

# What key stakeholders think about CLIL programmes: Commonalities and differences of perspective

JUAN DE DIOS MARTÍNEZ AGUDO

LAURA VICTORIA FIELDEN BURNS

*University of Extremadura (Spain)*

Received: 7 May 2020 / Accepted: 11 November 2020

DOI: <https://doi.org/10.30827/portalin.v0i35.15320>

ISSN paper edition: 1697-7467, ISSN digital edition: 2695-8244

**ABSTRACT:** This study aims to critically discuss some of the most controversial issues affecting the characterisation and implementation practices of CLIL (Content and Language Integrated Learning) programmes from key stakeholders' perspective (teachers, learners and parents). The focus will be placed on examining their voices on what CLIL education actually means, its alleged potential benefits in terms of enhanced L2 competence and increased L2 exposure and use, whether content knowledge is improved by CLIL instruction, whether content is emphasised over language in CLIL assessment and, lastly, whether teachers and learners feel motivated and satisfied with the CLIL experience. With this in mind, and based on methodological triangulation, this paper follows a mixed-method research design for investigating the commonalities and differences of perspective. While all stakeholders agree with the alleged benefits of CLIL in terms of enhanced language competence, surprisingly not all think the same about the need for more L2 exposure and use in bilingual classes and whether content knowledge is improved by CLIL instruction. On the whole, CLIL is positively evaluated by all stakeholders, provided these educational programmes are well articulated and effectively implemented in practice, although there is always room for further quality improvement and teacher training, according to teachers.

**Keywords:** CLIL programmes, stakeholders' views, commonalities and differences of perspective, more L2 exposure and use, quality improvement and teacher training

## **Lo que piensan los principales interesados de los programas AICLE: Coincidencias y diferencias de perspectiva**

**RESUMEN:** Este estudio analiza críticamente diversas cuestiones controvertidas relativas a la caracterización e implementación de los programas AICLE (Aprendizaje Integrado de Contenido y Lengua Extranjera) desde la perspectiva de los principales interesados (docentes, estudiantes y padres). Concretamente se examinan sus opiniones acerca de lo que supone la educación AICLE, sus beneficios en términos de competencia lingüística y de una mayor exposición y uso de la lengua, si el conocimiento del contenido mejora gracias a tales programas, si se enfatiza más el contenido que la lengua en la evaluación, y, por último, si los docentes y estudiantes se sienten motivados y satisfechos con la experiencia AICLE. Partiendo de la triangulación metodológica este artículo sigue un diseño de método mixto para investigar las coincidencias y diferencias de perspectiva. Mientras todos los interesados coinciden en los supuestos beneficios de AICLE en términos de mejora de la competencia lingüística, sorprendentemente no todos piensan lo mismo acerca de la necesidad de una

mayor exposición y uso de la lengua y de si el conocimiento del contenido mejora con AICLE. En general, se valora positivamente la educación AICLE, si tales programas están bien articulados y se implementan eficazmente en la práctica, aunque siempre hay margen para mejoras de calidad y formación del profesorado según los profesores.

**Palabras clave:** programas AICLE, opiniones de los interesados, coincidencias y diferencias de perspectiva, mayor exposición y uso de la lengua meta, mejora de la calidad y formación del profesorado

## 1. INTRODUCTION

Leaving aside all the initial *celebratory rhetoric* (Paran, 2013) or *unbridled enthusiasm* (Pérez-Cañado, 2016a) from the mid-1990s the CLIL approach has generously been viewed as a near panacea, 'THE' so-called current solution or effective alternative to traditional language education in the 21<sup>st</sup> century (Bruton, 2013). The purported positive effects or benefits (and of course also the possible limitations or shortcomings (Bruton, 2015)) of CLIL need to be critically reviewed as well as empirically confirmed (Cenoz et al., 2014; Lasagabaster & Doiz, 2017). It is a fact that CLIL programmes initially emerged in response to the need to overcome the communicative limitations of traditional language education, obsessively focused on grammatical correctness, or as Dalton-Puffer (2011, p.185) reminds us "CLIL is the way to transcend the perceived weaknesses of traditional foreign language teaching". Such discontent or dissatisfaction with the target language results as well as somewhat benevolent readings and interpretations of the research evidence and literature (Bruton, 2015) have undoubtedly contributed to fuel the overestimated benefits and unrealistic expectations of CLIL education (Dallinger et al., 2016; Harrop, 2012; Pladevall-Ballester, 2015). The widespread acceptance and perceived success of CLIL programmes across the world have been such that, as Hüttner et al. (2013, p.267) put it, "the enthusiasm with which this innovation is implemented by stakeholders and 'made a success' is not fully understood". In fact, its amazing spread "has surprised even its most ardent advocates" (Maljers et al., 2007, p.7). The blind acceptance of CLIL from its onset, without the apparent need or desire for scientific evidence supporting its alleged benefits and revealing its possible limitations, is disturbing. In this respect, Harrop (2012, p.68) warned us that "The risk of implementing CLIL under the weight of unrealistic expectations and without specifically addressing its emerging shortcomings is one that we cannot afford to run". Hence, as Dalton-Puffer (2011, p.185) justifiably put it, "Research is therefore called upon to verify in how far CLIL can fulfil these and other expectations", although perhaps it would be much better to question "what assumptions lie behind such expectations" (p. 193). With this backdrop in mind, Pérez-Cañado (2016a) concluded that CLIL research has witnessed a metaphorical pendulum effect by moving from a time of celebratory euphoria with numerous advocates praising the virtues of this innovative educational approach (Coyle et al., 2010; Marsh, 2002) towards a more critical and pessimistic perspective on its feasibility with some detractors or sceptics (Bruton, 2015; Paran, 2013). Indeed, certain critical voices have recently emerged on the scene, calling into question some of the problematic issues of CLIL programmes in terms of characterisation, implementation, and research (Bruton, 2015; Paran, 2013; Pérez-Cañado, 2016a). For example, Paran (2013) reminded us that not all the benefits attributed to CLIL are supported by CLIL research. Despite the overwhelming evidence supporting the potential

benefits of CLIL programmes as a result of the continued exposure to L2 input (Dalton-Puffer, 2008), conflicting evidence has also emerged quite recently, contrary to most expectations and to what has been reported in the CLIL literature. This gloomier forecast indicates that CLIL's positive effects are limited without statistical significance, which might confirm the assumption made by Bruton (2015, p.122): “the two-in-one seems illusory in CLIL classes”. In relation to this, Fernández-Sanjurjo et al. (2017, p.10) have dared to suggest that CLIL education might not be rendering expected results in content learning and, accordingly, “the integration of content and language is not being fully achieved”. Given that the possible limitations and shortcomings of CLIL have scarcely been subject to criticism, more discerning voices would certainly be desirable to deepen the understanding of the complexities of CLIL programmes. On the prevailing and unjustified tendency to positively speculate and overemphasise the potential benefits of CLIL, Bruton (2015, p.120) suggested the need for “more attention to detail in the whole picture”. With this in mind, this study aims to shed more light on the perceived effectiveness of the current implementation practices of CLIL programmes by investigating the commonalities and differences of perspective from main stakeholders (teachers, learners and parents).

### **1.1. What is meant by CLIL? The diversity of CLIL realities can no longer be seen as an excuse**

Efforts have certainly been made to explain what CLIL is not rather than what it is, since this educational approach still remains elusively unspecified (Cenoz et al., 2014). Given the different conceptualisations of CLIL (as for example “a foreign language enrichment measure packaged into content teaching” Dalton-Puffer, 2011, p.184), Coyle (2008, p.101) pointed out that “there is a lack of cohesion around CLIL pedagogies. There is neither one CLIL approach nor one theory of CLIL”. Before entering into heated discussions as to what CLIL actually entails, how it really works, to what extent the rather idealised integration is fully achieved in praxis and under what instructional conditions, one key point should be highlighted. The variability of CLIL learning outcomes is usually justified on the basis of the frequently invoked argument of the diversity of CLIL programme formats and practices across Europe, as indicated by Cenoz et al. (2014). In fact, Dalton-Puffer (2019) made it clear that the current diversity of CLIL realities or experiences today is mainly reflected when looking at research results. To a certain extent, the justification of such diversity viewed in terms of contextually-dependent variations of CLIL might, as correctly suggested by Bruton (2015, p.126), “sound more like excuses for not adopting clear CLIL practices, partly because they remain elusively unspecified (Cenoz et al., 2014)”. Having said that, there is much about the diversity of CLIL programmes in terms of implementation processes and practices (namely, quantity and quality of input/exposure to the target language, among other issues) that still lacks clarity. Far more important is to understand what CLIL in fact entails and how it is supposed to work. Many of the hasty interpretations and generalised conclusions drawn on the empirical evidence collected so far in the CLIL research agenda might be questionable by overlooking the issue of diversity. Indeed, similar contextual conditions have not been sufficiently controlled for in CLIL research since a large number of the CLIL settings examined have been investigated under somewhat contrived circumstances.

## 1.2. The main pretext for CLIL: An increased exposure to L2 input?

One of the most controversial issues in SLA research has been the hotly debated question of the role of L1 in L2 classrooms. In fact, the traditional prescriptive monolingual approach (L2-only input) has been challenged by the relatively recent bilingual approach (L1 can function as a cognitive tool or valuable learning resource) (Macaro, 2009; Storch & Aldosari, 2010; Swain & Lapkin, 2000). Perhaps the main controversy lies in the overuse of the L1 through common practices such as code-switching and translanguaging. On the controversial issue of L1 use in CLIL classes, which is practically non-existent in the CLIL research agenda (Lasagabaster, 2013; Lo, 2015; Méndez & Pavón, 2012; Zanoni, 2016), it has been concluded that “Some voices consider that the L1 only has a support function for explanation and its use should be minimised, whereas other voices state that the L1 has a learning function, as it can help to build up students’ lexicon and to foster their metalinguistic awareness” (Lasagabaster, 2013, p.1). Put simply, L1 use should be viewed as a helpful learning support or scaffolding to check understanding and facilitate the learning process in CLIL classes (Méndez & Pavón, 2012). In this regard, Lasagabaster (2013, p.17) argued that

the use of the first language, if judicious, can serve to scaffold language and content learning in CLIL contexts, as long as learning is maintained primarily through the L2. (...) More research is needed on how the L1 can be used/is being used in CLIL contexts to maximize L2 language and content learning (...) that should lead teachers to become aware of their code-switching and translanguaging practices and to reflect on the reasons for their choices.

From an extensive reading of the SLA literature one may come to the conclusion that comprehensible input and output constitute, in general, essential conditions for L2 learning. Besides the fact that both L1 and L2 are in direct and constant contact in CLIL classes, one of the most frequently given arguments, if not the most important, in favour of CLIL instruction is the provision of an increased L2 exposure and communicative practice or interaction (Hüttner et al., 2013), but also the emphasis on language learning through language use. Despite being perhaps an excessively simplistic and restrictive vision of bilingual education, and before looking more closely at whether the target language is used either extensively or minimally in CLIL classes (Pladevall-Ballester & Vallbona, 2016), increased language exposure is what likely makes CLIL so attractive today in current L2 pedagogy. On the whole, the importance of quantity and quality of input received is such that Cenoz et al. (2014, p.257) even dared to suggest that “Perhaps the same number of hours of direct language instruction would be as effective or more effective without a CLIL approach”, which we could not agree more with. In the same direction, Bruton (2015, p.124) also made clear that “any positive benefits in the FL for the CLIL students might be attributable to the quality or level of the instruction in the EFL classes, rather than the exposure to the FL of the CLIL content teachers, who are often less proficient in the FL”, which may lead us to conclude that the expected positive language outcomes cannot be explained only by CLIL. Beyond invoking relevant L2 acquisition theories (namely, the Input, Interaction and Output Hypotheses), perhaps we should examine the quality of L2 exposure offered by CLIL teachers.

### **1.3. Killing two birds with one stone? Is content knowledge affected by CLIL? Assessing only content in CLIL assessment?**

Although the content feature is ‘really the mainstay of the CLIL marketing enterprise’ (Bruton, 2015, p.123), Bruton (2015) reminded us that “the purpose behind the adoption of CLIL is usually FL-driven, typically to achieve more FL exposure, even though the actual selection and sequencing of the subject matter in the CLIL class is content-driven” (p. 120). Put simply, the principle motive behind a school’s choosing to do CLIL is the need for an improved L2 competence which becomes the main incentive for parents to enrol their children in a bilingual programme. In the same vein, Hüttner et al. (2013, p.270) claimed that “language-learning goals are the defining feature of CLIL in EU policy papers”. Notwithstanding, what has been invoked as the most distinguishing feature of CLIL is certainly its dual-focused nature that is assumed to give equal attention to both language and content learning (Mehisto et al., 2008), even though such idealised integration, which has been understood in different ways (Cenoz et al., 2014), seems to be far from being successfully implemented in practice since CLIL is mainly content-driven (Coyle et al., 2010). Perhaps CLIL programmes are not currently rendering the expected results since, as Cenoz et al. (2014, p.244) argued, “research conducted in actual CLIL classrooms shows that it is difficult to achieve a strict balance of language and content”, a finding Nikula et al. (2016) also highlighted. Recent critical voices have also questioned this idealised integration by concluding that CLIL might not be yielding expected results in content learning (Fernández-Sanjurjo et al., 2017). This is despite the fact that the aforementioned duality of content and language lies at the heart of CLIL. Consequently, further research is clearly needed to understand how content and language are best learnt in integration (Cenoz et al., 2014; Nikula et al., 2016).

Learning content subjects through a target language requires a greater cognitive processing effort, which means that a number of students find their CLIL programmes difficult (Broca, 2016). In this respect, Mehisto et al. (2008, p.20) acknowledged that although “Common sense seems to say that students studying in a second language cannot possibly learn the same amount of content as students studying in their first language”, actually they do. In order to address such challenges, Bruton (2015) also warned us that excessive scaffolding or support might contribute to the simplification of the content subject knowledge. It may even be that the content component of the equation is not being fully achieved in practice, as suggested by Fernández-Sanjurjo et al. (2017). Such opposing views become evident in the following thought-provoking claims made by Dalton-Puffer (2011, p.188-189):

It is a common concern of educators and parents how being taught in the foreign language will affect learners’ knowledge, skills, and understanding of the subject. Because the medium of learning is less perfectly known than the L1, it is feared that this will lead to reduced subject competence as a result of either imperfect understanding or the fact that teachers preempt this problem and simplify content (...) How is it possible that learners can produce equally good results even if they studied the content in an imperfectly known language?

All in all, one might conclude that there is no compelling reason to believe that learning academic content through a less perfectly known language than their native language produces better results.

Assessment often generates a great deal of uncertainty among CLIL teachers (Coyle et al., 2010). As Otto and Estrada (2019) put it, the way assessment is conducted becomes one of the most controversial issues in the CLIL literature, particularly in terms of how teachers deal with the integration of content and language and which methods and tools are best suited when assessing students' learning outcomes, among other aspects. As a result of the current diversity of CLIL models, the lack of a standardised CLIL roadmap (Cenoz et al., 2014) as well as the absence of established assessment criteria, Otto and Estrada (2019) highlighted a significant disparity among the CLIL assessment practices and processes. Given the dual-focused nature of CLIL which might somehow suggest the existence of two assessment processes involved, the crux of the matter is if CLIL teachers must only assess the content, or the language, or both. Faced with this dilemma, CLIL teachers should always bear in mind content first as the primary goal, then language competence, as laid out in curricular regulations imposed by education authorities. Additionally, content teachers are not usually prepared to face language-related issues since they are specialists in the content areas but less proficient in the target language. On the basis that "the content should always be the dominant element in terms of objectives" (Coyle et al., 2010, p.115), the emphasis would then be on content knowledge over language competence in CLIL assessment. Additionally, content teachers need to ensure that CLIL learners acquire similar levels in content competence so that they do not fall behind their mainstream counterparts, which cannot be allowed to happen.

#### **1.4. Another reason for CLIL: CLIL helps to foster motivation**

Numerous voices have stressed the affective potential of CLIL education by fostering learners' motivation and positive attitudes towards language learning (Coyle et al., 2010; Dalton-Puffer, 2011; Doiz et al., 2014; Marsh, 2000; Seikkula-Leino, 2007; Thompson & Sylvén, 2019). In particular, Marsh (2000, p.10) claimed that CLIL programmes can "nurture a *feel good* and *can do* attitude towards language learning in general". In relation to this, Doiz et al. (2014) reminded us that one of the main reasons put forward by the advocates of CLIL is that learners feel more motivated as a result of their participation in bilingual programmes. In the same vein, Thompson and Sylvén (2019, p.76) stated that "One of the benefits commonly attributed to CLIL is that it increases learner motivation". All in all, most studies conducted to date in Europe have confirmed a positive relationship between CLIL and motivation, focusing predominantly on the motivational effect of CLIL programmes on language attainment (Doiz et al., 2014; Möller, 2018; Seikkula-Leino, 2007; Thompson & Sylvén, 2019). Conversely, the affective sphere may also be negatively affected by CLIL education (Doiz et al., 2014), as conflicting findings in the research literature have pointed to, contrary to most expectations (Heras & Lasagabaster, 2015; Möller, 2018; Otwinowska & Forys, 2017).

## 2. RESEARCH QUESTIONS

Based on key stakeholders' beliefs about the characterisation and implementation practices of current CLIL programmes (Massler, 2012; Pladevall-Ballester, 2015; Van Kampen et al., 2018), the following research questions guided this mixed-method study:

- RQ1: Do teachers and parents involved in bilingual programmes think that they know what the CLIL approach entails?
- RQ2: Do all stakeholders think that CLIL enhances L2 competence?
- RQ3: Do teachers and learners believe that CLIL leads to increased L2 exposure and use?
- RQ4: Is content knowledge improved by CLIL education in the view of all stakeholders?
- RQ5: Is content learning emphasised over language competence in CLIL assessment according to all stakeholders?
- RQ6: Do teachers and learners feel motivated and satisfied with the CLIL experience?

## 3. METHOD

This investigation is framed within a broader research project (MONCLIL<sup>1</sup> – Content and Language Integrated Learning in Monolingual Contexts) focusing on a three-year longitudinal large-scale evaluation of CLIL programmes conducted in those Spanish monolingual communities with the least tradition in bilingual education (Andalusia, Extremadura, and the Canary Islands).

Since quantitative studies abound in the current CLIL research agenda, a two-pronged research design (combining both quantitative and qualitative paradigms) has been followed in this paper. Based on methodological triangulation, data have been collected from the main stakeholders involved in CLIL programmes (teachers, learners and parents) and through multiple data-gathering procedures such as closed-ended opinion questionnaires, in-depth interviews with teachers and learners covering a range of issues related to teachers' professional preparation and learners' language and content competence, as well as CLIL classroom observation.

### 3.1. Context and participants

The present investigation, which forms part of a large-scale evaluation of CLIL programmes in monolingual areas of Spain, was undertaken within the monolingual region of Extremadura, which is situated in the south-west of Spain on the border with Portugal. This Spanish autonomous community was chosen because it has very little tradition of bilingual education (from 2004 onwards). English is the most frequently chosen foreign language both in primary and secondary education. It is a fact that in some CLIL classes the foreign language is used extensively while in others minimally, which mainly depends on content teachers' language proficiency and the complexity of subject contents taught at schools, among other reasons.

---

<sup>1</sup> For further information visit the website: <https://www.monclil.com>



Regarding the sample under control in this investigation, it is worth noting that the participating stakeholders come from 8 public bilingual schools from both primary (n=5 schools) and secondary (n=3 schools) education, located in both urban and rural areas. The sampling process followed was that of probability sampling (random selection). Out of a broader sample, those schools which subsequently revealed the greatest homogeneity in terms of verbal intelligence, motivation and English level were selected as the final cohort for the study. Concrete details of the sample under scrutiny are displayed in Table 1.

*Table 1. The research sample*

CLIL learners	<i>n</i>	<i>Mean age</i>	<i>CLIL experience</i>
Primary education (6th grade)	82 (52.5%)	11.4 years	3.2 years
Secondary education (4th grade)	74 (47.4%)	15.6 years	5.9 years
Total	156 (100.0%)		
CLIL teachers (*3 native speakers)	<i>n</i>	<i>Mean age</i>	<i>CLIL experience</i>
Primary education (6th grade)	11 (52.3%)	35.7 years	2.8 years
Secondary education (4th grade)	10 (47.6%)	46.5 year	6.9 years
Total	21 (100.0%)		
CLIL learners' parents	<i>n</i>	<i>Mean age</i>	<i>CLIL experience</i>
Primary education (6th grade)	40 (57.1%)	43.1 years	
Secondary education (4th grade)	30 (42.8%)	45.1 years	

### 3.2. Data collection instrument and procedure

The data gathered in this study were collected through multiple sources (teachers, learners and parents) and data-gathering tools, including Likert-scale opinion questionnaires, semi-structured group interviews with both CLIL teachers and learners, and direct observation of several CLIL lessons. All the *schools agreed to participate in the study and, accordingly, signed the informed consent form.*

Three sets of questionnaires (one for each of the cohorts) have been administered (see Pérez-Cañado, 2016b for a detailed account of the design and validation of the questionnaires). The 1-4 Likert-scale surveys (strongly disagree to strongly agree, the neutral position was excluded from the questionnaires to avoid the central tendency error) included a number of closed response items (please see Pérez-Cañado, 2016b), which were organised into seven thematic blocks: (1) students' use and competence in English; (2) methodology; (3) materials and resources, (4) evaluation; (5) teacher training; (6) mobility; and finally (7) learners' improvement and motivation towards English (coordination and organisation for the teacher questionnaire). Practically the same aspects are addressed in the teacher, learner and parent survey (60, 49, and 40 items, respectively).



The questionnaires were administered by the researchers as teacher educators, who explained in Spanish through clear and simple instructions the overall purpose and potential usefulness of the survey. Specifically, various closed-response questionnaires (one for each cohort: teachers, learners, and parents), which included both demographic questions and opinion questions, were designed and validated in Spanish and in English. For the validation of the questionnaires, a double-fold pilot procedure was conducted, which comprised a first stage in which a group of experts provided their feedback on the tool designed, providing valuable comments and suggestions (reorganisation and rewording of items, elimination of certain items due to overlapping of questions, correction of typographical errors...). Subsequently, a second pilot phase with a representative sample of nearly 300 subjects (students, teachers, and parents with identical traits as the target respondents) allowed us to further refine the wording of the various items of the opinion questionnaires, and hence avoid ambiguities, confusion and redundancies. In order to guarantee internal consistency and reliability of the surveys, the Cronbach alpha coefficients were calculated for the student, teacher and parent questionnaires (0.940, 0.931, and 0.895, respectively) (see Pérez-Cañado, 2016b). The data collected from the questionnaires have been statistically analysed using the SPSS programme in its 21.0 version.

Qualitative research generally strengthens empirical findings by helping to interpret the statistically obtained data. In particular, as Cohen et al. (2007) put it, interviews provide researchers with in-depth information and might act as a complementary research instrument for gathering relevant quality data. For the purposes of this study, twelve semi-structured interviews (six with 3 teacher groups and another six with 10 learner groups) were undertaken in six schools (4 primary and 2 secondary) with randomly selected participants from the total sample under the supervision of the researchers, whose results were verified and triangulated with the data obtained from the surveys. The topics discussed in the face-to-face and in-depth interviews were sharply focused on the issues raised in the literature review and, accordingly, helped us to answer the research questions formulated in this study. These twelve 30-minute interviews containing predetermined questions were recorded with prior consent from the interviewees. In addition, four 50-minute CLIL lessons were observed and recorded by the researchers by means of a protocol originally drawn up and validated by experts for this purpose. Accordingly, both interview protocols and classroom observation protocols were filled out by the researchers during fieldwork and analysed later.

Data from interview and observation protocols were transcribed and analysed by means of theory-based content analysis, which is a widely used qualitative research technique (Cohen et al., 2007). In the process of analysis, the most frequently mentioned responses from the participants were in fact identified, categorized and discussed. The participants' responses were grouped at a more conceptual level, thus allowing the identification of general themes from the obtained data. For reasons of space here it is not possible to thoroughly discuss each item or theme, hence the emphasis has been placed only on those statements and responses focused on the research questions raised in this study.

**3.3. Findings and discussion**

Data collected from the surveys, interview protocols, and CLIL classroom observation protocols are reported below. As can be seen, Table 2 displays the data obtained from participants' responses to the questionnaires. For analyses purpose, positive responses (strongly agree and agree) were combined, while negative responses (strongly disagree and disagree) were also grouped.

*Table 2. Results from the questionnaires.*

Areas	Aspects	Teachers			Learners			Parents					
		*D	**NA	***A	Total	*D	**NA	***A	Total	*D	**NA	***A	Total
Area 1. Knowledge and preparation	Knowledge about CLIL	23.81%	4.76%	71.43%	100.00%					45.71%	4.29%	50.00%	100.00%
		Statement. I have knowledge/am well informed of the core principles of Content and Language Integrated Learning in bilingual education settings.											
		Statement. Content teachers need further CLIL training											
Area 2. Competences and assessment	L2 competence	23.81%	4.76%	71.43%	100.00%								
		Statement. Learners' English has improved as a consequence of their participation in the bilingual programme											
		4.76%	0.00%	95.24%	100.00%	10.59%	0.00%	89.41%	100.00%	7.14%	0.00%	92.86%	100.00%
Area 2. Competences and assessment	L2 exposure and use	47.62%	4.76%	47.62%	100.00%	43.53%	0.59%	55.88%	100.00%				
		Statement. More use of English is needed in the bilingual classroom											
		Statement. Learners' knowledge of the subject contents taught in English has improved due to their participation in the bilingual programme											
Area 2. Competences and assessment	Content competence	14.29%	9.52%	76.19%	100.00%	18.82%	0.59%	80.59%	100.00%	31.43%	2.86%	65.71%	100.00%
		Statement. When assessing, mastery of content is emphasised over language competence											
		33.33%	14.29%	52.38%	100.00%	37.06%	4.12%	58.82%	100.00%	24.29%	17.14%	58.57%	100.00%
Area 3. Motivation and evaluation	Students' motivation towards English	4.76%	0.00%	95.24%	100.00%	12.94%	1.18%	85.88%	100.00%	18.57%	2.86%	78.57%	100.00%
		Statement. Learner motivation towards English learning has increased as a result of their participation in a bilingual programme (learner and parent) / My students are interested in the bilingual class (teachers)											
		19.05%	4.76%	76.19%	100.00%	9.41%	1.76%	88.82%	100.00%	12.86%	1.43%	85.1%	100.00%
Statement. Being part of a bilingual section compensates for the resulting workload													

\*D=Disagreement (strongly disagree and disagree)

\*\*NA=No answer

\*\*\*A=Agreement (strongly agree and agree)

With respect to RQ1 (*Do teachers and parents involved in bilingual programmes think that they know what the CLIL approach entails?*), as displayed in Table 2, it is noteworthy that while almost three-quarters of the participating CLIL teachers acknowledged that they had an extensive knowledge about the core principles of CLIL pedagogy, a similar percentage of respondents also recognised that further CLIL training was needed. Conversely, nearly half of parents were uncertain about what CLIL pedagogy actually involves. It is a fact that many parents hold unrealistic beliefs and overestimated expectations about CLIL education (Hüttner et al., 2013; Pladevall-Ballester, 2015), which are likely fuelled by an EU-backed pro-CLIL propaganda campaign. While many parents believe that CLIL education promises an added value to their children in terms of employment (Hüttner et al., 2013), Pladevall-Ballester (2015) felt that parents unrealistically perceive CLIL as the only remedy or solution to their children's low level of English. Broca (2016) also moves in the same direction by claiming that secondary education CLIL learners' parental expectations and involvement are different from those of non-CLIL counterparts. The truth is that parents are not necessarily required to have detailed knowledge of the intricacies of bilingual programmes. All this is in line with what Cenoz et al. (2014, pp.243-244) pointed out, that "the core characteristics of CLIL are understood in different ways (...) definitions of CLIL and the varied interpretations of this approach within Europe indicate that it is understood in different ways by its advocates".

Qualitative data were also gathered from teacher interview protocols and classroom observation protocols. Surprisingly, one shared observation reported by many CLIL teachers in interviews was that they were not at all sure whether CLIL programmes were rightly or wrongly being implemented these days. Additionally, most respondents agreed that the CLIL pedagogy training they received is adequate, but not sufficient, as it could always be improved. In particular, several CLIL teachers stressed the need for more practice-oriented training, above all in terms of *how to engage learners* with CLIL lessons given their lack of motivation and participation.

Regarding RQ2 (*Do all stakeholders think that CLIL enhances L2 competence?*), it should be noted that all the stakeholders involved in CLIL programmes expressed their full agreement with the assumption that bilingual education greatly helps improve the target language competence, which is in line with most expectations and with what has been extensively reported in the literature to date (Coyle et al., 2010; Dalton-Puffer, 2008; Marsh, 2000; Mehisto et al., 2008; Pladevall-Ballester, 2015). In particular, Pladevall-Ballester (2015) reported that primary school CLIL students, except for low achievers, expressed overall satisfaction with CLIL programmes in terms of use of English and understanding ability. Put simply, CLIL's positive effects on L2 competence are widely-acknowledged by key stakeholders, as displayed in Table 2.

On whether CLIL leads to increased L2 exposure and use (RQ3), contrary to expectations, opposing views were found in this regard both among teachers and among learners. Surprisingly, around half of the respondents in each group believed that there is no need to increase L2 use, which would lead us to think that L1 use is accepted and allowed in CLIL settings, perhaps viewed as a helpful resource to facilitate the understanding and learning processes. This finding runs counter to the view by Pladevall-Ballester and Vallbona (2016) who suggested that more promising CLIL learning outcomes (specifically, EFL receptive skills) might only be observable in the long run with more intensive exposure to the target

language. As for what percentage of class time is taught in English, as shown in Figures 1 and 2, nearly half of the participating CLIL teachers recognised a more extensive use of the target language ( $\geq 50\%$  of class time) in comparison with CLIL learners. In relation to this, several CLIL teachers made it clear that the use of English mainly varies according to the content subject being taught in terms of complexity of technical vocabulary (for example, chemistry and biology, when compared to physical education and history).

In the background section of the questionnaire, both teachers and learners were asked about the amount of exposure to English learners are usually exposed to in the bilingual class, thus recognising differences of view in this regard, as can be seen in Figures 1 and 2.

Figure 1. L2 exposure time (CLIL teachers)

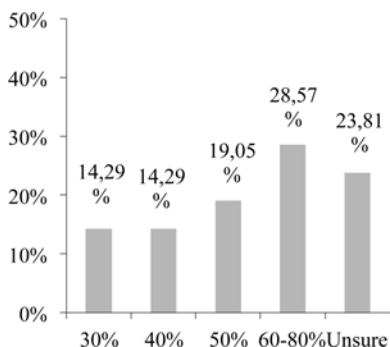
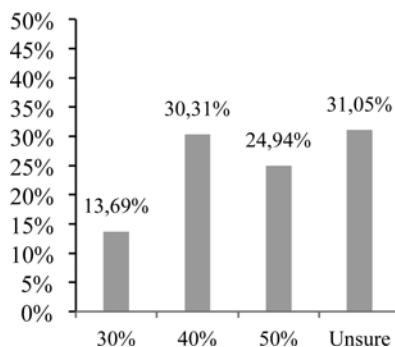


Figure 2. L2 exposure time (CLIL learners)



Teacher and learner interviews and CLIL lesson observation also provided enlightening data for both RQ2 and RQ3. It is widely recognised by CLIL teachers that learners' L2 competence had substantially improved as a result of their engagement and participation in bilingual education programmes. Concerning the common practice of codeswitching, the teacher respondents made clear that L2 use mainly depends on the class in terms of student participation but also acknowledged that L1 is sometimes overused for convenience in the classroom context. Similarly, CLIL learners felt that their language competence had improved (Pladevall-Ballester, 2015) and were satisfied with their teachers' language proficiency, though they also pointed out that it takes some time and effort to assimilate the theoretical contents of the subjects being taught through a target language, especially vocabulary, which might account for the fact that some learners are not so participative in CLIL classes.

Bearing in mind that CLIL is assumed to stimulate cognitive flexibility (Coyle et al., 2010), RQ4 addresses whether content knowledge is improved by CLIL instruction in the view of all stakeholders. On the effects of CLIL education on content subject learning, conflicting evidence has been recently reported by CLIL research, with empirical data not only in favour (Martínez, 2020; Ouazizi, 2016; Surmont et al. 2016) but also against bilingual programmes (Dallinger et al., 2016; Fernández-Sanjurjo et al., 2017; Piesche et al., 2016). As mentioned above, the expected results in content learning may be negatively affected by CLIL programmes (Fernández-Sanjurjo et al., 2017). Such suspicion is surprisingly

evidenced in about one-third of the participating parents (and a lower percentage of CLIL teachers) who expressed their disagreement with the purported benefits of CLIL in terms of content competence, which is fully congruent with the position of Dalton-Puffer (2011). In the same vein, Pladevall-Ballester (2015) pointed out that parents also fear that bilingual programmes might be an obstacle or threat to their children's L1 and content knowledge. This reservation is mainly backed up by the assumption that CLIL learners find it difficult to learn academic contents through a target language (Broca, 2016) due, among other reasons, to the high linguistic demands of the content areas and, above all, to the greater cognitive effort required to be able to process or assimilate the theoretical contents of the subject matter (Coyle et al., 2010). This greater effort makes learners progress more slowly, hence the need to receive more input and additional support to advance in their learning process (Dallinger et al., 2016).

Contrary to such emerging distrust, one shared response by most teachers in interviews was that content knowledge is not negatively affected by CLIL instruction, with content competence being within the normal values when compared to that of non-CLIL learners. However, what is widely recognised by many learners is the additional effort and difficulty in learning the theoretical contents of the subjects being taught through a target language.

Concerning the weight of content in CLIL programmes, and more precisely, whether content competence is emphasised over language competence in CLIL assessment according to all stakeholders (RQ5), over half of all respondents fully agreed that content knowledge is emphasised over language competence not only when conducting CLIL lessons but also when assessing CLIL learning outcomes. This is a logical consequence of the fact that CLIL is mainly content-driven in practice as pointed out by Coyle et al. (2010) and Bruton (2015). Additionally, this is backed up by Otto and Estrada (2019) who reported that content knowledge is always prioritised over language competence in CLIL assessment for teachers, leaving language competence behind. Accordingly, such a view is congruent with the position of Cenoz et al. (2014, p.244) who reported, "research conducted in actual CLIL classrooms shows that it is difficult to achieve a strict balance of language and content", thus challenging the esteemed integration of CLIL.

From the responses gathered in interviews with CLIL teachers and learners, it can be concluded that content competence (70/80%) is generally emphasised over language competence (30/20%) in CLIL assessment. While written communication skills (reading and writing) are preferably assessed in content subjects, oral skills (listening and speaking) receive special emphasis in the language subject in CLIL programmes. A mixture of summative and continuous assessment is mainly conducted in content subjects, giving emphasis first to content aspects and then to language aspects (in particular, reading and writing), whilst continuous assessment is mainly promoted in language subjects but also self-assessment is treated. Daily assigned tasks, project-work, attitude and behaviour, and student participation are also assessed in CLIL settings.

As regards the last research question (RQ6: *Do teachers and learners feel motivated and satisfied with the CLIL experience?*), there is a wide consensus among all the stakeholders that CLIL education fosters student and teacher motivation and, above all, that their involvement and participation in bilingual programmes compensate for all efforts made. Notwithstanding, some evidence of frustration on the part of some teachers was also found in

view of the lack of student engagement in CLIL lessons. Such frustration is also echoed by Pladevall-Ballester (2015) who pointed out that CLIL teachers' main concerns included their students' low level of English, lack of materials, and lack of institutional and peer support.

Additional qualitative data were also collected from interview and observation protocols. In fact, one widely shared remark by both teachers and learners was that the CLIL experience was positive and productive, as evidenced in several illustrating comments 'despite the great deal of work and the tremendous effort undertaken, it has been a worthwhile experience' and 'apart from personal dedication and additional preparation of the CLIL lessons, we feel glad and satisfied with the work done', highlighting their satisfaction with and motivation for CLIL education. However, despite the additional effort, it was also clarified that such experiences could be improved still further in the view of teachers.

#### 4. CONCLUSIONES

This mixed-method study aims to critically explore what the main stakeholders believe about the characterisation and implementation practices of current CLIL programmes, and accordingly, whether they feel satisfied and motivated with the CLIL experience. The teacher, learner and parent voices we draw on in this paper revealed both commonalities and differences of perspective. This is congruent with most expectations and with what has been reported in the CLIL literature concerning the frequently invoked argument of the current diversity of CLIL realities, experiences and practices, as we are reminded by Cenoz et al. (2014) and Dalton-Puffer (2019). For example, while all stakeholders fully agreed with the alleged potential benefits of CLIL education in terms of enhanced language competence, surprisingly not all think the same about the need for more L2 exposure and use in bilingual classes and whether content knowledge is improved by CLIL instruction. Overall, CLIL education is positively evaluated by all stakeholders, provided that these educational programmes are well articulated and effectively implemented in practice, although there is always room for further quality improvement and teacher training, according to teachers.

Not unexpectedly, this study presents several limitations and blind spots, namely the sample size, and the specificity of CLIL programme formats and practices under scrutiny, among others. Needless to say, more qualitative CLIL research is needed to deepen the understanding of the reality of CLIL programmes from the perspective of key stakeholders. In particular, further attention should be placed on conditions of exposure and quality of input, rather than simply the amount of exposure.

Bearing in mind the seemingly contradictory evidence recently reported by the research agenda which challenges, to a certain extent, the very theoretical underpinnings of the CLIL approach, research must move forward towards a more critical positioning that thoroughly discusses not only the potential of CLIL but also its possible limitations and shortcomings. Certainly, many questions still remain unresolved in the research agenda as CLIL is still a highly promising research area with many exciting avenues to explore in the near future, as Pérez-Cañado (2016a) reminds us.

## 5. REFERENCES

- Broca, A. (2016). CLIL and non-CLIL: Differences from the outset. *ELT Journal*, 70(3), 320-331. <https://doi.org/10.1093/elt/ccw011>
- Bruton, A. (2013). CLIL: some of the reasons why .... and why not. *System*, 41(3), 587-597. <https://doi.org/10.1016/j.system.2013.07.0>
- Bruton, A. (2015). CLIL: Detail matters in the whole picture. More than a reply to J. Hüttner and U. Smit (2014). *System*, 53, 119-128. <https://doi.org/10.1016/j.system.2015.07.005>
- 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>
- Cohen, L., Manion, L., & Morrison, K. (2007). *Research methods in Education* (6th ed.). Routledge.
- Coyle, D. (2008). CLIL-A pedagogical approach from the European perspective. In N. Van Dusen-Scholl, & N. H. Hornberger (Eds.) *Encyclopedia of Language and Education* (pp. 97-111). Springer.
- Coyle, D., Hood, P., & Marsh, D. (2010). *CLIL: Content and Language Integrated Learning*. Cambridge University Press.
- Dallinger, S., Jonkman, K., Hollm, J., & Fieg, Ch. (2016). The effects of content and language integrated learning on students' English and history competences –Killing two birds with one stone? *Learning and Instruction*, 41, 23–31.
- Dalton-Puffer, C. (2008). Outcomes and processes in content and language integrated learning (CLIL): current research from Europe. In W. Delanoy, & L. Volkman (Eds.) *Future Perspectives for English Language Teaching* (pp. 139–157). Universitätsverlag Winter.
- Dalton-Puffer, C. (2011). Content-and-language integrated learning: from practice to principles? *Annual Review of Applied Linguistics*, 31, 182-204.
- Dalton-Puffer, C. (2019). CLIL in practice: what does the research tell us? Goethe-Institut. Available online: <https://www.goethe.de/en/spr/unt/kum/clg/20984546.html>
- Doiz, A., Lasagabaster, D., & Sierra, J. M. (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>
- Fernández-Sanjurjo, J., Fernández-Costales, A., & Arias-Blanco, J. M. (2017). Analysing students' content-learning in Science in CLIL vs. non-CLIL programmes: Empirical evidence from Spain. *International Journal of Bilingual Education and Bilingualism*, 20, 1-14. <https://doi.org/10.1080/13670050.2017.1294142>.
- Harrop, E. (2012). Content and Language Integrated Learning (CLIL): Limitations and possibilities. *Encuentro*, 21, 57-70.
- Heras, A., & Lasagabaster, D. (2015). The impact of CLIL on affective factors and vocabulary learning. *Language Teaching Research*, 19(1), 70–88.
- Hüttner, J., Dalton-Puffer, Ch., & Smit, U. (2013). The power of beliefs: lay theories and their influence on the implementation of CLIL programmes. *International Journal of Bilingual Education and Bilingualism*, 16(3), 267-284. <https://doi.org/10.1080/13670050.2013.777385>
- Lasagabaster, D. (2013). The use of the L1 in CLIL classes: The teachers' perspective. *Latin American Journal of Content and Language Integrated Learning*, 6(2), 1-21. <https://doi.org/10.5294/laclil.2013.6.2.1>
- Lasagabaster, D., & Doiz, A. (2017). A longitudinal study on the impact of CLIL on affective factors. *Applied Linguistics*, 38(5), 688-712.



- Lo, Y. Y. (2015). How much L1 is too much? Teachers' language use in response to students' abilities and classroom interaction in content and language integrated learning. *International Journal of Bilingual Education and Bilingualism*, 18(3), 270-288.
- Macaro, E. (2009). Teacher use of codeswitching in the second language classroom: Exploring 'optimal' use. In M. Turnbull, & J. Dailey-O'Cain (Eds.) *First language use in second and foreign language learning* (pp. 35-49). Multilingual Matters.
- Maljers, A., Marsh, D., & Wolff, D. (Eds.) (2007). *Windows on CLIL. Content and Language Integrated Learning in the spotlight*. European Platform for Dutch Education.
- Marsh, D. (2000). *Using languages to learn and learning to use languages*. University of Jyväskylä.
- Marsh, D. (2002). *Content and Language Integrated Learning: The European dimension-actions, trends and foresight potential*. DG Education & Culture, European Commission.
- Martínez, J. d. D. (2020). To what extent does parental educational background affect CLIL learners' content subject learning? Evidence from research in Spain. *International Journal of Bilingual Education and Bilingualism*, 23. <https://doi.org/10.1080/13670050.2020.1715916>
- Massler, U. (2012). Primary CLIL and its stakeholders: What children, parents and teachers think of the potential merits and pitfalls of CLIL modules in primary teaching. *International CLIL Research Journal*, 1(4), 36-46.
- Mehisto, P., Marsh, D., & Frigols, M<sup>a</sup> J. (2008). *Uncovering CLIL: Content and Language Integrated Learning in bilingual and multilingual education*. Macmillan.
- Méndez, M. C., & Pavón, V. (2012). Investigating the coexistence of the mother tongue and the foreign language through teacher collaboration in CLIL contexts: perceptions and practice of the teachers involved in the plurilingual programme in Andalusia. *International Journal of Bilingual Education and Bilingualism*, 15, 573-592.
- Möller, V. (2018). Promoting bilingualism at the primary and secondary level: The role of intelligence, motivation and anxiety. *International Journal of Bilingual Education and Bilingualism*. <https://doi.org/10.1080/13670050.2018.1559795>
- Nikula, T., Dalton-Puffer, Ch., Llinares, A., & Lorenzo, F. (2016). More than content and language: the complexity of integration in CLIL and multilingual education. In T. Nikkula, E. Dafouz, P. Moore, & U. Smit (Eds.) *Conceptualising integration in CLIL and multilingual education* (pp. 1-25). Multilingual Matters.
- Otto, A., & Estrada, J. L. (2019). Towards an understanding of CLIL in a European context: Main assessment tools and the role of language in content subjects. *CLIL Journal of Innovation and Research in Plurilingual and Pluricultural Education*, 2(1), 31-42. <https://doi.org/10.5565/rev/clil.11>
- Otwinowska, A., & Foryś, M. (2017). They learn the CLIL way, but do they like it? Affectivity and cognition in upper-primary CLIL classes. *International Journal of Bilingual Education and Bilingualism*, 20, 457-480. <https://doi.org/10.1080/13670050.2015.1051944>
- Ouazizi, K. (2016). The effects of CLIL education on the subject matter (mathematics) and the target language (English). *Latin American Journal of Content & Language Integrated Learning*, 9(1), 110-137.
- Paran, A. (2013). Content and Language Integrated Learning: Panacea or policy borrowing myth? *Applied Linguistics Review*, 4(2), 317-342. <https://doi.org/10.1515/applirev-2013-0014>
- Pérez-Cañado, M<sup>a</sup> L. (2016a). From the CLIL craze to the CLIL conundrum: Addressing the current CLIL controversy. *Bellaterra Journal of Teaching & Learning Language & Literature*, 9(1), 9-31. <https://doi.org/10.5565/rev/jtl3.667>

- Pérez-Cañado, M<sup>a</sup>. L. (2016b). Evaluating CLIL programmes: Instrument design and validation. *Pulso. Revista de Educación*, 39, 79-112.
- Piesche, N., K. Jonkmann, Ch. Fiege, & KeBler, J.U. (2016). CLIL for All? A randomised controlled field experiment with sixth-grade students on the effects of Content and Language Integrated Science Learning. *Learning and Instruction*, 44, 108–116.
- Pladevall-Ballester, E. (2015). Exploring primary school CLIL perceptions in Catalonia: students', teachers' and parents' opinions and expectations. *International Journal of Bilingual Education and Bilingualism*, 18(1), 45-59. <https://doi.org/10.1080/13670050.2013.874972>.
- Pladevall-Ballester, E., & Vallbona, A. (2016). CLIL in minimal input contexts: A longitudinal study of primary school learners' receptive skills. *System*, 58, 37–48.
- 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>
- Storch, N., & Aldosari, A. (2010). Learners' use of first language (Arabic) in pair work in an EFL class. *Language Teaching Research*, 14, 355-375.
- Surmont, J., E. Struys, M. Van Den Noort, and P. Van De Craen (2016). The effects of CLIL on mathematical content learning: A longitudinal study. *Studies in Second Language Learning and Teaching*, 6(2), 319–337.
- Swain, M., & Lapkin, S. (2000). Task-based second language learning: the uses of the first language. *Language Teaching Research*, 4, 251-274.
- Thompson, A. S., & Sylvén, L. K. (2019). CLIL and motivation revisited: A longitudinal perspective. In L. K. Sylvén (Ed.) *Investigating content and language integrated learning. Insights from Swedish high schools* (pp. 76-95). Multilingual Matters.
- Van Kampen, E., Admiraal, W. & Berry. A. (2018). Content and language integrated learning in the Netherlands: teachers' self-reported pedagogical practices. *International Journal of Bilingual Education and Bilingualism*, 21(2), 222-236, <https://doi.org/10.1080/13670050.2016.1154004>
- Zanoni, F. (2016). Code-switching in CLIL classes: A case study. *ELLE*, 5(2), 279-295.

## FUNDING

This study was supported by the Spanish Ministry of Economy and Competitiveness [research grant number FFI2012-32221] and the Government of Andalusia (Spain) [research grant number P12-HUM-2348].

