnyei's framework of L2 motivation (Dörnyei, 2001:113)	
LANGUAGE LEVEL	Integrative motivational subsystem Instrumental motivational subsystem
LEARNER LEVEL	Need for achievement. Self-confidence Language use anxiety Perceived L2 competence Causal attribution Self-efficacy
LEARNING SITUATION LEVEL Course-specific motivational components	Interest (in the course) Relevance (of the course to one's needs) Expectancy (of success) Satisfaction (one has in the outcome)

Table 2. Framework of L2 motivation (Continuation)

Dörnyei's framework of L2 motivation (Dörnyei, 2001:113)	
Teacher-specific motivational components	Affiliative motive (to please the teacher) Authoritative motive (controlling vs autonomy-supporting) Direct socialization of motivation <ul style="list-style-type: none"> - Modelling - Task presentation - Feedback
Group-specific motivational components	Goal-orientedness Norm and reward system Group cohesiveness Classroom goal structure (cooperative, competitive or individualistic)

Although Gardner's and Dörnyei's models are frequently referred to in studies on motivation (Alejo and Piquer, 2016; Coyle, 2013; Doiz *et al.* 2014; Kormos and Csizer, 2011; Lasagabaster, 2015; Pinner, 2013), in recent years there has been a considerable proliferation of theories related to motivation (De la Fuente, 2004). A number of different models derived from a variety of theoretical approaches coexist nowadays and different authors have created their own theoretical constructs and introduced new terminology related to motivation. The introduction of the concept of "selves" into the field of language learning motivation has the potential to open new insights into language learning and motivation research. An important line of research in social psychology distinguishes between "ideal" and "possible" selves. On the one hand, the ideal self represents "the attributes that a person would like to possess (e.g., hopes, aspirations, desires)". Possible selves, on the other hand, "represent individuals' ideas of what they might become, what they would like to become, and what they are afraid of becoming, and thus provide a conceptual link between cognition and motivation" (Dörnyei and Csizer, 2002: 453-454).

However, as Dörnyei claims, the numerous emerging theories, which incorporate new labels and nuances, make it difficult to reach a consensus regarding motivation. Therefore, the literature presents "the partly overlapping notions of self-confidence, self-efficacy, self-competence, task specific self-concept, self-worth, and academic self-esteem" (Dörnyei 2001: 213). Thus, in order to clarify the concepts that are central to the present study, Table 3 synthesizes some of the theoretical constructs used in relation to motivation and L2 learning that emerged from the revision of the relevant literature (Daniel *et al.*, 2008; De la Fuente, 2004; Hidi, 2006; Kormos and Csizer, 2011; Pinner, 2013; Pintrich, 1999; Seikkula-Leino, 2007).

Table 3. *Self-related beliefs*

Constructs	Definition
Self-efficacy (Bandura, 1986)	Self-efficacy beliefs express one's views as to whether one is capable of performing a given learning task and are consequently future-oriented.
Self-concept (Shavelson <i>et al.</i> , 1976 & Burns, 1979, 1982)	Self-concept beliefs are based on past experiences and are broader evaluations of one's general self-worth or esteem.
Self-esteem (Borba, 1989, 1994)	Self-esteem consists of five basic building blocks: security, selfhood, affiliation, mission and competence.
Self-confidence (Clément, 1980)	Self-confidence in general refers to the belief that a person has the ability to produce results, accomplish goals or perform tasks competently.

As stated above, this study analyses the effects of motivation on students' language attainment. In order to do so, affective aspects pertaining to motivation and anxiety are grouped around four clusters of factors: self-esteem, self-demand, anxiety and lack of interest. The former two can be considered as potential positive factors for student motivation. In the present study, self-esteem is related to Borba's conceptualization of self-esteem and Shavelson's definition of self-concept (see Table 3), as both refer to the person's ideas and beliefs about themselves and their own competences. Self-demand is, on the other hand and according to the literature on motivation, related to "volition," which is defined by Corno (1993:16) as a "dynamic system of psychological control processes that protect concentration and directed effort in the face of personal and/or environmental distractions, and so aid learning and performance". Furthermore, Corno also adds volition as a parallel construct within the field of language learning "motivated behaviour", Volition consists of effort and persistence, both factors mentioned in the relevant literature (Dörnyei, 2009; Gardner, 1972).

The other two motivational aspects considered in this study are anxiety and lack of interest. Both are factors that may have a negative influence on the L2 learning process. According to Lasagabaster (2015: 5), affect is made up of intrinsic motivation, liking for challenge and anxiety, which he defines as "the lack of confidence in oneself as a learner, uneasiness, frustration, self-doubt, apprehension and tension which are specifically related to language learning situations". Finally, interest is also an important component in motivation, considered to be "a unique motivational variable, as well as a psychological state (...) characterized by increased attention, concentration and affect" (Hidi, 2006: 70).

The literature on foreign language learning has often indicated that motivation has a definitive influence on learners' language attainment (Doiz, Lagasabaster & Sierra, 2014; Dörnyei, 2009; Gardner, 2015; Lasagabaster, 2011; Pintrich, 1999). The present study focuses on the four motivational factors described above in order to assess the extent of such influence in CLIL settings.

3. RESEARCH ON MOTIVATION

This section offers an overview of the main studies on motivation in CLIL settings. For example, after examining the relationship between achievement level and affective factors, Seikkula-Leino (2007) concluded that, in spite of their low self-concept in foreign languages, CLIL learners also have a strong motivation to learn. In the Spanish context, Lasagabaster has conducted a series of research studies on the relationship between English language achievement and student motivation in CLIL and English as Foreign Language (EFL) settings. These studies, which involve secondary school students in the Basque Country, compare the attitude of students toward CLIL programs with the attitude of students enrolled in regular EFL programs. In 2009, Lagasabaster and Sierra (2009) carried out a study with 287 students. The tool used was a seven-point semantic differential questionnaire based on Gardner (1985). The sociocultural variable and the independent variable of gender were scrutinized in this study. Their results showed that “CLIL programmes help to foster positive attitudes towards language learning in general” and that “[s]tudents enrolled in CLIL groups held more positive attitudes towards English”. In 2011, Lasagabaster published a paper in which he analysed the data from a cross-sectional study involving 191 secondary school students and found that both EFL and CLIL students were highly motivated to learn English, although CLIL students were significantly more enthusiastic than those in traditional EFL classrooms. Thus, he concluded: “there is a strong relationship between the CLIL approach and motivation” (Lasagabaster, 2011: 14). In this study, three motivational factors were analysed: 1) interest and instrumental orientation; 2) attitudes towards learning English in class; and 3) effort. In 2014, Doiz, Lagasabaster and Sierra provided an in-depth analysis of a three-year longitudinal study. They focused on the influence of the CLIL methodological approach and its interaction with individual and contextual variables. They considered the individual variables of students’ age and sex and the contextual variable concerning parental socio-cultural level. The questionnaire, based on scales used by Gardner (1985) and Schmidt and Watanabe (2001), excluded integrative motivation, as it was not considered relevant to the context, and included anxiety, as it was hypothesized to play an important role in L2 motivation. Thus, six scales were used, namely, instrumental orientation, parental support (these two scales taken from Gardner, 1985), intrinsic motivation, interest in foreign languages/cultures, anxiety and motivational strength (the last four scales taken from Schmidt and Watanabe, 2001). On this occasion, the findings revealed that CLIL students had the lowest means in all the scales measured except for anxiety. The differences between both groups regarding anxiety were not statistically significant; in other words, both groups showed a similar degree of anxiety when they had to speak English. The other scales showed similar results to those of previous studies, with CLIL students being intrinsically more motivated, more instrumentally oriented and showing a higher interest in foreign languages than non-CLIL students.

One year later, in 2015, Doiz and Lasagabaster published the results of another study on the impact of CLIL on affective factors in order to provide more empirical data to confirm the purported benefits and the positive influence of these programs on students’ affective stance. Again, the questionnaire used was based on the instruments created by Gardner (1985) and Schmidt and Watanabe (2001), although Doiz and Lasagabaster reduced the number of variables in the questionnaire and modified the instrument, reorganizing the division of items into factors to carry out an exploratory analysis. Thus, the analyses allowed them to have

lower cut-off values for factor loadings. They administered the questionnaire in two age groups and in three phases of the survey, considering the following seven factors: intrinsic motivation, instrumental motivation, global integration, interest in foreign languages and cultures, anxiety, motivational strength and parental encouragement. The data were compared across time. The findings revealed that motivation was maintained over time in non-CLIL classes, whereas there was a motivational decline in some of the affective dimensions of the younger CLIL students. In the case of younger students, significant differences were found only on the anxiety scale. Anxiety seems to show higher levels over time. In the case of the older CLIL students, they remained more motivated in all the scales except for interest and anxiety.

Also in 2015, Otwinoska and Forys (2015) studied the links between affectivity and cognition in upper-primary CLIL classes in order to verify whether negative emotions inhibit cognitive processes. They did so by obtaining data from children's performance in Science, Mathematics and English, and found that young CLIL learners experience symptoms of intellectual helplessness (IH) in Science and Mathematics CLIL classes, being negative affectivity a predictor of IH. However, these results were different for English, as they demonstrated that grades in English did not significantly predict IH in CLIL. The explanation for this may be found in the differences in the type of language used in the modules observed and the distinction between basic interpersonal communication skills (BICS) and cognitive academic language proficiency (CALP) (Cummins, 1979: 12-13).

More recently, Arribas (2016) has examined a whole CLIL school analysing students' attitudes, motivation and receptive vocabulary outcomes. The results showed that CLIL learners scored higher in receptive vocabulary tests due to their greater motivation. However, as in other studies, differences between CLIL and non-CLIL groups were not statistically significant. The reason for these findings was found in the irregular implementation of the CLIL program in the school and its lack of experience with the methodology.

The role played by external factors, such as input-related variables (extramural exposition, amount of input, intensity of exposition), age of onset, age across time, gender, social setting, type of school (state or private) or socioeconomic status, in motivation has been challenged by a number of researchers (Fernández and Canga, 2014; Larson-Hall, 2008; Lasagabaster, 2011; Unsworth *et al.* 2014; Pfenninger, 2016). Alejo and Piquer (2016) analysed input-related variables, motivation and language attainment in two different social settings (rural and urban). They focused on the socio-economic status as a proxy. They also studied motivational profiles associated with social classes. They administered a motivation questionnaire that incorporated most of the theoretical constructs on motivation elaborated by Dörnyei (2001, 2006) and included questions on (1) attitudes towards learning experiences, (2) integrativeness, (3) ideal self, (4) instrumentality, (5) anxiety and (6) effort. The test performed comparing motivational profiles did not show significant differences between the two schools under study. Nevertheless, when considering Dörnyei's motivational system (2005), they found significant differences between certain motivational dimensions such as anxiety and effort.

Both motivation and language attainment are recognized as relevant variables in foreign language learning. As has been shown, it is possible to identify a positive link between affective factors and scores in language attainment. However, many questions remain unanswered due to the fact that motivational variables were not controlled in those studies. As Lasagabaster

(2011: 7) states: “Although motivation was mentioned as one of the key factors in these results, this variable was not controlled”. To this end, the present study aims to analyse two variables in CLIL implementation, examining both the impact of the CLIL educational approach and the influence of affective factors on pupils’ language achievement. The ultimate aims are (1) to contribute empirically sound data to promote further implementation of CLIL, and (2) to encourage reflection on how CLIL is implemented in a particular context, its effect on students’ linguistic competence and the interrelation between this methodology and students’ affective stance. In order to do so, this study controls motivational variables with an initial homogenization of the sample. More information about this process is provided in the description of the study.

4. THE STUDY

4.1. Research questions

In this paper, CLIL and non-CLIL learners’ language attainment is examined, compared and contrasted. Furthermore, the interaction of motivation and language attainment is measured considering pupils’ achievement levels regarding use of English, vocabulary, listening, speaking and reading. Affective aspects pertaining to motivation are grouped around four clusters: self-esteem, self-demand, anxiety and lack of interest. Consequently, the study correlates primary and secondary students’ language attainment outcomes with their motivation. The following research questions form the basis of this research:

1. Are there statistically significant differences between the achievement levels of CLIL and non-CLIL learners?
2. Are CLIL students more motivated to learn English than non-CLIL students?
3. What is the impact of motivational variables on CLIL and non-CLIL students’ language attainment?

4.2. Sample

The CLIL model under scrutiny in this study was implemented in a monolingual context. Out of the initial 428 participants, 352 were considered for the study. A test was performed in order to control the effect of students’ verbal intelligence and to avoid the possible influence of this variable on the results. Thus, it was possible to determine whether CLIL was truly responsible for the potential differences observed or whether they could be ascribed to other variables. To this end, CLIL and non-CLIL students were matched in terms of verbal intelligence and motivation discarding from the original sample students with higher and lower scores. The final sample comprised 352 students divided into non-CLIL learners (61.4%) and CLIL learners (38.6%). Thus, this study guarantees the homogeneity of the sample and, accordingly, the reliability of the results.

The students were enrolled in seven state schools and one charter school in the province of Seville, in both Primary (12-year-old students) and Compulsory Secondary Education (14-year-old students). As regards the setting, and following the criteria of the regional government (see references: Junta de Andalucía), five urban and three rural schools took part in

the study. The number of participants was approximately equal across genders (50.4% boys and 49.6% girls). Participants were streamed into two different instruction types: students who were enrolled in CLIL programs and students who followed an EFL approach (from now on non-CLIL) (see Table 4).

Table 4. Description of the sample

Centres	Ownership	Area	Level	Program
1 (n=42)	Charter	Urban	Primary	Non-CLIL= 42
2 (n=55)	State	Rural	Primary	non-CLIL= 31
				CLIL= 24
3 (n=46)	State	Urban	Primary	non-CLIL= 28
				CLIL= 18
4 (n=51)	State	Rural	Primary	non-CLIL= 32
				CLIL= 19
5 (n=56)	State	Urban	Secondary	non-CLIL= 26
				CLIL= 30
6 (n=28)	State	Urban	Secondary	non-CLIL= 10
				CLIL= 18
7 (n=23)	State	Rural	Secondary	non-CLIL= 13
				CLIL= 10
8 (n=51)	State	Urban	Secondary	non-CLIL= 34
				CLIL= 17

For the speaking test (assessed with several scales, namely: speaking, grammar, lexical range, fluency, pronunciation and task fulfilment), students were arranged in pairs. This made it possible to assess peer interaction and to reduce test time. The primary school test was made up of two parts, whereas the secondary school test was composed of three parts. In both tests, the first part focused on a series of questions about the students' personal lives. In the second part of the primary school test, students were provided with an image and some questions related to it. They had to hold a conversation based on the image and the questions. In the second part of the secondary school test (spoken interaction: two-way dialogue), students were presented with three different situations (e.g. you are going on a trip to New York; you are going to organize a surprise birthday party; you are going to do a school project on animals). They had to agree on which situation to choose, and then answer a few questions related to the situation. Finally, the last part of the test was a three-way dialogue in which a topic was provided and students in pairs had to hold a conversation engaging

the examiner too. The primary school test was 15 minutes long, whereas the secondary school test was between 20 and 30 minutes long. Due to the nature of these tests, the time needed to administer them and the requirements to be met (students had to be arranged in mixed-gender pairs in order to facilitate voice recognition during data analysis), it was necessary to select the students that would take the test. A minimum of two CLIL and two non-CLIL pairs were selected in each school, making up a total of 38 students (19 CLIL and 19 non-CLIL learners). The selection of students was made in collaboration with their teachers on the basis of a degree of homogeneity in their linguistic competence.

4.3 Instruments

For this study, students' scores in two tests were analysed: a language proficiency test and a motivational test. The first one was originally designed and validated to guarantee content and construct validity. Three different batteries of six subtests each (grammar, vocabulary, reading, writing, listening, and speaking) were designed, following the Common European Framework of reference (CEFR), the national Decrees, and the regional Orders that establish the official curriculum for the educational stages assessed. They corresponded to each of the three levels on which the overall study was centered (6th grade of Primary Education, 4th grade of Compulsory Secondary Education, and 1st grade of Baccalaureate).

The motivational test used was the MA test by Pelechano (1994). The test is made up of 35 items and focuses on four motivational aspects related to achievement and anxiety: (i) desire to work and self-esteem (containing 10 items); (ii) realistic personal self-demand (consisting of 7 elements); (iii) anxiety in the face of exams (with a negative-inhibitory content and made up of 9 elements); (iv) lack of interest in learning (comprising 9 items).

4.4. Data analysis

The results were described using the mean score and the standard deviation. Significant levels between CLIL and non-CLIL learners regarding their L2 language achievements were set using a Student-Fisher t-test for independent groups. Analyses of covariance were also carried out in order to determine the effect of the motivational variables on the students' English language achievement (dependent variable). The aim was to analyse the interaction of motivation and language attainment and to establish the degree of interaction of motivational variables (covariables) and CLIL programs (independent variable). To this end, a contrast regarding the equality of variances was performed in CLIL and non-CLIL groups using the Levene test.

No statistically significant differences were found between CLIL and non-CLIL groups regarding their variances in self-esteem, anxiety, lack of interest or self-demand scales, in both primary and secondary students. Similarly, correlation analysis determined the possible multicollinearity between motivational variables and English language competence results. There were negative statistically significant correlations between the lack of interest scale and the English language results and positive correlations between self-demand and the English language results. These correlations were higher in primary school students; however, all of them were less than 0.70 (the highest was 0.48), which means that there was no multicollinearity.

5. RESULTS

5.1. The achievement levels of CLIL and non-CLIL learners

The level of attainment of CLIL learners and non-CLIL learners was different for both primary and secondary education. The results showed that there were statistically significant differences as regards use of English, vocabulary and all subtests for speaking in primary school students; a significance level of 0.05 was assumed (see Table 5) that is lower than the p value of those tests. Interestingly, no statistically differences were found in any of the receptive skills.

Table 5. Mean and standard deviation of test scores in Primary School students

Tests	Program	Mean	Standard deviation	p value
Use of English	Non-CLIL learners	8.92	5.44	0.002
	CLIL learners	11.49	5.08	
Vocabulary	Non-CLIL learners	5.77	3.13	0.001
	CLIL learners	8.05	3.07	
Listening	Non-CLIL learners	11.87	2.50	0.900
	CLIL learners	11.82	3.06	
Reading	Non-CLIL learners	5.20	3.55	0.194
	CLIL learners	5.93	3.90	
Speaking Total	Non-CLIL learners	5.23	2.04	0.001
	CLIL learners	8.12	1.79	
Grammar	Non-CLIL learners	0.98	0.48	0.001
	CLIL learners	1.52	0.45	
Lexical Range	Non-CLIL learners	1.00	0.44	0.001
	CLIL learners	1.64	0.35	
Fluency Interaction	Non-CLIL learners	1.04	0.51	0.001
	CLIL learners	1.66	0.48	
Pronunciation	Non-CLIL learners	1.19	0.38	0.001
	CLIL learners	1.66	0.36	
Task Fulfilment	Non-CLIL learners	1.02	0.40	0.001
	CLIL learners	1.66	0.36	

In secondary education, we also found statistically significant differences as regards use of English and vocabulary, listening, reading and all subtests of speaking; a significance level of 0.05 was assumed (see Table 6) between the level of attainment of CLIL learners and non-CLIL learners. The *p* value in those tests is lower than the significance level.

Table 6. Mean and standard deviation of tests scores in Secondary School students

Tests	Group	Mean	Standard deviation	p value
Use of English	Non-CLIL learners	17.16	10.77	0.001
	CLIL learners	28.85	10.64	
Vocabulary	Non-CLIL learners	6.64	3.53	0.001
	CLIL learners	9.92	3.05	
Listening	Non-CLIL learners	3.55	1.51	0.001
	CLIL learners	4.71	1.41	
Reading	Non-CLIL learners	2.64	1.63	0.001
	CLIL learners	3.91	1.55	
Speaking Total	Non-CLIL learners	5.00	2.05	0.001
	CLIL learners	8.32	1.60	
Grammatical	Non-CLIL learners	0.92	0.48	0.001
	CLIL learners	1.58	0.38	
Lexical Range	Non-CLIL learners	0.974	0.46	0.001
	CLIL learners	1.71	0.35	
Fluency Interaction	Non-CLIL learners	0.97	0.49	0.001
	CLIL learners	1.71	0.42	
Pronunciation	Non-CLIL learners	1.16	0.37	0.001
	CLIL learners	1.61	0.32	
Task_Fulfilment	Non-CLIL learners	0.97	0.42	0.001
	CLIL learners	1.71	0.35	

In the case of secondary school students, there are no differences between the results in receptive and productive skills. CLIL groups outperformed non-CLIL students on all the tests.

5.2. The impact of students' motivation on CLIL and non-CLIL programs

Once it was established that the level of attainment of CLIL and non-CLIL students differed, it was appropriate to research if motivational factors played a major role in this, or whether such differences might have been due to other aspects of the CLIL program.

Table 7 shows the effects of the independent variable (CLIL program) and the corresponding effects of the covariance on the dependent variable (English language achievement), the latter measured by four subtests (use of English, vocabulary, listening and reading) at primary level. No statistically significant effect of self-esteem was found regarding any of the subtests. The p -value (Sign.=0.384) was higher than the significance level 0.05. The anxiety variable only had a positive statistically significant effect in use of English (p -value=0.003), with a small effect size but close to 5% (Partial $\eta^2=0.047$); in other words, anxiety was associated with the *use of English* results.

Lack of interest was the motivational variable with the highest effect on the students' performance. In fact, it had a statistically significant effect on the results obtained in all subtests since all p -values were lower than the significant level 0.05. The effect size was low in all the subtests, the highest corresponding to the vocabulary test, which was about 6% (Partial $\eta^2=0.055$).

Self-demand had a statistically significant effect on both use of English and *vocabulary*, the p -value being lower than the significant level 0.05. The effect size of this covariant in each of the tests is lower than 3%.

Table 7. Tests of between-subject effects (Primary School students)

	Covariates								Independent	
	Self-esteem		Anxiety		Lack of interest		Self-demand		Program	
Dependents	Sign.	η^2	Sign.	η^2	Sign.	η^2	Sign.	η^2	Sign.	η^2
Use of English	0.384	0.004	0.003	0.047	0.023	0.027	0.019	0.029	0.010	0.035
Vocabulary	0.554	0.002	0.396	0.004	0.001	0.055	0.022	0.028	0.001	0.086
Listening	0.760	0.001	0.357	0.005	0.004	0.043	0.529	0.002	0.541	0.002
Reading	0.364	0.004	0.350	0.005	0.007	0.039	0.087	0.015	0.429	0.003

The effect of the independent variable (program) on use of English showed statistically significant differences between CLIL and non-CLIL learners (p -value=0.010). The mean in the CLIL group was higher (11.106) than in the non-CLIL group (9.094), as shown in Table 8. However, the effect size was low since only 3.5 % of the variance in use of English was associated with the program (Partial $\eta^2 = 0.035$). The total variance was also low, $R^2 = 0.185$ being adjusted. However, when the effect of the covariates was discounted, the difference between the two groups diminished (compare means in Table 5 to means in Table 8).

The effect of the independent variable also showed statistically significant differences between CLIL and non-CLIL learners in vocabulary (p -value=0.001). The mean was higher in the CLIL groups (7.283) than in the non-CLIL groups (5.878), as can be seen in Table 8. The effect size was higher than in use of English; this means that nearly 9% of the variance in vocabulary was associated with the program (Partial $\eta^2 = 0.086$). The total variance was also low, $R^2 = 0.189$ being adjusted. Again, when the effect of the covariates was discounted, the difference between the two groups is diminished (compare means in Table 5 to means in Table 8).

No statistically significant differences were found regarding listening or reading between CLIL and non-CLIL learners, being the p -value associated with the mean of the effect of the program higher than the significant level 0.05.

The effect of the covariate lack of interest was low since, when it was removed, it was not possible to reject the null hypothesis of statistically significant differences between CLIL and non-CLIL learners (listening: p -value=0.541; reading: p -value=0.429).

Table 8. Estimated marginal means (Primary School students)

Dependents	Program	Mean	Standard deviation
Use of English	non-CLIL	9.09	0.42
	CLIL	11.10	0.63
Vocabulary	non-CLIL	5.87	0.25
	CLIL	7.28	0.38
Listening	non-CLIL	11.93	0.22
	CLIL	11.68	0.34
Reading	non-CLIL	5.28	0.31
	CLIL	5.73	0.46

Table 9 shows the effects of the independent variable and the covariates on the dependent variable in secondary students. All the motivational variables, except for self-demand, had a statistically significant effect on each of the four subtests used to measure the dependent variable (L2 language achievement). Thus, there were statistically significant differences regarding use of English, due to the effect of the covariates anxiety and lack of interest; in vocabulary, due to the effect of the covariant anxiety; and in listening, due to the effect of the covariant self-esteem. In all the covariates the p -value was lower than the significant level 0.05. The effect size of the covariates in the results of the tests was essentially irrelevant, the highest being 4.7% (Partial $\eta^2=0.047$) and it corresponded to the effect of the covariant anxiety on use of English.

The independent variable had a statistically significant effect on all the subtests. The effect sizes were relevant exceeding 22% in use of English.

Table 9. Tests of between-subject effects (Secondary School students)

	Covariates								Independent	
	Self-esteem		Anxiety		Lack of interest		Self-demand		Program	
Dependents	Sign.	Eta ²	Sign.	Eta ²	Sign.	Eta ²	Sign.	Eta ²	Sign.	Eta ²
Use of English	0.502	0.003	0.007	0.047	0.040	0.028	0.085	0.019	0.001	0.223
Vocabulary	0.905	0.001	0.030	0.031	0.551	0.002	0.510	0.003	0.001	0.193
Listening	0.042	0.027	0.619	0.002	0.297	0.007	0.675	0.001	0.001	0.128
Reading	0.135	0.015	0.850	0.001	0.337	0.006	0.571	0.002	0.001	0.125

Table 10 shows the estimated means for CLIL and non-CLIL learners, once the effect of the motivational variables was discounted. Means are in favor of CLIL learners in all the subtests. The total variance explained by the independent variable program regarding use of English was close to 30 % (Adjusted $R^2 = 0.293$); in vocabulary, close to 20 % (Adjusted $R^2 = 0.203$); in listening, close to 14% (Adjusted $R^2 = 0.141$); and in reading, close to 13 % (Adjusted $R^2 = 0.130$).

Despite the modest effect that motivational variables have on L2 language achievement, the effect can be observed in the different mean scores corresponding to the tests taken by CLIL and non-CLIL learners. Thus, comparing the means in Table 10 to the means in Table 6, the differences in means between CLIL and non-CLIL learners are slightly lower in Table 10. The reason for this decrease is that in Table 10 the effect of motivational variables was discounted.

Table 10. Estimated marginal means (Secondary School students)

Dependents	Program	Mean	Standard deviation
Use of English	non-CLIL	17.52	1.13
	CLIL	28.44	1.19
Vocabulary	non-CLIL	6.67	0.36
	CLIL	9.88	0.38
Listening	non-CLIL	3.57	0.16
	CLIL	4.68	0.16
Reading	non-CLIL	2.67	0.17
	CLIL	3.87	0.18

6. CONCLUSIONS

To answer the research question regarding the existence of statistically significant differences between the achievement levels of CLIL and non-CLIL learners, the findings present slightly different results depending on the educational level analysed. In primary education, the results show that CLIL learners outperformed non-CLIL learners on all the subtests. However, the differences were not statistically significant for all of them. CLIL seems to have a lower effect on receptive skills (listening and reading) than on productive skills (speaking and writing). We agree with Pfenninger (2016: 137) that the reason for the higher effect of CLIL on productive skills may be found in the “oral-based, communicative, pedagogical approach used in CLIL programmes”.

In secondary education, there are statistically significant differences for all the subtests. These differences are always in favor of the CLIL learners. Their means are always higher than those in non-CLIL programs. The findings also suggest that significant differences in favor of the CLIL learners increase at secondary level.

The second and third research questions may be addressed jointly. Our findings suggest a positive answer to the second question on whether CLIL students are more motivated to learn English than non-CLIL students. In this regard, this study partly agrees with previous research findings (Seikkula-Leino, 2007; Lagasabaster, 2009, 2011 and 2014; Pfenninger, 2016) that acknowledge the interaction between motivation and language achievement. For our sample, the confluence of motivational variables seems to play a role in language achievement, since the motivational variables have a statistically significant effect on the differences observed in the subtests carried out. Therefore, motivation is, in general terms, an unequivocally important factor for the learning of a second language, and it plays a more important role in CLIL than in non-CLIL settings. However, a more nuanced approach is needed when considering the impact of motivation variables on CLIL and non-CLIL students' language attainment. This is due to the fact that, when motivational factors are disaggregated, as was done for this study, the results do not seem to indicate that all affective factors have the same influence on learners' L2 attainment. At least, this was the case in the bilingual schools where this study took place. For example, lack of interest is the variable with the largest effect on the results of the subtests, especially at primary level. Furthermore, the effects of motivational variables seem to be more consistent at primary level than at secondary level. This finding further supports the idea put forward by Doiz (Doiz *et al*, 2014: 222) that students' motivation “diminishes progressively with time”. Further research that takes age and motivational variables into account needs to be undertaken in order to shed more light on this matter. The differences observed when considering motivational factors independently could explain the results of previous studies (Lasagabaster and Doiz, 2015; Arribas, 2016) that found that differences between CLIL and non-CLIL groups were not statistically significant when the effect of motivation was analysed.

In this study, on the whole, the partial effect of the motivational variables on the subtests has been low or moderate. In a way, it can be argued that motivational variables can be considered covariates because they had an effect on students' language achievement. However, not all variables had the same effect or to the same degree, since there was no statistically significant effect on the learners' final language attainment in all cases. Different variables other than motivational ones, such as the type of instruction, might be stronger

predictors of L2 attainment. In fact, the results show that when the methodology (CLIL or non-CLIL) is considered as an independent variable and the effect of motivational aspects is discounted, there are differences in means between CLIL and non-CLIL learners in favor of the former. This is in accord with Pfenninger's suggestion (2016) that CLIL instruction has a positive effect on students' language achievement. The results of this study seem to indicate that type of instruction is indeed an essential factor that can help explain the differences observed in the different subtests, namely, use of English, vocabulary, listening, speaking and reading, both at primary and secondary levels.

Further research is still needed in order to advance in our understanding, not only of which motivational factors affect more noticeably students' language attainment in CLIL settings, but also the degree to which they do so. This will make it possible to design and plan more gratifying and efficient CLIL experiences for teachers and learners alike.

7. REFERENCES

- Alejo, R. and Piquer, A. (2016). "Urban vs. rural CLIL: An analysis of input-related variables, motivation and language attainment. Language", in *Language, Culture and Curriculum*, 29, 3: 245-262.
- Arribas, M. (2016). "Analysing a whole CLIL school: Students' attitudes, motivation, and receptive vocabulary outcomes", in *Latin American Journal of Content and Language Integrated Learning*, 9, 2: 267-292.
- Breidbach, S. and Viebrock, B. (2012). CLIL in Germany: Results from recent research in a contested field of education", in *International CLIL Research Journal*, 1, 4: 5-16.
- Casal, S. and Moore, P. (2009). "The Andalusian bilingual sections scheme: Evaluation and con-sultancy", in *International CLIL Research Journal* 1: 36-46.
- Coyle, D. (2013). "Listening to learners: An investigation into 'successful learning' across CLIL contexts", in *International Journal of Bilingual Education and Bilingualism*, 16: 244-266.
- Corno, L. (1993). "The best-laid plans: Modern conceptions of volition and educational research", in *Educational Researcher*, 22: 14-22.
- Daniels, L. M., Haynes, T. L., Stupnisky, R. H., Parry, R. P., Newall, N. E. and Pekrum, R. (2008). "Individual differences in achievement goals: A longitudinal study of cognitive, emotional, and achievement outcomes", in *Contemporary Educational Psychology*, 33: 584-608.
- De la Fuente, J. (2004). "Recent perspectives in the study of motivation: Goal orientation theory", in *Electronic Journal of Research in Educational Psychology*, 2, 1: 35-62.
- Dalton-Puffer, C. (2011). "Content and language integrated learning: From practice to principles?", in *Annual Review of Applied Linguistics*, DOI: 10.1017/S0267190511000092.
- Dalton-Puffer, C. and Smit, U. (2013a). Content and language integrated learning: A research agenda", in *Language Teaching*, 46: 545-559.
- Dalton-Puffer, C. and Smit, U. (2013b). "The power of beliefs: Lay theories and their influence on the implementation of CLIL programmes", in *International Journal of Bilingual Education and Bilingualism*, 16, 3: 267-284.
- Doiz, A., Lasagabaster, D. and Sierra, J. M. (2014). "CLIL and motivation: The effect of individual and contextual variables", in *Language Learning Journal*, 42: 209-24.

- Dörnyei, Z. (2001a). *Motivation*. England: Pearson Education Limited.
- Dörnyei, Z. (2001b). *Motivational strategies in the language classroom*. Newcastle-upon-Tyne: Cambridge University Press.
- Dörnyei, Z. (2009). "The L2 motivational self-system" in Z. Dörnyei and E. Ushioda (eds.), *Motivation, language identity and the L2 self*. Bristol: Multilingual Matters, 9-42.
- Fernández, A. and Canga, A. (2014). "A preliminary study on motivation and gender in CLIL and non-CLIL types of instruction", in *International Journal of English Studies*, 4, 1: 21-36.
- Gardner, D. and Lambert, W. (1972). *Attitudes and motivation in second language learning*. USA: Newbury House Publishers.
- Gardner, D. and Yung, K. (2015). "Learner motivation in self-access language learning", in *Innovation in Language Learning and Teaching*, DOI: 10.1080/17501229.2015.1088545.
- Graaff, R. and Westhoff, G. (2007). "An observation tool for effective L2 pedagogy in content and language integrated learning (CLIL)", in *International Journal of Bilingual Education and Bilingualism*, 10, 5: 603-624.
- Hidi, S. (2016). "Interest: A unique motivational variable", in *Educational Research Review*, 1: 69-82.
- Kerstin, L. (2013). "CLIL in Sweden: Why does it not work? A metaperspective on CLIL across contexts in Europe", in *International Journal of Bilingual Education and Bilingualism*, 16, 3: 301-320.
- Kormos, J, Kiddle, T. and Csizer, K. (2011). "Systems of goals, attitudes, and self-related beliefs in second-language-learning motivation", in *Applied Linguistics*, 32, 5: 495-516.
- Lagasabaster, D. and Sierra, J. M. (2009). "Language attitudes in CLIL and traditional EFL classes", in *International CLIL Research Journal*, 1, 2: 4-17.
- Lasagabaster, D. (2011). "English achievement and student motivation in CLIL and EFL settings", in *Innovation in Language Learning and Teaching*, 5, 1: 3-18.
- Lasagabaster, D. and Doiz, A. (2015). "A longitudinal study on the impact of CLIL affective factors", in *Applied Linguistics*, 38, 5: 688-712.
- Lorenzo, F., Casal, S. and Moore, P. (2009). "The effects of content and language integrated learning in European education: Key findings from the Andalusian bilingual sections evaluation project", in *Applied Linguistics*, 31, 3: 418-42.
- Macintyre, P., Mackinnon, S. and Clément, R. (2009). "The baby, the bathwater, and the future of language learning motivation research", in Z. Dörnyei and E. Ushioda (Eds.), *Motivation, Language Identity and the L2 Self*. Bristol: Multilingual Matters, 43-65.
- Nieto, E. (2016). "The impact of CLIL on the acquisition of the learning to learn competence in secondary school education in the bilingual programmes of Castilla-La Mancha", in *Porta Linguarum*, 2: 21-34.
- Pavón, V. and Rubio, F. (2010). "Teacher' concerns and uncertainties about the introduction of CLIL programmes", in *Porta Linguarum*, 14: 45-58.
- Pavón, V. and Ellison, M. (2013). "Examining teacher roles and competences in Content and Language Integrated Learning (CLIL)", in *Linguarum Arena*, 4: 65-78.
- Pinner, R. (2013). "Authenticity of purpose: CLIL as a way to bring meaning and motivation into EFL contexts", in *Asian EFL Journal*, 15, 4: 138-159.
- Otwinowska, A. and Forys, M. (2015). "They learn the CLIL way, but do they like it? Affectivity and cognition in upper-primary CLIL classes", in *International Journal of Bilingual Education and Bilingualism*, DOI:10.1080/13670050.2015.1051944.

- Pfenninger, S. (2016). "All good things come in threes: Early English learning, CLIL and motivation in Switzerland", in *Cahiers de l'ILSL*, 48: 119-147.
- Pérez Cañado, M. L. (2012). "CLIL research in Europe: past, present and future", in *International Journal of Bilingual Education and Bilingualism*, 15, 3: 315-341.
- Pérez Cañado, M. L. (2014). "Teacher training needs for bilingual education: In-service teacher perceptions", in *International Journal of Bilingual Education and Bilingualism*, 19, 3: 1-30.
- Pérez Cañado, M. L. (2016). "Are teachers ready for CLIL? Evidence from a European study", in *European Journal of Teacher Education*, 39, 2: 79-112.
- Seikkula-Leino, J. (2007). "CLIL learning: Achievement levels and affective factors", in *Language and Education*, 21, 4: 328-341.

ACKNOWLEDGEMENTS

This work was supported by the Spanish Ministry of Economy and Competitiveness, under Grant FFI2012-32221, and by the Junta de Andalucía, under Grant P12-HUM-23480. We would also like to thank the school management and the students who participated in the study.