From the CLIL craze to the CLIL conundrum: Addressing the current CLIL controversy

María Luisa Pérez Cañado
Department of English Philology, University of Jaén, Spain

Article received 28 March 2016, final version received 31 March 2016
http://dx.doi.org/10.5565/rev/jtl3.667

Abstract
This article provides an updated account of the evolution of Content and Language Integrated Learning (CLIL), from an initial period of CLIL craze to one of CLIL critique to, at present, what could be considered a CLIL conundrum. The controversies which currently affect this approach are documented on three main fronts (characterization, implementation, and research), illustrating how the so-called pendulum effect is at work in all of them. The concomitant challenges posed by these controversies are identified and specific ways to redress them are provided via concrete research-based proposals stemming from two governmentally-funded research projects. The ultimate aim is to identify the chief hurdles which need to be tackled within the CLIL arena in the very near future and to signpost possible ways of superseding them in order to continue advancing smoothly into the next decade of CLIL development.

Key words: CLIL, controversy, characterization, implementation, research

Resumen
Este artículo realiza una revisión actualizada de la evolución del Aprendizaje Integrado de Contenidos y Lenguas Extranjeras (AICLE), desde un periodo inicial de defensa de este enfoque hasta otro de crítica posterior, desembocando en una etapa de controversia. El debate que actualmente rodea a este enfoque afecta a tres frentes principales (su caracterización, su implementación y su investigación), todos los cuales se abordan, ilustrando cómo el llamado efecto péndulo se puede discernir en cada uno de ellos. Se identifican los principales retos que derivan de la controversia existente en estas tres grandes áreas del AICLE y se proponen formas concretas de afrontarlos mediante ejemplos provenientes de dos proyectos de I+D sobre el tema. El fin último es identificar los principales obstáculos que se han de superar en el campo del AICLE en el futuro inmediato y realizar propuestas concretas sobre cómo afrontarlos con garantías para continuar avanzando en el desarrollo de este enfoque.

Palabras clave: AICLE, controversia, caracterización, implementación, investigación

Précis
Cet article fournit une mise à jour de l'évolution de l'Enseignement de Matières par l'Intégration d'une Langue Étrangère (EMILE), d'une période initiale de la défense de cette approche à d'autres critiques ultérieures, conduisant à une période de controverse. Le débat actuel entourant cette approche affecte trois fronts principaux (la caractérisation, la mise en œuvre et la recherche), qui sont tous traités, illustrant comment le soi-disant effet pendule peut être discerné dans chacun d'eux. Les principaux défis posés par le conflit dans ces trois domaines d'EMILE sont identifiés et des moyens concrets pour y faire face sont proposés par des exemples tirés de deux projets de recherche sur le sujet. Le but ultime est d'identifier les principaux obstacles à surmonter dans le domaine de l'EMILE dans l'avenir immédiat et faire des propositions concrètes pour les traiter avec des garanties pour continuer à progresser dans le développement de cette approche.

Mots-clefs: EMILE, controverse, caractérisation, implémentation, recherche
Introduction

The 21st century has been characterized by great upheaval in the language teaching arena. In these past 15 years, considerable strides have been taken towards the “multilingual turn” (May, 2014, p. 1) in language education. At a time when teaching through a single language is regarded as “drip-feed” (Vez, 2009, p. 8) or “second rate” (Lorenzo, 2007, p. 35) education, encouraging polylinguaging, translanguaging, or plurilingualism has become paramount, particularly in the European scenario, where the “mother tongue + 2” objective (namely, the need for EU citizens to be proficient in their mother tongue plus two other European languages) has been targeted for over three decades (European Commission, 1995).

Against this backdrop, a specific approach to language teaching has forcefully come to the fore and embedded itself in the language teaching scenario: CLIL (Content and Language Integrated Learning) in English, AICLE (Aprendizaje Integrado de Contenidos y Lenguas Extranjeras) in Spanish, or EMILE (l’Enseignement de Matières par l’Intégration d’une Langue Étrangère) in French. It has been “embraced quickly and enthusiastically by stakeholders: Parents, students, language/educational policy-makers all over the world, but especially in Europe” (Lasagabaster & Doiz, 2016, p. 1). Indeed, in our continent, it has had an exponential uptake particularly over the course of the past two decades, and has swiftly been “put into practice from primary education through vocational education to university” (Merino & Lasagabaster, 2015, p. 1). If, according to authors such as Hughes (2010), CLIL initiatives are expected to come to fruition in 20 years, the time is ripe to step back and do some much-needed stocktaking into how they have played out.

And this is especially the case since CLIL has undergone a very interesting evolution since it first entered the European scene. It was initially heralded as the potential lynchpin to tackle the foreign language deficit on our continent and was embraced as “a lever for change and success in language learning” (Pérez Cañado & Ráez Padilla, 2015, p. 1), as “awesome innovation” (Tobin & Abello-Contesse, 2013, p. 224), or as “the ultimate opportunity to practice and improve a foreign language” (Pérez-Vidal, 2013, p. 59). However, after this period of unbridled enthusiasm, over the course of the past half a decade, a more critical attitude has emerged (Cabezas Cabello, 2010; Bruton, 2011a, 2011b, 2013, 2015; Pérez Cañado, 2011, 2012; Cenoz, Genesee, & Gorter, 2013; Paran, 2013), calling into question some of the core underpinnings of CLIL and shaking CLIL advocates out of their complacency. As Paran (2013, p. 334) has put it, we have moved from a “celebratory rhetoric” which saw CLIL as a near panacea to dwelling almost exclusively “on the
problematic issues of CLIL”. This so-called “pendulum effect” (Swan, 1985, p. 86) which has characterized language teaching history has just made itself conspicuous in the CLIL scenario (Pérez Cañado, in press), leading to CLIL controversy on different fronts. Great debate has been sparked off and contradictory opinions have been harbored vis-à-vis pivotal aspects of CLIL characterization, implementation, and research, thereby creating the need to revisit some taken-for-granted issues affecting this approach and constituting challenges to be addressed in the present and very near future of CLIL theory and praxis.

It is precisely on these challenges that the present article seeks to focus. It will identify these three areas of current contention in CLIL (characterization, implementation, and research), canvassing the chief challenges they have generated and offering routes to address and overcome them through the provision of concrete examples from two ongoing governmentally-funded research projects (cf. Acknowledgements). The ultimate aim is to identify the chief hurdles which need to be tackled in the very near future and to signpost possible ways of superseding them in order to continue advancing smoothly into the next decade of CLIL development.

**CLIL challenges**

*The controversy in CLIL characterization*

An initial controversy which is affecting CLIL pertains to its very characterization. When CLIL was coined and launched in the mid-1990s by UNICOM, the University of Jyväskylä (Finland), and the European Platform for Dutch education (Marsh, 2006; Fortanet-Gómez & Ruiz-Garrido, 2009), it was defined as “a dual-focused education approach in which an additional language is used for the learning and teaching of both content and language” (Marsh & Langé, 2000, p. 2). The dual-focused component underscores the fact that CLIL has two aims: one subject- or theme-related, and the other, language-focused. The additional language, in turn, is normally not the most widely used one in the environment. Finally, the emphasis on both teaching and content points to the very hallmark of CLIL: the fact that it straddles these two aspects of learning, involving the fusion of previously fragmented elements of the curriculum and requiring teachers to forego their respective mindsets grounded on a single subject and to pool their skills and knowledge (Coyle, Hood, & Marsh, 2010).

Although it is widely consensual that CLIL is a “well-recognized and useful construct for promoting L2/foreign language teaching” (Cenoz et al., 2013, p. 16), its exact limits are
very difficult, if not impossible, to pin down. CLIL has been challenged for its “ill-defined nature” (Paran, 2013, p. 318), “convenient vagueness” (Bruton, 2013, p. 588), and internal ambiguity (Cenoz et al., 2013, p. 2). Indeed, according to Paran (2013, p. 319), it is “afflicted with a high lack of terminological clarity, starting with the confusion between CLIL, CBI, and Immersion Education”. This is why, initially, the prevalent tendency was to distill the core features which differentiate CLIL from other types of immersion approaches and which make it a foreign language teaching trend in its own right, and not a mere offshoot of other types of bilingual programs (Lasagabaster & Sierra, 2009; Pérez Cañado, 2012; Pérez-Vidal, 2013; Dalton-Puffer, Llinares, Lorenzo & Nikula, 2014). These affect the language of instruction (generally not present in the students’ context), the languages taught through CLIL (mostly major international linguae francæ, with English holding a hegemonic position), the methodology used (which involves the integration of language and content, with foreign language teaching and CLIL lessons being timetabled alongside each other), the language level targeted (a functional vs. native-like competence of the language studied), the linguistic command of teachers (which, in line with the foregoing, need no longer be native-like), the amount of exposure to the second or foreign language (lower, as age of onset of language learning tends to be pushed back in CLIL contexts), or the types of materials employed (adapted or originally designed, as opposed to authentic ones).

However, the metaphorical pendulum has of late swung to the other extreme, calling into question this reductionist, isolationist view of CLIL as detrimental for practitioners and researchers (Cenoz et al., 2013, p. 1): “We argue that attempts to define CLIL by distinguishing it from immersion approaches to L2 education are often misguided”. In this vein, Somers and Surmont (2010), Cenoz, Genesee, and Gorter (2013), Hüttner and Smit (2013), Cenoz (2015), and Cenoz and Ruiz de Zarobe (2015) expound on the similarities rather than differences between CLIL, immersion, and Content-Based Instruction (CBI), and advocate a more inclusive, integrative, and constructivist stance which does not attempt to provide “a detailed, theoretically ‘tight’ definition of what is (not) CLIL” or enter into “heated discussion into where to draw the borders and what (not) to include” (Hüttner & Smit, 2014, p. 164).

Indeed, authors such as Somers and Surmont (2010) or Cenoz et al. (2013) have capitalized on the similarities between CLIL and immersion education. To begin with, vis-à-vis the language of instruction and the language taught through CLIL, they maintain that immersion programs also incorporate target languages which are not always present in students’ contexts and they also point out that CLIL is equally used to teach regional and
minority languages which are official in certain communities (e.g., French in Belgium or Catalan/Basque in Spain). *Methodologically*, the balance between language and content also runs through immersion education, which is seen as an instance of a content-driven approach rather than a language-driven one by Met (1998) or Genesee (2004). Both types of programs call for what Cenoz et al. (2013, p. 10) consider a more “systematic, explicit and coherent integration of language and content instruction”. The *language level targeted* is no different from CLIL, according to the afore-mentioned authors, as instrumental motivation drives both CLIL and immersion programs, and an advanced functional proficiency (vs. a native-like one) is also the goal of immersion programs in North America. Late and partial immersion programs cannot hope to reach a native-like proficiency either. In line with the foregoing, the demystification of the native speaker as the ideal *teacher* is now also present in immersion, as it is in CLIL. Furthermore, the *starting age* is no longer different in both types of programs, since there are both middle and late immersion programs and early CLIL ones (from infant education). Finally, in regards to *materials*, Somers and Surmont (2010) and Cenoz et al. (2013) uphold that they are originally designed in both contexts. Thus, in view of these arguments, it is “difficult, if not impossible, to identify features that are uniquely characteristic of CLIL in contrast with immersion education” (Cenoz et al., 2013, p. 13).

The same occurs with CLIL and CBI. As Cenoz et al. (2013: 11) underscore, “some consider CLIL to be the same as CBI and, thus, immersion, which is clearly a form of CBI”. Indeed, according to Ruiz de Zarobe (2008, p. 61). “Content and Language Integrated Learning (CLIL) and Content-based Instruction (CBI) can be considered synonymous. The former is used more frequently in Europe while the latter has gained more popularity in the United States and Canada”. Indeed, Cenoz (2015) has recently maintained that the essential properties of CLIL and CBI (use of the L2 as a medium of instruction, societal aims, or the typical type of student) are the same. There are only what she terms accidental differences between both approaches, which are “linked to the specific educational contexts where the programmes take place” (Cenoz, 2015, p. 22). CLIL and CBI are thus now considered “labels for the same reality” (Cenoz & Ruiz de Zarobe, 2015, p. 90).

Thus, the way out of this “terminological puzzle” (Dalton-Puffer et al., 2014, p. 2) is held to lie in “integration” (Cenoz & Ruiz de Zarobe, 2015, p. 90). A much broader, all-encompassing view of CLIL is now proposed, where this acronym is regarded as an “umbrella construct” which includes immersion education (Cenoz et al., 2013, p. 13), a “blanket term” (Cenoz et al., 2013, p. 5), or a “holistic view of what we do with language use and languages in a pluralistic sense” (Coyle, in Piquer Vives & Lorenzo Galés, 2015, p. 89).
Coyle (in Piquer Vives & Lorenzo Galés, 2015, p. 89) goes as far as to claim, in this sense, that “My vision for the future is that CLIL, as a concept or term, won’t even be used”. Thus, the onus is now on recognizing the diversity of formats which can be subsumed within CLIL and on ensuring that the results and effects of all types of multilingual programs (be they CLIL, CBI, or immersion) are shared so that the pedagogical and research community can benefit from them. As Cenoz et al. (2013, p. 16) claim, “A taxonomy or delineation of alternative formats for CLIL would help bring order to these matters”.

The controversy in CLIL implementation

This lack of conceptual clarity affecting CLIL trickles down to on-the-ground practice and has clear implications for CLIL implementation. Just as the definition of CLIL has been plagued with ambiguity, so has its implementation been criticized for lacking cohesion (Coyle, 2008), clarity (Bruton, 2011b), and coherence (Cenoz et al., 2013). Although CLIL may be historically unique, it is not unique pedagogically (Cenoz et al., 2013).

Indeed, criticism has recently been leveled at CLIL due to the plethora of models or variants which can be identified within it. This wide spectrum of models which CLIL encompasses is held to be dependent on a series of factors or parameters. For Coyle et al. (2010), these are operating factors -among which they subsume teacher availability, levels of teacher and student language fluency, amount of time available, ways of integrating content and language, out-of-school opportunities and networking with other countries, and assessment processes- and scale of the CLIL program –which rests on extensive instruction through the vehicular language, where the latter is almost exclusively used, or partial teaching, where code-switching or translanguaging are present to a greater extent. According to Wolff (2005), CLIL variants are determined by environmental parameters, which he outlines as involving the degree of FL and content teaching, choice of subjects, time of exposure, and linguistic situation (monolingual/monocultural – multilingual-multicultural). In turn, Smit’s (2007) proposal also encompasses extent of content and language teaching and adds as further criteria population segments (elite-mainstream), age groups, monolingual-multilingual settings, types of teachers involved, learner assessment, type and amount of target language usage, and language taught. Finally, for Rimmer (2009, p. 4), the variables which intervene in what he terms “the CLIL mix” are degree and depth of content, L1/L2 balance, involvement of subject specialists, and the extent to which CLIL is present in the curriculum.
The flexible combination of these factors is conducive to a remarkably broad array of CLIL programs and this has been regarded as detrimental by certain scholars for the pedagogically coherent evolution of CLIL: “Identifying the programmatic, instructional, and student-related properties that are specific and perhaps unique to CLIL is complicated by the diverse and ill-defined range of learning contexts/opportunities that can be classified as CLIL” (Cenoz et al., 2013, pp. 12-13).

However, another notable batch of authors has recently countered this view, crafting a compelling argument that the variegated types of approaches which can be subsumed within CLIL have, far from hampering its development, helped it to accommodate the linguistic diversity of the European landscape (Wolff, 2005; Coyle & Baetens-Beardsmore, 2007; Lasagabaster, 2008; Pérez Cañado, in press), thereby avoiding the one-size-fits-all model (Smit, 2007) which has “failed miserably” (Lorenzo, Moore, & Casal, 2011, p. 454). This “context-sensitive stance on CLIL”, as Hüttner and Smit (2014, p. 164) term it, is necessary, as “CLIL practice is informed by local realisations of language teaching methodologies (…) and, most importantly of all, a host of content subjects” (Hüttner & Smit, 2014, p. 163). It is furthermore fully commensurate with Kumaravadivelu’s (2001, p. 538) post-method pedagogy of particularity, a claim which Durán-Martínez and Beltrán-Llavador (2016, p. 89) also endorse:

The CLIL approach is stretching some commonly assumed practices and theories of teaching and of second language acquisition beyond their boundaries to the extent that the concept of method itself is being challenged and suggestions have been made to replace it with the pedagogic parameters of particularity, practicality and possibility as organizing principles for L2 teaching and teacher education.

Thus, CLIL is, in Dickey’s (2004, p. 13) terms, like a “blanket on a large bed shared by many children, each pulling in their own direction”, and it is precisely its flexible nature and numerous variations which have allowed it to “stretch to meet all needs” rather than be “torn to shreds” (Ibid.).

Irrespective of the camp with which one sides, it remains incontrovertible that we stand in need of characterizing “representative pedagogical practices” (Bruton, 2011a, p. 5) of CLIL and of knowing exactly “what it looks like in practice” (Bruton, 2011b, p. 254). Its linguistic, methodological, and organizational traits need to be further honed, sharpened, and fine-tuned in line with the demands of the diverse contexts where it is being applied.

This is, however, not the only controversy which has been aroused vis-à-vis CLIL implementation. Contention has also repeatedly underpinned the discussion on the (lack of)
egalitarianism in CLIL teaching. An initial set of authors maintain that CLIL promotes social inclusion and equity, as the introduction of this approach in mainstream education provides a greater range of students with opportunities for linguistic development which they were previously denied. In this sense, Marsh (2002, p. 10) claims that “Egalitarianism has been one success factor because this approach is seen to open doors on languages for a broader range of learners”. Coyle et al. (2010, p. 2) also incide on this issue, underscoring that CLIL is appropriate “for a broad range of learners, not only those from privileged or otherwise elite backgrounds”. To take a case in point, in Andalusia, where bilingual programs have been running for a decade, CLIL is currently being applied school-wide in all compulsory public education stages. The goal for 2020 is to extend CLIL to the whole of Primary and Compulsory Secondary Education (CSE), an objective which is well underway in Primary Education, where all public schools which have been implementing CLIL programs for five or more years now only have bilingual classes for the whole of this educational stage, with monolingual streams no longer existing.

However, this egalitarianism has again been called into question by another set of scholars, who have sounded a note of caution as regards the level of self-selection in CLIL strands, with its corollary inadequacy for attention to diversity (Lorenzo, Casal, & Moore, 2009; Hughes, 2010). Mehisto (2007, p. 63) warns that “CLIL can attract a disproportionally large number of academically bright students” and Bruton (2011a, 2011b, 2013, 2015) and Paran (2013) are particularly adamant on this score. The thrust of their argument is that CLIL branches normally comprise the more motivated, intelligent, and linguistically proficient students and that these differences are conducive to prejudice and discrimination against non-CLIL learners. The latter are considered “remnants” by Bruton (2013, p. 593), who maintains that CLIL is favoring elitism: “Implicitly, CLIL is likely to be elitist and cream off certain students” (Bruton, 2013, p. 595); “rather than increasing the equality of opportunity, CLIL in certain contexts is subtly selecting students out” (Bruton, 2013, p. 593).

In order to address this second controversy affecting CLIL implementation, it becomes incumbent on practitioners to cater to diversity and to ensure CLIL enhances language and content learning in over- and under-achievers alike. As Durán-Martínez and Beltrán-Llavador (2016, p. 88) put it, we are now faced with the “difficulty of catering for inclusive alternatives for SEN children and the need to become fully confident and proficient in their use of English”.

Bellaterra Journal of Teaching & Learning Language & Literature. 9.1 (Feb-Mar 2016) ISSN 2013-6196
The controversy in CLIL research

However, is there is an area where the so-called pendulum effect has been at work, that is CLIL research (cf. Pérez Cañado, in press). Two clear moments can be discerned if we canvass the research hitherto conducted into the effects of CLIL. In an initial phase, CLIL advocates vastly outnumber its detractors or skeptics, and investigations on CLIL paint its outcomes in the most positive light possible, almost exclusively singing the praises of this approach (Marsh & Langé, 2000; Madrid & García Sánchez, 2001; Coyle, 2002, 2006, 2008, 2009a, 2009b, 2010; Marsh, 2002, 2008; Wolff, 2003; Coonan, 2005; Järvinen, 2005; Lyster, 2007; Muñoz, 2007; Lasagabaster, 2008; Ruiz de Zarobe, 2008; Gimeno Sanz, 2009; Navés, 2009; Coyle et al., 2010). This is what Cenoz et al. (2013, p. 14) term “the bandwagon effect”: since CLIL makes “all the right noises” (Rimmer, 2009, p. 5) to stakeholders, these authors hastened to jump on the CLIL bandwagon, given the “evangelical picture” (Banegas, 2011, p. 183) that was offered of this approach.

In the past few years, the pendulum has violently swerved to the opposite extreme, initiating a second phase in CLIL research which harbors a pessimistic outlook on its effects and feasibility (Cabezas Cabello, 2010; Bruton, 2011a, 2011b, 2013, 2015; Paran, 2013); questions the validity of the research conducted (Pérez Cañado, 2011, 2012; Bruton, 2011a, 2011b, 2013, 2015; Paran, 2013; Pérez Cañado & Ráez Padilla, 2015); and warns against the wholesale adoption of CLIL and the dangers inherent in the rush to embrace it: “It is very possible that deficient FL teaching might become even more deficient, especially for the less academically able, the less linguistically proficient, or the less economically privileged” (Bruton, 2013, p. 595).

Some of the studies which were interpreted by their authors as yielding positive effects of CLIL programs (e.g., Ruiz de Zarobe, 2007; Alonso, Grisaleña, & Campo, 2008) are now reinterpreted from a negative stance (cf. Bruton 2011a). The methodological shortcomings of these studies are now also pinpointed for the first time and attention is drawn to the fact that they could potentially compromise the validity of the outcomes obtained (cf. Bruton, 2011b, 2013; Pérez Cañado, 2012; Pérez Cañado & Ráez Padilla, 2015). These caveats can be classified in terms of variables (the homogeneity of the experimental and control groups has very rarely been guaranteed; moderating variables have not been factored in or controlled for; the L1 and content knowledge of the subjects taught through CLIL have rarely been worked in as dependent variables), research design (there is a clear need for longitudinal as opposed to cross-sectional research; treatment and comparison groups should be matched and compared; an eclectic or mixed research design should be favored; multiple
triangulation should be employed), and statistical methodology (statements on the effects of CLIL should be based on empirical analyses and not merely on stakeholder appreciations; inter-rater reliability and inter-coder agreement should always be calculated; multivariate procedures should be used to isolate those variables which are truly responsible for the possible differences ascertained) (cf. Pérez Cañado & Ráez Padilla, 2015 and Pérez Cañado, in press for a more detailed account of these lacunae).

In order to bring this metaphorical pendulum to a standstill in CLIL research, it behooves future investigators to ensure these methodological flaws are superseded in their research. Rather than interpret the same (methodologically skewed) studies from opposing perspectives, new ones devoid of research design and statistical problems should be conducted in order to have unbiased, balanced, and methodologically sound research shed light on the true effects of CLIL.

**Future research routes and ways forward**

How exactly to go about this? This next section provides concrete instances on how to step up to the challenges posed by the controversy on CLIL characterization, implementation, and research. They derive from two governmentally-funded R&D projects financed by the Ministry of Economy and Competitiveness and the Junta de Andalucía (cf. Acknowledgements).

To begin with, in order to counter the terminological and pedagogical vagueness of CLIL and to offer practitioners clear-cut guidelines which can foster successful CLIL implementation, we have conducted extensive classroom observation in diverse CLIL contexts. Indeed, direct observation has been employed as one of the data-collection techniques within the qualitative part of the investigation. Two researchers per class have observed and videotaped one hour of a content subject being taught through CLIL and one hour of the English as a foreign language class participating in the CLIL program across a broad range of different contexts: public and private schools, Primary and Secondary Education, rural and urban contexts, and 12 different provinces within three autonomous communities (Andalusia, Extremadura, and the Canary Islands).

The outcomes of the observation have been collated between the researchers and followed up with a short face-to-face interview with the teacher being observed in order to complete the closed-response items with more open-ended information. The basis for this behavior observation has been a protocol which has been originally drawn up and validated...
via the expert ratings approach within the projects (cf. Appendix). It has allowed the researchers to paint a more precise picture of what CLIL looks like at the grassroots level vis-à-vis aspects such as L2 use in class; L2 development and discursive functions; oral, written, and general competence development in CLIL classes; methodology and types of groupings; materials and resources; coordination and organization; and evaluation. The insights gleaned from such fine-grained and extensive observation, which can be replicated in many additional contexts, should allow us to make headway in characterizing representative pedagogical CLIL practices and to shed light on this initial controversy.

In turn, as regards the equity conundrum, steps can be taken to redress it on two major fronts. First and foremost, future research should determine whether there is indeed self-selection in CLIL and whether, as Bruton (2011a, 2011b, 2013) has put it, the more motivated, intelligent, and linguistically proficient students can be found in CLIL groups. This has been done in our projects by including an initial year-long phase devoted to determining the homogeneity of CLIL and non-CLIL groups. Contrary to what Bruton has claimed, it has been possible to match bilingual and non-bilingual learners within and across schools. CLIL and non-CLIL classes have been found to be homogeneous on four different variables: verbal intelligence, motivation (where four factors have been considered: will, anxiety, disinterest, and self-demand), socioeconomic status, and extramural exposure to the foreign language. Thus, a random, non-probabilistic sample of over 1,500 CLIL and non-CLIL students in 36 schools in both urban and rural areas of 12 provinces and three autonomous communities has been found to be homogenous on all these fronts, thereby shooting down the belief—in our context, clearly unsubstantiated—that the most intelligent, motivated, and socially privileged students are those found in CLIL streams. It would thus be necessary to administer tests such as those applied in the first phase of our projects in other contexts where comparative research on the effects of CLIL is going to be undertaken.

It would furthermore be desirable to factor in moderator or intervening variables in future research in order to determine how CLIL is functioning with the diverse types of learners who are now increasingly involved in dual-focused programs and to whose diversity it now becomes essential to cater. In this sense, our studies have considered 11 different intervening variables in order to determine the possible modulating (differential) effect they exert on CLIL and non-CLIL students’ L1, L2, and content learning: type of school (public, private, charter), province, setting (rural-urban), gender, sociocultural status, motivation, verbal intelligence, English level, time of exposure to English inside and outside school, and linguistic competence of the teacher. Singling out these learner variables will undoubtedly...
contribute to fostering successful learning for all the different types of students who are increasingly participating in bilingual programs.

Finally, vis-à-vis the research arena, the swings of the pendulum could be mitigated by carving out a research agenda which supersedes the methodological shortcomings which have compromised the validity of CLIL investigations which have hitherto been conducted. We again provide a practical example from our research projects.

To begin with, in terms of variables, our studies have begun by guaranteeing the homogeneity of the CLIL and non-CLIL strands in terms of verbal intelligence, motivation, socioeconomic status, and extramural exposure to the language. Twice the amount of schools which finally partook in the projects were selected (two public rural ones for Primary Education, two public urban ones for Primary Education, two public rural ones for CSE, two public urban ones for CSE, two private ones, and two charter ones per province), the four tests were administered, and the existence of statistically significant differences was calculated within and across groups and schools: between the CLIL and mainstream EFL classes in public schools; between the CLIL groups in private and public schools; and between the non-CLIL classes in charter and public schools. The schools which evinced the greatest homogeneity within and across cohorts were selected for participation in the subsequent phases of the projects, thereby ensuring the comparability of the experimental and control groups within and across schools. Secondly, as mentioned in previous paragraphs, moderating variables were factored in and controlled for in order to determine their possible modulating effect on the L1, L2, and content subjects taught through English. Finally, the impact of CLIL programs (the independent variable) was gauged not only on English language (L2) competence (grammar, vocabulary, and the four skills), but also on the students’ Spanish language competence (L1) and on the level of mastery of the contents of those subjects implemented through CLIL, so that three different dependent variables were considered.

In terms of research design, our studies are longitudinal rather than cross-sectional, as they have examined the impact of CLIL on L1, L2, and content mastery over the course of a year and a half, administering post-tests at the end of 6th grade of Primary Education and 4th grade of CSE, and delayed post-tests halfway through the first year of Baccalaureate. An eclectic research design has also been followed, as our studies have combined a quantitative section of applied, primary, quasi-experimental research with a pre-/post-test control group design, to which a delayed post-test has also been added, and a qualitative part involving
survey research. The latter has also employed multiple triangulation, specifically of four types:
- *Data triangulation*, as multiple sources of information have been consulted to mediate biases interjected by people with different roles in the language teaching context: students, parents, and teachers (and within the latter, non-linguistic area teachers, English language teachers, and teaching assistants).
- *Methodological triangulation*, since multiple data-gathering procedures have been drawn on: questionnaires, interviews, and observation.
- *Investigator triangulation*, due to the fact that three different researchers have analyzed the open-response items on the questionnaire and interviews, written up their conclusions, and collated their findings.
- *Location triangulation*, given that language learning data have been collected from multiple data-gathering sites: Primary Schools, Secondary Schools, and the provincial educational administration.

Finally, concerning *statistical methodology*, Cronbach α, the Kuder-Richardson reliability coefficient, ANOVA, and the t test have been employed to ensure that findings are grounded on solid statistical evidence, and factor and discriminant analyses have been used to determine whether CLIL is truly responsible for the differences ascertained between the bilingual and non-bilingual branches or whether they can be ascribed to the intervening variables considered.

By controlling these methodological issues and thereby remediating the flaws of prior investigations on these fronts, the outcomes obtained will provide reliable, empirically solid information on the true effects of CLIL, thereby bringing the research pendulum to a standstill.

**Conclusion**

The present article has provided an updated account of the exciting and challenging time which Content and Language Integrated Learning is currently living. Its hard-and-fast appearance in the field of language education, its swift uptake across the continent (and even beyond it), and the phenomenal amount of attention it has attracted have caused a vibrant research scene to burgeon around it, leading from what we could term an initial *CLIL craze* to a period of *CLIL critique* and, at present, to a *CLIL conundrum*. It is on these controversies which currently affect the characterization, implementation, and research on CLIL that we
have focused, as well as on how to best redress the concomitant challenges they pose through concrete research-based proposals.

It transpires from the foregoing that CLIL is still a thriving area of research, that there are still many exciting avenues to explore in the future, and that much road is still to be paved in the CLIL enterprise. Controversies are always healthy and, like Cenoz et al. (2013, p. 16), we celebrate the more “critical empirical examination of CLIL in its diverse forms”, since we consider it has infused the field with renewed life and enriched the multiple perspectives from which it can be examined. However, we also believe that these seemingly contradictory camps are not irreconcilable and have offered ways out of the current CLIL conundrum. If time and patience dovetail with continuous stocktaking, rigorous research, and ongoing collaboration, we firmly believe that a solid template can be built for the future, where the CLIL agenda will continue advancing strongly and steadily.

Acknowledgements
This work has been supported by the research projects FFI2012-32221 and P12-HUM-2348, funded by the Ministerio de Economía y Competitividad and the Junta de Andalucía, respectively.

References


Tobin, N. A. & Abello-Contesse, C. (2013). The use of native assistants as language and cultural resources in Andalusia’s bilingual schools. In C. Abello-Contesse, P. M.


Author information
Dr. María Luisa Pérez Cañado is Associate Professor at the Department of English Philology of the University of Jáen, Spain, where she has also been Vicedean of the Faculty of Humanities and Education for twelve years.

Email: mlperez@ujaen.es

To cite this article:
### PROYECTO MON-CLIL: Los Efectos del Aprendizaje Integrado de Contenidos y Lenguas Extranjeras en Comunidades Monolingües: Un Estudio Longitudinal

### PROTOCOLO DE OBSERVACIÓN

#### VARIABLES DE IDENTIFICACIÓN

1. CENTRO: 

2. CURSO: 6º EP 4º ESO

3. ASIGNATURA: 

4. TIPO DE PROFESORADO:  
   - Lengua extranjera  
   - Área no lingüística  
   - Auxiliar lingüístico

5. ¿ES COORDINADOR/A DE SU SECCIÓN BILINGÜE? Sí No

6. EDAD: 

7. SEXO: Hombre Mujer

8. NACIONALIDAD: 

9. SITUACIÓN ADMINISTRATIVA:  
   - Funcionario/a con destino definitivo  
   - Funcionario/a con destino provisional  
   - Interino/a  
   - Otro: 

10. EXPERIENCIA DOCENTE GENERAL:  
    - Menos de 1 año  
    - 1-10 años  
    - 11-20 años  
    - 21-30 años  
    - Más de 30 años

11. EXPERIENCIA DOCENTE EN UN CENTRO BILINGÜE:  
    - Menos de 1 año  
    - 1-5 años  
    - 6-10 años  
    - 11-15 años  
    - Más de 15 años

### 1. USO DE LA L2 EN CLASE

1. El nivel de competencia lingüística del profesor en clase se asemeja a:  
   - A1  
   - A2  
   - B1  
   - B2  
   - C1  
   - C2

2. El profesor utiliza el inglés para el desarrollo de la clase:  
   - Entre 0%-25%  
   - Entre 25%-50%  
   - Entre 50%-75%  
   - Entre 75%-100%

3. El profesor traduce del español al inglés en el desarrollo de la clase:  
   - Mucho  
   - Bastante
Poco
Nada

4. El profesor practica el “code-switching” de manera razonada y sistemática
Mucho
 Bastante
Poco
Nada

5. El nivel de competencia lingüística del alumnado es adecuado para su etapa educativa
Mucho
Bastante
Poco
Nada

6. El alumnado utiliza el inglés en clase
Entre 0%-25%
Entre 25%-50%
Entre 50%-75%
Entre 75%-100%

7. El alumnado traduce del español al inglés en el desarrollo de la clase
Mucho
Bastante
Poco
Nada

8. El alumnado practica el “code-switching” de manera razonada y sistemática
Mucho
Bastante
Poco
Nada

Otras observaciones: ........................................................................................................................................
.......................................................................................................................................................
.......................................................................................................................................................
.......................................................................................................................................................
.....................................................................................................................................................

2. DESARROLLO DE LA L2 EN CLASE: FUNCIONES DISCURSIVAS

Se utiliza el inglés en clase para:

<table>
<thead>
<tr>
<th></th>
<th>Dar instrucciones</th>
<th>Introducir el tema</th>
<th>Transmitir contenidos</th>
<th>Realizar actividades</th>
<th>Aclarar dudas</th>
<th>Formular preguntas</th>
<th>Corregir tareas</th>
<th>Repasar conocimientos</th>
<th>Organizar los distintos tipos de agrupamiento</th>
<th>Interactuar con los compañeros/alumnos/profesores</th>
<th>Suministrar feedback sobre las actuaciones de clase</th>
</tr>
</thead>
<tbody>
<tr>
<td>9</td>
<td>Mucho</td>
<td>Bastante</td>
<td>Poco</td>
<td>Nada</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>Mucho</td>
<td>Bastante</td>
<td>Poco</td>
<td>Nada</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>11</td>
<td>Mucho</td>
<td>Bastante</td>
<td>Poco</td>
<td>Nada</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>12</td>
<td>Mucho</td>
<td>Bastante</td>
<td>Poco</td>
<td>Nada</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>13</td>
<td>Mucho</td>
<td>Bastante</td>
<td>Poco</td>
<td>Nada</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>14</td>
<td>Mucho</td>
<td>Bastante</td>
<td>Poco</td>
<td>Nada</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>15</td>
<td>Mucho</td>
<td>Bastante</td>
<td>Poco</td>
<td>Nada</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>16</td>
<td>Mucho</td>
<td>Bastante</td>
<td>Poco</td>
<td>Nada</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>17</td>
<td>Mucho</td>
<td>Bastante</td>
<td>Poco</td>
<td>Nada</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>18</td>
<td>Mucho</td>
<td>Bastante</td>
<td>Poco</td>
<td>Nada</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>19</td>
<td>Mucho</td>
<td>Bastante</td>
<td>Poco</td>
<td>Nada</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Otras observaciones: ........................................................................................................................................
3. DESARROLLO DE COMPETENCIAS EN CLASE

En la clase, se favorece el desarrollo de:

<table>
<thead>
<tr>
<th>Competencia</th>
<th>Mucho</th>
<th>Bastante</th>
<th>Poco</th>
<th>Nada</th>
</tr>
</thead>
<tbody>
<tr>
<td>La comprensión oral</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>La expresión oral</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>La comprensión escrita</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>La expresión escrita</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>La interacción comunicativa oral</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>La interacción comunicativa escrita</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>La capacidad crítica</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>La creatividad</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>La autonomía en el aprendizaje</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>La conciencia metalingüística</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Aspectos interculturales de la lengua extranjera</td>
<td>Mucho</td>
<td>Bastante</td>
<td>Poco</td>
<td>Nada</td>
</tr>
</tbody>
</table>

**Otras observaciones:**

<table>
<thead>
<tr>
<th>Observación</th>
<th>Mucho</th>
<th>Bastante</th>
<th>Poco</th>
<th>Nada</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

4. METODOLOGÍA Y TIPOS DE AGRUPAMIENTO

<table>
<thead>
<tr>
<th>Método/Actividad</th>
<th>Mucho</th>
<th>Bastante</th>
<th>Poco</th>
<th>Nada</th>
</tr>
</thead>
<tbody>
<tr>
<td>Se utiliza el aprendizaje basado en tareas en clase</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Se utiliza el aprendizaje basado en proyectos en clase</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Se da prioridad a la dimensión léxica en la clase bilingüe</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Se utiliza aprendizaje cooperativo en la clase bilingüe</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Se utiliza el método transmisivo de gramática-traducción en clase</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Se utiliza el método audiolingual en clase</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Se realizan actividades abiertas</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Se realizan actividades de respuesta única</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Se realizan actividades que requieren únicamente la activación de procesos cognitivos de nivel bajo (tales como memorizar, comprender y aplicar)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Se realizan actividades que exigen movilizar procesos cognitivos complejos (tales como analizar, evaluar y crear)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>El docente favorece el andamiaje lingüístico (mediante paráfrasis, repeticiones, ejemplos, definiciones, sinónimos y antónimos, etc.)</td>
<td>Mucho</td>
<td>Bastante</td>
<td>Poco</td>
<td>Nada</td>
</tr>
</tbody>
</table>
12. El docente favorece que los alumnos aprendan y usen estrategias de compensación y de aprendizaje (e.g., para resolver problemas de comprensión lingüística)  
   | Mucho | Bastante | Poco | Nada |
13. Se siguen las recomendaciones del Marco Común Europeo de Referencia  
   | Mucho | Bastante | Poco | Nada |
14. Se siguen las recomendaciones del Portafolio Europeo de Lenguas  
   | Mucho | Bastante | Poco | Nada |
15. Se utiliza el agrupamiento “lockstep” en clase  
   | Mucho | Bastante | Poco | Nada |
16. Se utiliza el trabajo en grupo en clase  
   | Mucho | Bastante | Poco | Nada |
17. Se utiliza el trabajo en parejas en clase  
   | Mucho | Bastante | Poco | Nada |
18. Se utiliza el trabajo individual en clase  
   | Mucho | Bastante | Poco | Nada |

Otras observaciones:

5. MATERIALES Y RECURSOS

19. Se utilizan materiales auténticos para la enseñanza bilingüe  
   | Mucho | Bastante | Poco | Nada |
20. Se utilizan materiales adaptados para la enseñanza bilingüe  
   | Mucho | Bastante | Poco | Nada |
21. Se utilizan materiales originales diseñados por el profesorado para la enseñanza bilingüe  
   | Mucho | Bastante | Poco | Nada |
22. Se tiene en cuenta la atención a la diversidad en los materiales que se emplean  
   | Mucho | Bastante | Poco | Nada |
23. Se utiliza software multimedia en clase  
   | Mucho | Bastante | Poco | Nada |
24. Se utilizan materiales de referencia online en clase  
   | Mucho | Bastante | Poco | Nada |
25. Se utilizan blogs, Wikis (herramientas Web 2.0) y webquests en clase  
   | Mucho | Bastante | Poco | Nada |
26. Se utilizan pizarras electrónicas interactivas en clase  
   | Mucho | Bastante | Poco | Nada |
27. Se utiliza comunicación mediada por ordenador en clase (e.g., e-Twinning)  
   | Mucho | Bastante | Poco | Nada |

Otras observaciones:

6. COORDINACIÓN Y ORGANIZACIÓN

28. Se constata la coordinación entre el profesorado de ANLs y los auxiliares de conversación  
<p>| Mucho | Bastante | Poco | Nada |</p>
<table>
<thead>
<tr>
<th>Núm.</th>
<th>Pregunta</th>
<th>Mucho</th>
<th>Bastante</th>
<th>Poco</th>
<th>Nada</th>
</tr>
</thead>
<tbody>
<tr>
<td>59</td>
<td>Se constata la coordinación entre el profesorado de ANLs y el profesorado de inglés como lengua extranjera</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>60</td>
<td>Se constata la coordinación entre el profesorado de inglés como lengua extranjera y los auxiliares de conversación</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>61</td>
<td>Existe integración curricular (se integran contenidos de distintas asignaturas y campos de conocimiento)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>62</td>
<td>Se apoya el aprendizaje lingüístico en clases de contenido</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>63</td>
<td>Se apoya el aprendizaje de contenidos en clases lingüísticas</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>64</td>
<td>Se enfatiza la conexión entre la L1, L2 y L3</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>65</td>
<td>Se colabora en la preparación y diseño de materiales</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Otras observaciones:**

<table>
<thead>
<tr>
<th>Núm.</th>
<th>Pregunta</th>
<th>Mucho</th>
<th>Bastante</th>
<th>Poco</th>
<th>Nada</th>
</tr>
</thead>
<tbody>
<tr>
<td>66</td>
<td>A la hora de evaluar, se da prioridad al dominio de los contenidos frente a la competencia lingüística</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>67</td>
<td>A la hora de evaluar, se incluye un componente oral</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>68</td>
<td>Se practica la evaluación diversificada</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>69</td>
<td>Se practica la evaluación formativa</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>70</td>
<td>Se practica la evaluación sumativa</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>71</td>
<td>Se utiliza la autoevaluación (e.g., a través del Portfolio Europeo de Lenguas)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Otras observaciones:**

<table>
<thead>
<tr>
<th>Núm.</th>
<th>Pregunta</th>
<th>Mucho</th>
<th>Bastante</th>
<th>Poco</th>
<th>Nada</th>
</tr>
</thead>
<tbody>
<tr>
<td>72</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>73</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>74</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>75</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>76</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**7. EVALUACIÓN**