Multilingual Education

María Luisa Pérez Cañado *Editor*

Content and Language Integrated Learning in Monolingual Settings

New Insights from the Spanish Context

Springer

Multilingual Education

Volume 38

The book series Multilingual Education publishes top quality monographs and edited volumes containing empirical research on multilingual language acquisition, language contact and the respective roles of languages in contexts where the languages are not cognate and where the scripts are often different, in order to be able to better understand the processes and issues involved and to inform governments and language policy makers. The volumes in this series are aimed primarily at researchers in education, especially multilingual education and other related fields, and those who are involved in the education of (language) teachers. Others who will be interested include key stakeholders and policy makers in the field of language policy and education. The editors welcome proposals and ideas for books that fit the series. For more information on how you can submit a proposal, please contact the Associate Publishing Editor, Natalie Rieborn. E-mail: Natalie.Rieborn@springer.com

More information about this series at http://www.springer.com/series/8836

María Luisa Pérez Cañado
Editor

Content and Language Integrated Learning in Monolingual Settings

New Insights from the Spanish Context

Springer

Editor
María Luisa Pérez Cañado
Departamento de Filología Inglesa
Universidad de Jaén
Jaén, Spain

ISSN 2213-3208 ISSN 2213-3216 (electronic)
Multilingual Education
ISBN 978-3-030-68328-3 ISBN 978-3-030-68329-0 (eBook)
https://doi.org/10.1007/978-3-030-68329-0

This Springer imprint is published by the registered company Springer Nature Switzerland AG
The registered company address is: Gewerbestrasse 11, 6330 Cham, Switzerland

Contents

Introduction

María Luisa Pérez Cañado

Abstract This introductory chapter sets the scene for the broader volume and has a three-pronged objective. To begin with, it presents the rationale and justifies the need for a monograph of this nature. It then frames the volume against the backdrop of CLIL practice and policy in Spain. Finally, it introduces the ten chapters and four blocks into which the book is structured and illustrates how they provide updated insights into CLIL characterisation, implementation, and research from a multifaceted perspective.

1 Rationale

This monograph seeks to focus on the innovations and challenges affecting a teaching approach which has enjoyed a massive uptake over the past two decades in very diverse educational settings: Content and Language Integrated Learning (CLIL). The latter has undergone a very interesting evolution since it first entered the European scene in 1994. 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. However, after this period of unbridled enthusiasm, over the course of the past half a decade, a more critical attitude has emerged, 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 characterised language teaching history has just made itself conspicuous in the CLIL scenario, leading to CLIL controversy on different fronts (cf. Pérez Cañado, 2016a, 2017). Great debate has been sparked off and contradictory opinions have been harboured vis-à-vis pivotal aspects of CLIL characterisation,

M. L. Pérez Cañado (✉)
University of Jaén, Jaén, Spain
e-mail: mlperez@ujaen.es

M. L. Pérez Cañado (eds.), *Content and Language Integrated Learning in Monolingual Settings*, Multilingual Education 38,
https://doi.org/10.1007/978-3-030-68329-0_1

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 innovations and challenges that this monograph seeks to focus. It will provide updated research evidence on CLIL characterisation, implementation, and research fundamentally stemming from two governmentally funded research projects (R&D projects FFI2012-32221 and P12-HUM-2348, funded by the Spanish *Ministerio de Economía y Competitividad* and the *Junta de Andalucía*, respectively).[1] Through them, a quantitative, longitudinal study has been conducted into the effects of CLIL on the English language competence, Spanish language competence, and content knowledge of those subjects taught through the foreign language (FL) of Primary (6th grade) and Secondary (4th grade of Compulsory Secondary Education) Education students in 12 different provinces of Spain, considered to be a representative microcosm of the multifaceted CLIL landscape. It has matched the CLIL and non-CLIL students of 53 Primary and Secondary schools in terms of verbal intelligence, motivation, and extramural exposure to English and worked with a total of 2,245 students, 333 teachers, and 595 parents (3,173 subjects in all). It has employed 11 different types of tests (verbal intelligence, motivation, Spanish Language and Literature, two content tests corresponding to the subjects taught through CLIL, and English grammar, vocabulary, reading, writing, listening, and speaking tests) in order to investigate the impact of CLIL on different cognitive, contextual, and affective variables: context (rural–urban), type of school (public, private, charter), educational level (Primary, Secondary, Baccaulaurate), motivation, verbal intelligence, extramural exposure to English, and socioeconomic status. It has furthermore done so from a longitudinal perspective, as pre-, post-, and delayed post-tests have been administered to Primary, Compulsory Secondary, and non-compulsory Secondary Education students. Finally, factor and discriminant analyses have been performed to determine the interaction among all these variables and ascertain whether CLIL is truly responsible for the potential differences observed.

From a qualitative standpoint, it has probed students', teachers', and parents' satisfaction with all the curricular and organisational aspects of CLIL schemes and carried out a detailed SWOT analysis of the way in which they are functioning, according to the key players involved in their grassroots implementation. Three types of instruments have been designed and validated for the qualitative side of the study: questionnaires, semi-structured individual and focus group interviews, and direct behaviour observation. Multiple triangulation procedures have also been employed: data, methodological, investigator, and location triangulation.

The study has thus superseded some of the main lacunae pinpointed for prior investigations of this nature. It has, to begin with, guaranteed the homogeneity of the experimental and control groups at the outset of the study (Langé, 2007; Lasagabaster,

[1] In addition, it also incorporates the findings of three additional research projects: 2018-1-ES01-KA201-050356, RTI2018-093390-B-I00, and FFI2014-54179-C2-2-P, funded by the European Union, the *Ministerio de Ciencia, Innovación y Universidades*, and the *Ministerio de Economía, Industria y Competitividad,* respectively.

2008; Lyster, 2007; Madrid Fernández, 2006; Pérez-Vidal, 2007; Ruiz de Zarobe, 2008; Ruiz de Zarobe & Lasagabaster, 2010). It has also examined the impact of CLIL not only on the foreign language (English), but also on L1 and content knowledge, factoring in an important number of moderating variables (Cenoz et al., 2013; Dalton-Puffer et al., 2014; Lasagabaster & Ruiz de Zarobe, 2010). It has equally determined whether CLIL is truly responsible for the possible differences ascertained or whether they can be ascribed to these other variables (Cenoz et al., 2013; Dalton-Puffer et al., 2014; Pérez Cañado, 2011, 2012). Finally, it is longitudinal rather than cross-sectional (Lasagabaster & Ruiz de Zarobe, 2010; Lasagabaster & Sierra, 2010), in order to determine whether the effects of CLIL pervade when this type of programme is discontinued (cf. Pérez Cañado, 2017 for a detailed proposal of the research agenda which needs to be carved out in order to address these niches).

This book presents the key findings of the study on both these quantitative and qualitative fronts. Through them, it offers new empirically grounded insights into the current state of CLIL characterisation (through an innovative proposal to link CLIL to ELF), implementation (via the observation protocols employed and the SWOT analyses conducted), and research (by examining the effects of CLIL on the L1, FL, key competences, and content subjects taught through English). And it does so by focussing on a country which is very conspicuous on the CLIL map (Spain) and within it, on monolingual contexts, where there is a more notable 'shortage of research in CLIL' (Fernández-Sanjurjo et al., 2019, p. 662).

2 Backdrop: CLIL in Spain

Indeed, Spain particularly stands out within the European landscape, since, as Coyle (2010, p. viii) contends, 'Spain is rapidly becoming one of the European leaders in CLIL practice and research'. As is the case with the broader continental ambit, this educational approach has blossomed particularly over the course of the past ten years: 'In the last decade CLIL … has undergone a rapid development in the Spanish scenario' (Ruiz de Zarobe & Lasagabaster, 2010, p. ix).

This expansion of CLIL is often attributed to the deficient language proficiency levels generally attained in Spain: 'The dismal foreign language proficiency usually obtained through its conventional teaching as a school subject has led many primary, secondary and tertiary instituitions to put CLIL programmes into practice' (Lasagabaster & Ruiz de Zarobe, 2010, pp. 290–291). Indeed, the unsuccessful foreign language learning experiences in the Spanish context have been documented by a notable number of scholars: 'dissatisfaction is the common denominator when the proficiency in English of Spanish students is scrutinized, despite many having spent quite a few years trying to learn the language' (Lasagabaster & Sierra, 2009, p. 7).

These lacunae in language learning can be traced to the lack of teaching tradition and social concern for language education, which have been clearly reflected in the official legislation guiding foreign language instruction. Indeed, for most of the

twentieth century, the latter was restricted to Baccalaureate and mainly involved the teaching of French, which was not taught following communicative principles. The general tendency since the *Ley General de Educación de 1970* was passed has been to push foreign language teaching to increasingly earlier stages of the education ladder: from the age of 11–12 with the aforementioned Law, to 6 with the 1990 *Ley Orgánica General del Sistema Educativo* (LOGSE) and the 2002 *Ley Orgánica de Calidad de la Educación* (LOCE), to the second cycle of Infant Education with the 2006 *Ley Orgánica de Educación* (LOE). An optional second foreign language has also been introduced since 2006 in the third cycle of Primary Education. Furthermore, under the auspices of the 2006 LOE, which rewards instruction in the foreign language in bilingual centres, an increasing number of CLIL projects and programmes fostering innovative education methods have been incorporated across the country.

Indeed, all regional education authorities are now endorsing plurilingual policies. The most outstanding official CLIL initiatives include:

The Spanish Ministry and British Council Project: It began in 1996–1997 with a view to providing children from ages 3 to 16 with bilingual and bicultural education. Aragón, Asturias, the Balearic Islands, Cantabria, Castilla y León, Castilla-La Mancha, Ceuta, Extremadura, Madrid, Melilla, Murcia, and Navarra are all involved in this initiative.

Programa de Inmersión Lingüística: Supported by the Spanish Ministry of Education and Science, it offers summer courses for students in the last cycle of Primary Education and the first grade of Compulsory Secondary Education.

PALE (Programa de Apoyo a la Enseñanza y el Aprendizaje de Lenguas Extranjeras): It has involved 13 autonomous communities (Andalusia, Aragón, the Canary Islands, Castilla-La Mancha, Catalonia, Extremadura, Galicia, Murcia, La Rioja, Asturias, Castilla y León, Madrid, and Valencia) and is geared at aiding CLIL teachers to improve their FL competence via 200 hours of training and a two-week study visit abroad.

Aulas Europeas: These European classrooms are language and culture immersion programmes in France and the UK aimed at Infant, Primary, and Secondary teachers of any subject. They are based on an agreement between the Spanish MEC and the French Embassy, in collaboration with the French Institute in Madrid, and affect the entire Spanish territory.

PILC (Proyectos de Innovación Lingüísitica en Centros): These language innovation projects started in La Rioja in 2004–2005 and target non-university teachers of any subject willing to implement CLIL in their classrooms.

ETC*(English Through Content)*: This initiative affects Navarre and was launched in 2001 with 36 Infant and Primary schools. Since 2003, it has been applied to all schools in this autonomous community. It involves the application of a CLIL approach organised around a series of topics through 43 lesson units adapted to the aforementioned educational levels.

Secciones Europeas / Secciones Bilingües: These sections entail CLIL instruction in Primary and Secondary schools of diverse monolingual communities (Aragón, Andalusia, Asturias, Canarias, Cantabria, Castilla-La Mancha, Extremadura or Madrid) and in certain bilingual ones (such as Galicia or the Balearic Islands).

Proyecto Bilingüe: This is the official name which the bilingual project of the community of Madrid receives. It was initially set up in 26 public Infant and Primary schools in 2003–2004 and has since then been extended to 147. Here, the CLIL model can be applied to any subject except Spanish and Mathematics and is taught through English, French, or German. It includes its own specific teacher training programme.

Plan de Fomento del Plurilingüismo: This is the denomination of the ambitious CLIL plan which has been set in motion in Andalusia since 2005. It is based on five pillars and 74 official actions and has been updated by the *Plan Estratégico de Desarrollo de las Lenguas* (PEDLA) in 2017.

As can be observed, these specific CLIL actions are geared at Primary and Secondary level. At these stages, they seek to foster a more communicative, participative, active, and motivating approach to the teaching of languages. Thus, the foreign language teaching situation in Spain is currently under change and CLIL is sparking increased interest and attention in our educational panorama. In our country, this approach is distinctive on two counts. First, it encompasses a diversity of models practically tantamount to the number of regions where it is applied, given the decentralisation of our educational system, which transfers educational powers to each autonomous community. Thus, in our context, the gap between EU policy and CLIL grassroots action is bridged via regional rather than national educational initiatives and no single blueprint exists: 'Spain is a mixture of heterogeneous language situations that lead to different ways of understanding and managing L2 education' (Fernández Fontecha, 2009, p. 4). And second, dual-focussed education has been developed in Spain with both second (co-official) and foreign (other European) languages, and in both bilingual communities where English is the third language taught through CLIL (The Basque Country, Catalonia, Valencia, the Balearic Islands, Galicia) and in monolingual communities conspicuous for their lack of tradition in foreign language teaching (e.g. Extremadura, Castilla-La Mancha, or Andalusia). For these reasons, Spain could well serve as a model for the multiple possibilities offered by the broader CLIL spectrum and thus for other countries seeking to implement it.

3 Structure

In offering new insights into CLIL characterisation, implementation, and research in this multifaceted Spanish context,[2] the present volume comprises ten chapters classified into four main blocks. It begins with a chapter by *Diego Rascón Moreno* and *Antonio Vicente Casas Pedrosa* which describes the linguistic, methodological, and

[2]For a specific focus on contextual, cognitive, and affective variables in CLIL research (e.g. rural–urban context, socioeconomic status, amount and type of exposure), please see the special issue of *Porta Linguarum* (2018) also stemming from the two main research projects presented here. However, this volume offers a completely different perspective by focussing on the current state of CLIL characterisation, implementation, and research into the effects of CLIL on L1, L2, and content learning, from a global (and not variable-based) outlook. In addition, the effects of CLIL on key competences (such as digital competence) are also examined here, and a broader transcontinental

organisational traits of CLIL in both FL and non-linguistic area subjects taught in English with a CLIL methodology vis-à-vis the seven main fields of interest which have been canvassed: EFL use in class, discursive functions, competence development, methodology and types of groupings, materials and resources, coordination and organisation, and evaluation. The results allow the authors to paint a clearer picture of what CLIL looks like at the grassroots level and to thereby make headway in characterising its *implementation.*

María Luisa Pérez Cañado then sets forth an innovative proposal to shed light on the CLIL *characterisation* controversy by examining CLIL in connection to English as a lingua franca (ELF), two acronyms which take centre-stage in the current ELT literature. CLIL is described on a multiplicity of levels, fostering the connection to ELF, unpacking ten main fronts on which both concepts conflate, and showcasing the view that they are in fact compatible, parallel, and can be incorporated in the language classroom concomitantly to the benefit of the stakeholders involved.

The remaining chapters go on to focus on the effects of CLIL *research. Daniel Madrid Fernández*, *Antonio Bueno González*, and *Juan Ráez Padilla* walk the reader through the double-fold pilot process which has been painstakingly followed to draw up a batch of tests to measure the effects of CLIL on English language competence and then provide the actual instruments[3] for their application in further iterations of this type of research. *Esther Nieto Moreno de Diezmas* empirically examines the impact of CLIL on the acquisition of digital competence, while *Elvira Barrios* explores its effects on L1 competence development and *Juan de Dios Martínez Agudo* centres on how it impinges on subject matter learning (specifically, on Science in Primary and Secondary Education). *Macarena Navarro-Pablo* and *María del Mar Gálvez Gómez* finish by determining the effects of CLIL on FL competence, in terms of linguistic components (grammar and vocabulary) and skills (reading, listening, and speaking), respectively.

The book draws to a close with a wrap-up piece by *Cristian Aquino-Sterling*, who, based on all these findings, expounds on future challenges and ways forward for CLIL characterisation, implementation, and research. The ultimate aim is to provide an updated picture of where we stand in the CLIL arena, to identify the chief hurdles which need to be tackled in the very near future, and to signpost possible ways of overcoming them in order to continue advancing smoothly into the next decade of CLIL development.

---

perspective on CLIL is offered in the final chapter. In this sense, this monograph also incorporates the findings of three additional research projects: 2018-1-ES01-KA201-050356, RTI2018-093390-B-I00, and FFI2014-54179-C2-2-P, funded by the European Union, the *Ministerio de Ciencia, Innovación y Universidades*, and the *Ministerio de Economía, Industria y Competitividad*, respectively.

[3]For the remaining qualitative instruments designed, validated, and applied in the projects, please see Pérez Cañado (2016a) (observation protocols), Pérez Cañado (2016b) (questionnaires) and Pérez Cañado (2018) (interview protocols).

References

Cenoz, J., Genesee, F., & Gorter, D. (2013). Critical analysis of CLIL: Taking stock and looking forward. *Applied Linguistics, 35*(3), 243–262. https://doi.org/10.1093/applin/amt011.

Coyle, D. (2010). Foreword. In D. Lasagabaster & Y. Ruiz de Zarobe (Eds.), *CLIL in Spain: Implementation, results and teacher training* (pp. vii–viii). Newcastle upon Tyne: Cambridge Scholars Publishing.

Dalton-Puffer, C., Llinares, A., Lorenzo, F., & Nikula, T. (2014). You can stand under my umbrella. Immersion, CLIL and bilingual education. A response to Cenoz, Genesee and Gorter (2013). *Applied Linguistics, 35*(2), 213–218. https://doi.org/10.1093/applin/amu010.

Fernández Fontecha, A. (2009). Spanish CLIL: Research and official actions. In Y. Ruiz de Zarobe & R. M. Jiménez Catalán (Eds.), *Content and language integrated learning: Evidence from research in Europe* (pp. 3–21). Bristol: Multilingual Matters.

Fernández-Sanjurjo, J., Fernández-Costales, A., & Arias Blanco, J. M. (2019). Analysing students' content-learning in science in CLIL vs. non-CLIL programmes: Empirical evidence from Spain. *International Journal of Bilingual Education and Bilingualism, 22*(6), 661–674.

Langé, G. (2007). Postscript to CLIL 2006 and future action. In D. Marsh & D. Wolff (Eds.), *Diverse contexts, converging goals: CLIL in Europe* (pp. 350–358). Frankfurt am Main: Peter Lang.

Lasagabaster, D. (2008). Foreign language competence in content and language integrated courses. *The Open Applied Linguistics Journal, 1,* 31–42.

Lasagabaster, D., & Ruiz de Zarobe, Y. (2010). Ways forward in CLIL: Provision issues and future planning. In D. Lasagabaster & Y. Ruiz de Zarobe (Eds.), *CLIL in Spain: Implementation, results and teacher training* (pp. 278–295). Newcastle upon Tyne: Cambridge Scholars Publishing.

Lasagabaster, D., & Sierra, J. M. (2009). Language attitudes in CLIL and traditional EFL classes. *International CLIL Research Journal, 1*(2), 4–17.

Lasagabaster, D., & Sierra, J. M. (2010). Immersion and CLIL in English: More differences than similarities. *ELT Journal, 64*(4), 367–375.

Lyster, R. (2007). *Learning and teaching languages through content: A counterbalanced approach.* Amsterdam: John Benjamins.

Madrid , D. (2006). Bilingual and plurilingual education in the European and Andalusian context. *International Journal of Learning, 12*(4), 177–185.

Paran, A. (2013). Content and language integrated learning: Panacea or policy borrowing myth? *Applied Linguistics Review, 4*(2), 317–342.

Pérez Cañado, M. L. (2011). The effects of CLIL within the APPP: Lessons learned and ways forward. In R. Crespo & M. A. García (Eds.), *Studies in honour of Ángeles Linde López* (pp. 13–30). Granada: Universidad de Granada.

Pérez Cañado, M. L. (2012). CLIL research in Europe: Past, present, and future. *International Journal of Bilingual Education and Bilingualism, 15*(3), 315–341.

Pérez Cañado, M. 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.

Pérez Cañado, M. L. (2016b). Evaluating CLIL programmes: Instrument design and validation. *Pulso. Revista De Educación, 39,* 79–112.

Pérez Cañado, M. L. (2017). Stopping the "pendulum effect" in CLIL research: Finding the balance between Pollyanna and Scrooge. *Applied Linguistics Review, 8*(1), 79–99. https://doi.org/10.1515/applirev-2016-2001.

Pérez Cañado, M. L. (2018). The evolution of bilingual education in monolingual settings: An Andalusian case study. In P. Romanowski & M. Jedynak (Eds.), *Current research in bilingualism and bilingual education* (pp. 207–241). Amsterdam: Springer.

Pérez-Vidal, C. (2007). The need for focus on form (FoF) in content and language integrated approaches: An exploratory study. *RESLA Vol. Extra, 1,* 39–54.

Ruiz de Zarobe, Y. (2008). CLIL and foreign language learning: A longitudinal study in the Basque Country. *International CLIL Research Journal, 1*(1), 60–73.

Ruiz de Zarobe, Y., & Lasagabaster, D. (2010). CLIL in a bilingual community: The Basque autonomous community. In D. Lasagabaster & Y. Ruiz de (Eds.), *CLIL in Spain: Implementation, results and teacher training* (pp. 12–29). Newcastle upon Tyne: Cambridge Scholars Publishing.
Swan, M. (1985). A critical look at the communicative approach (2). *ELT Journal, 39*(2), 76–87.

Characterising Representative CLIL Practices: An Andalusian Case Study

Diego Rascón Moreno and Antonio Vicente Casas Pedrosa

Abstract The characterisation of Content and Language Integrated Learning (CLIL) has recently come to the forefront and acquired a new significance. Both their conceptualisation and pedagogical implementation have of late started to be questioned and are considered as excessively vague and ambiguous, since CLIL is held to encompass too broad an array of possible programme alternatives, thus making its exact limits very difficult to pin down. Thus, we need to characterise representative CLIL practices and to know exactly what it looks like in practice. This chapter reports on the outcomes of two governmentally funded R&D projects (FFI2012-32221 and P12-HUM-2348), within which an observation protocol has been designed, validated, and applied in 53 public, private, and charter schools in 12 provinces belonging to Andalusia, the Canary Islands, and Extremadura. English as a Foreign Language and Non-Linguistic Area subjects taught in English with a CLIL methodology have been observed and the linguistic, methodological, and organisational traits of CLIL are here described with a representative sample in the provinces of Jaén and Granada vis-à-vis the seven main fields of interest which have been canvassed: foreign language use in class, discursive functions, competence development, methodology and types of groupings, materials and resources, coordination and organisation, and evaluation. The results allow us to paint a clearer picture of what CLIL looks like at the grassroots level and to thereby make headway in characterising representative pedagogical CLIL practices which will hopefully contribute to honing and fine-tuning its characterisation.

D. Rascón Moreno (✉) · A. V. Casas Pedrosa
University of Jaén, Jaén, Spain
e-mail: drascon@ujaen.es

A. V. Casas Pedrosa
e-mail: avcasas@ujaen.es

M. L. Pérez Cañado (eds.), *Content and Language Integrated Learning in Monolingual Settings*, Multilingual Education 38,
https://doi.org/10.1007/978-3-030-68329-0_2

1 Introduction

We stand in dire need of characterising representative Content and Language Integrated Learning (CLIL) practices and of knowing exactly what this approach looks like at the grassroots level. In the last few years, CLIL has been facing challenges posed by controversies regarding its characterisation and implementation. These have been identified by Pérez Cañado, who has also provided specific ways to redress them (2016 and 2017).

On the one hand, its conceptualisation has been questioned. For example, it is confused with its predecessors (Paran, 2013, p. 319): Canadian immersion, North American bilingual education, and European international schools (Pérez Cañado, 2012, p. 316). In addition, it is considered to have an 'ill-defined nature' (Paran, 2013, p. 318). This has resulted in distilling the core features that make it a foreign language teaching trend in itself, but these attempts to isolate CLIL from other types of bilingual programmes have been criticised too (Cenoz, et al. 2013). Pérez Cañado offers a way out of this terminological conundrum (2016, p. 18) and calls for conducting 'extensive classroom observation in diverse CLIL contexts' to address this controversy over characterisation.

On the other hand, the implementation of CLIL is considered to be vague and ambiguous as well (Bruton, 2011a; Cenoz, et al. 2013), mainly due to its lack of a unique pedagogy and the 'plethora of models or variants which can be identified within it' (Pérez Cañado, 2016, p. 14). The controversy here lies in the argument that there should not be a 'one-size-fits-all model' (Smit, 2007 cited in Pérez Cañado, 2016, p. 15), and that there is a flexible combination of factors conducive to a broad array of CLIL programmes, up to 216 according to Coyle (2007), which allows it to accommodate the linguistic diversity in Europe. Here again, characterising pedagogical practices that are representative is deemed necessary, 'irrespective of the camp with which one sides' (Pérez Cañado, 2016, p. 15).

This chapter stems from two research projects funded by the Spanish Ministry of Economy and Competitiveness and the Andalusian government. One of the various lines of action stemming from them is the design, validation, and application of observation protocols in 53 public, private, and charter schools in 12 provinces belonging to the three monolingual communities in Spain with the least tradition in bilingual education, in order to precisely counter the aforementioned conceptual and pedagogical vagueness.

In particular, this chapter describes a representative sample of 44 observations in the Andalusian provinces of Jaén and Granada as regards the following fields of interest which have been canvassed: foreign language use in class, discursive functions, competence development, methodology and types of groupings, materials and resources, coordination and organisation, and evaluation. English as a Foreign Language and Non-Linguistic Area subjects taught through English with a CLIL methodology (Science to a greater extent, and Mathematics and Ethics to a lesser extent) were observed in search of linguistic, methodological, and organisational

traits of CLIL which allow us to paint a clearer picture of what it looks like at the grassroots level, and, ultimately, to hone and fine-tune its characterisation. The English as a Foreign Language classes were also studied, as their teachers play an essential role in the language teaching context and must coordinate with non-linguistic area teachers and teaching assistants in the common interests of providing a successful bilingual experience at school.

2 Prior Research

Doubts about the conceptualisation and implementation of Content and Language Integrated Learning (CLIL) have of late begun to be raised as both features are considered to be excessively vague and ambiguous. With regard to the former, some authors have criticised its ambiguity, vagueness, ill-defined nature, and lack of exact limits or terminological clarity that cause it to be confused with Content-Based Instruction (CBI) or Immersion Education (Cenoz, et al. 2013; Paran, 2013). Probably influenced by this criticism, interest in highlighting the core features of this approach has increased in the last few years. Authors like Dalton-Puffer et al. (2014), Lasagabaster and Sierra (2009), Pérez Cañado (2012), and Pérez-Vidal (2013) (cited in Pérez Cañado, 2016, p. 12) have distilled some traits of CLIL in relation to its language of instruction, the languages taught through it, the methodology used, the language level targeted, the linguistic command of teachers, the amount of exposure to the second or foreign language, or the types of materials employed.

Nevertheless, this tendency to specify the features of CLIL to make it a foreign language teaching approach in itself has in turn been called into question by authors such as Cenoz (2015), Cenoz and Ruiz de Zarobe (2015), Cenoz, et al. (2013), Hüttner and Smit (2014) and Somers and Surmont (2011) (cited in Pérez Cañado, 2016, p. 12) on the claim that there are more similarities than differences between CLIL and immersion or CBI, and that a reductionist or isolationist view of this approach is negative for teachers and researchers.

We fully agree with those experts who believe that the solution to this conceptual crossroads is to adopt an integrative and encompassing view of CLIL, i.e. to regard it as an umbrella term that can include various forms. As Cenoz, et al. (2013) put it, 'rather than insisting on the uniqueness of CLIL, efforts might be better spent establishing a taxonomy of different common forms of CLIL/CBI so as to circumscribe the diverse contexts in which CLIL is found' (p. 258). For Pérez Cañado (2016), 'the onus is now on recognising the diversity of formats which can be subsumed within CLIL and on ensuring that the results and effects of all types of multilingual programmes (be they CLIL, CBI, or immersion) are shared so that the pedagogical and research community can benefit from them' (p. 14).

In line with the foregoing, Pérez Cañado (2012) suggests several strategies to make headway in this area: reflection, classroom observation, and communication. To conduct the second, a protocol has been designed and validated (via the expert

ratings approach) as part of the actions derived from two governmentally funded research projects in which the authors of this chapter participate.

As far as implementation is concerned, CLIL has been criticised for not being unique pedagogically. The aforementioned conceptual ambiguity has reached 'on-the-ground practice' (Pérez Cañado, 2016, p. 14). A wide variety of models can be subsumed within CLIL, dependent on many parameters such as the following: operating factors, scale of the CLIL programme, environmental parameters, population segments, age groups, monolingual-multilingual settings, types of teachers involved, learner assessment, type and amount of target language usage, language taught, degree, and depth of content, first language (L1)/second language (L2) balance, involvement of subject specialists, and the extent to which CLIL is present in the curriculum (see Pérez Cañado, 2016, p. 14).

This amount of possible variants is considered by many scholars to be detrimental to a coherent evolution of Content and Language Integrated Learning (CLIL), pedagogically speaking. For Cenoz, et al. (2013), the diversity of learning contexts that can be classified as CLIL, and which they mention are 'ill-defined', is an obstacle to identify 'the programmatic, instructional, and student-related properties that are specific, and perhaps unique, to CLIL' (p. 254), although in our view, they do not need to be unique to it.

On this occasion, the swinging of the pendulum has to do with the positive opinion that a different group of authors has about the nature of CLIL. For Lasagabaster (2008) and Pérez Cañado (2017) not only does this flexibility not hamper its development but it also helps to accommodate the diverse linguistic landscape in Europe. It stretches 'to meet all needs' according to Dickey (2004, p. 13), and avoids the 'one-size-fits-all model' mentioned by Smit (2007) and considered a failure by Lorenzo, et al. (2011, p. 454) (cited in Pérez Cañado, 2016, p. 15).

Again here, it is important to describe the 'representative pedagogical practices' of CLIL that Bruton (2011b, p. 240) claimed for, regardless of the aforementioned camp 'with which one sides' (see Pérez Cañado, 2016, p. 15). It is essential to know exactly what this approach looks like in practice since 'its linguistic, methodological, and organisational traits need to be further honed, sharpened, and fine-tuned in line with the demands of the diverse contexts where it is being applied' (Pérez Cañado, 2016, p. 15).

Some practices in the same Southern Spanish region of our sampled schools have already been described, albeit not under the classroom observation conditions which will be explained in Sect. 4. Ruiz Gómez (2015) expounds on a project carried out by the Educational Administration of Andalusia at the beginning of this decade. Among its objectives was 'to agree on a full-fledged methodological model based on the CLIL approach' (Ruiz Gómez, 2015, p. 16). The results obtained from different bilingual schools are highly heterogeneous due to their various interpretations of CLIL, 'some of which were clearly not effective' (Ruiz Gómez, 2015, p. 16).

Ruiz Gómez (2015) reports on the existence of a diversity of approaches, which he framed within two general groups. First, he describes situations in which language exploration was not given special attention, but there were repetitive language

showers through which the second language (L2) acquisition took place. Thus, the active use of the language to transmit the subject contents was the only way to fulfil the L2 teaching objectives. For Ruiz Gómez, this teaching behaviour is very appropriate, especially when no extra time in the curriculum is given to the non-linguistic areas despite the difficulty of covering them through the target language. With regard to materials, we are informed that either authentic ones are borrowed from any of the countries speaking English, German, or French, or the staff use Spanish ones that they have translated ad hoc.

Within the second group, Ruiz Gómez (2015) includes those cases in which Content and Language Integrated Learning (CLIL) was regarded as an approach that consciously integrates target language and subject content objectives. This interpretation generated a plethora of versions depending on the management of language and content in the teaching agenda. On the one hand, some versions were successful due to the elaboration of an integrated curriculum including balanced language and non-language objectives, as well as carefully selected content and materials, under the consideration of effective second language acquisition processes. Practitioners can also be given credit for good CLIL practices on the following grounds:

> (1) ability to successfully blend language and non-language aims; (2) a medium-high mastery of the L2 (usually B2 or higher) that let them manipulate the language with ease; (3) readiness to accept new educational proposals; and (4) working in permanent cooperation with the L2 teacher. (Ruiz Gómez, 2015, p. 17)

Sadly, on the other hand, the project described by this author found bilingual settings in Andalusia that presented clear flaws in the implementation of the programme, mainly because of 'the non-existence of a reliable theoretical reference model' that could ease the integration of the two types of learning, which consequently triggered ineffective methodological CLIL practices basically derived from intuition. Ruiz Gómez (2015) cites some examples such as didactic proposals with L2 objectives reduced to 'the presentation of a group of new lexical items', the use of allegedly adapted materials that generally turned out to be very difficult for students, 'an unsystematic and even neglected consideration of skills', and the exploitation of materials 'which showed a deficient representation of discourse models and functions' (pp. 18–19).

The undertaking of empirically solid classroom observation like the one presented below will enable us to identify cases of ineffective methodological practices of the CLIL programme, such as those found by Ruiz Gómez (2015). Thanks to this detection, it will be possible to develop remedial actions against practices deviating from the challenging path of CLIL, a path which we cannot afford to not follow, in light of its results and the many benefits it brings to students.

Pérez Cañado and Lancaster's research (2017) should also be mentioned, as it shows the outcomes of an intervention conducted in a school in the province of Jaén. It is particularly worth mentioning here because it can guide CLIL pedagogical practice vis-à-vis oral communication. The authors reveal the positive results of a longitudinal case study in terms of all speaking skills in the medium term, and of oral

production skills also in the long term. This research is focused on oral comprehension and production. It does not include the written skills.

With respect to the methodology of the CLIL classes, the authors mention that the sessions of Physical Education and Ethics given to the experimental group were essentially communicative, involving some research, discussions, and oral presentations on the part of students. As for the English as a Foreign Language (EFL) lessons, the bilingual group received them in a classroom equipped with technology and with access to a bilingual library, following again a communicative approach to learning, where the time of exposure to English was as high as 80%, and student-centred, independent, task-based, and collaborative learning prevailed.

3 Objectives

This chapter is aimed at characterising representative Content and Language Integrated Learning (CLIL) practices in order to know what it looks like at the grassroots level, and at determining the existence of within-group differences according to a series of variables. Linguistic, methodological, and organisational traits of the approach are here described, with a sample in the Spanish provinces of Granada and Jaén, through the application of a previously designed and validated observation protocol.

This general objective can be divided into two key metaconcerns which prove essential for this research project and consist of several component corollaries:

Metaconcern 1

(a) To describe traits of CLIL vis-à-vis foreign language use in class.
(b) To describe traits of CLIL in relation to discursive functions.
(c) To describe traits of CLIL as regards competence development.
(d) To describe traits of CLIL concerning methodology and types of groupings.
(e) To describe traits of CLIL relating to materials and resources.
(f) To describe traits of CLIL pertaining to coordination and organisation.
(g) To describe traits of CLIL regarding evaluation.

Metaconcern 2

In turn, the second main objective intends to shed light on the modulating (differential) effect exerted on the sessions observed—in terms of the abovementioned fields of interest– by the intervening variables of school, province, level, subject, type of teacher, bilingual coordination, age, gender, nationality, administrative situation, overall teaching experience, and teaching experience in a bilingual school.

4 Method

4.1 Research Design

This study reports on the outcomes of two governmentally funded research and development (R&D) projects (see Acknowledgements) aimed at evaluating CLIL programmes in three of the monolingual communities in Spain with the least tradition in bilingual education: Andalusia, the Canary Islands, and Extremadura. Although the project has consisted of different stages focusing on several aspects of CLIL, this chapter is related to the direct behaviour observation of both content subjects (the ones taught through the foreign language) and English as a Foreign Language (EFL) sessions taught to both Primary (6th grade) and Secondary (4th grade) Education students in the Andalusian provinces of Granada and Jaén.

This investigation is an instance of primary research, and within it, of survey research, as it includes questionnaires (Brown, 2001). According to this author, survey research is characterised by being based on data, employing interviews and questionnaires, and being mid-way between qualitative and statistical research since it can use both types of techniques. More specifically, an observation protocol which had been previously designed and validated has been applied.[1]

Another feature of this study is what Denzin (1970) calls 'multiple triangulation', since the following types of triangulation have been employed:

- Data triangulation, given that, apart from the protocols themselves, after the direct behaviour observation, brief interviews were held with the educators (and that involves, within them, not only foreign language (FL) and non-linguistic area (NLA) teachers, but also teacher assistants (TAs).
- Investigator triangulation, since, for each province, two members of the research team have observed each of the lessons under analysis and have completed the observation protocols.
- Location triangulation, due to the fact that data has been collected from both Primary and Secondary schools in the provinces of Granada and Jaén.

This study is inserted within the qualitative side of the investigation. It seeks to examine the teachers' and students' behaviour in both content and EFL sessions through direct observation by members of the research team and to determine the existence of within-group differences in their perceptions. This analysis will allow us to clearly describe how CLIL is implemented at the grassroots level in terms of the seven main fields of interest which have been canvassed.

[1]The observation protocol can be obtained from Pérez Cañado (2016, pp. 27–31).

4.2 Sample

Although the two above-mentioned projects have involved 53 public, private, and charter schools in the 12 provinces of Andalusia, the Canary Islands, and Extremadura, this chapter focuses on the schools in the provinces of Granada (five) and Jaén (six). Whereas four of these 11 schools are Primary, there are seven Secondary schools (Fig. 1). More specifically, this chapter only focuses on the EFL and content teachers whose sessions were observed by two members of the research group in charge of the Andalusian provinces of Granada and Jaén. Thus the number of teachers involved is 22.

Moreover, in terms of gender, nearly two-thirds are female teachers (64.9%) whereas 35.1% are males (Fig. 2). All the educators are of Spanish nationality (97.1%) except for an American TA (2.9%). With regard to their age, two-thirds are 45 or younger (68%) and 32% are older than 45. Half the sessions observed

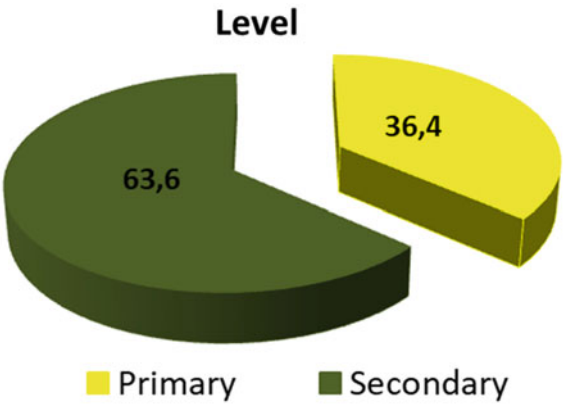


Fig. 1 Breakdown of the overall sample in terms of level

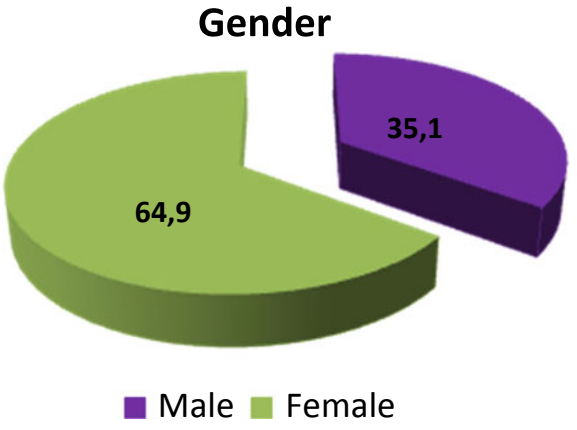


Fig. 2 Breakdown of the overall sample in terms of gender

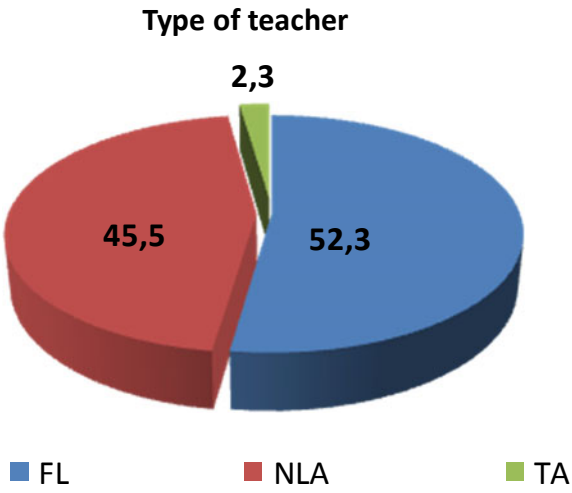


Fig. 3 Breakdown of the overall sample in terms of type of teacher

were taught by EFL teachers, whereas the rest were run by Science teachers (8), as well as an Ethics teacher and a Mathematics teacher. Another session was run by an FL teacher together with the TA (Fig. 3). Furthermore, nearly half of the educators (48.4%) are bilingual coordinators (Fig. 4).

In terms of experience, nine teachers in our sample have one to ten years of teaching experience, followed by eight teachers who have between 21 and 30 years of teaching experience, four teachers who have been teaching for less than a year, and two teachers who have more than 30 years of experience. In addition, the vast majority have been involved in bilingual education for either between six and ten years (seven teachers) or for more than 15 years (six teachers). They outnumber the rest and this suggests that they have acquired a high level of experience with respect to bilingualism, in which they have been involved from the outset of the Andalusian Plurilingualism Promotion Plan (APPP) (*Plan de fomento del plurilingüismo en Andalucía*, Junta de Andalucía, 2005), or, on some occasions, even before its official implementation. The rest have done so for between 11 and 15 years (two teachers),

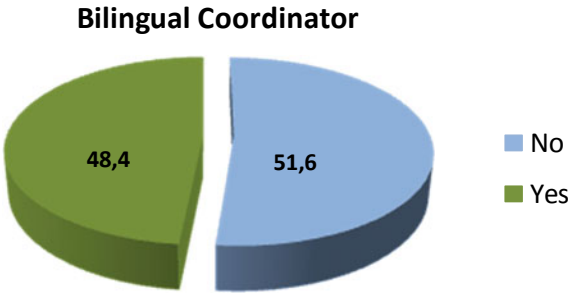


Fig. 4 Breakdown of the overall sample in terms of bilingual coordinator

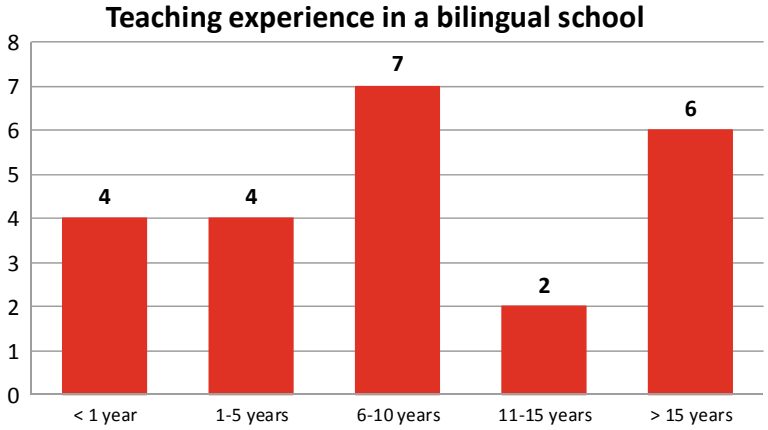


Fig. 5 Breakdown of the overall sample in terms of their teaching experience in a bilingual school

between one and five years (four teachers), and for less than a year (four teachers) (Fig. 5).

Most of these teachers are civil servants with a permanent destination (44.4%), whereas 22.2% are interns. The rest (33.3%) consider themselves to be in a different administrative situation from the above-mentioned ones.

4.3 Variables

A series of identification (subject) variables have been considered. They are specified below:

- School
- Province
- Level (Primary/Secondary)
- Subject
- Type of teacher (NLA/FL/TA)
- Bilingual coordinator
- Age
- Gender
- Nationality
- Administrative situation (civil servant with permanent destination/civil servant with provisional destination/intern/other)
- Overall teaching experience
- Teaching experience in a bilingual school.[2]

[2]Although all these variables were included in the observation protocol, only some of them have been considered for this study. The ones left aside are school, province, subject, nationality, and

4.4 Instrument

Since the study focuses on direct behaviour observation, it has employed an observation protocol. It was designed, validated, and applied not only in the schools in the provinces of Granada and Jaén involved in this research, but also in the rest of schools participating in the national projects within which it is subsumed. This protocol first includes 11 demographic or background questions, in line with Patton's (1987) question types. These provide mainly biographical information from the teachers in charge of the sessions, which has allowed us to control the identification variables mentioned above.

Then the protocol is divided into the seven main fields of interest which have been canvassed: foreign language use in class (1–8, eight items), discursive functions (9–19, eleven items), competence development (20–30, eleven items), methodology and types of groupings (31–48, eighteen items), materials and resources (49–57, nine items), coordination and organisation (58–65, eight items), and evaluation (66–71, six items). The number of items within each section has been included in brackets and the tool includes some space at the end of each section for the researcher to comment on any other aspects considered interesting for the project. This could be done during the session itself and/or afterwards, in the brief interviews held between the researchers and the educators.

The former type of questions are fill-in and short-answer ones (following Brown's [2001] classification) and the latter, Likert-scale ones from 1 to 4 in order to avoid the central tendency error.

4.5 Statistical Methodology

The data collected has been statistically analysed by means of the Statistical Package for the Social Sciences (SPSS) programme, version 21.0. For *Metaconcern* 1 (objectives a-g), descriptive statistics has been used. Therefore, both central tendency (mean, median, and mode) and dispersion measures (range, low-high, and standard deviation) have been calculated. In turn, for *Metaconcern* 2 several statistical tests have been employed, namely: the analysis of variance (ANOVA), t-test, and the Mann-Whitney *U* test, with the aim of finding statistically significant differences exerted on the sessions by the aforementioned identification variables. Grounded Theory analysis (Glaser and Strauss, 1967) has been used as the framework of reference for data coding, memoing, and drawing of conclusions, in order to categorise, synthesise, and identify emerging patterns in the open-response data.

---

administrative situation. Thus, for example, since all educators were Spanish except for one, as explained in Sect. 4.2, the nationality variable has not been taken into account.

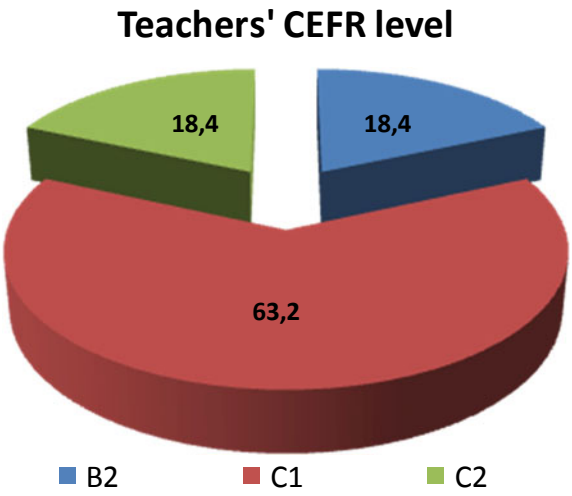


Fig. 6 Breakdown of the overall sample in terms of the teachers' CEFR level

5 Results and Discussion

This section is aimed at analysing the results drawn from this project bearing in mind the above-mentioned component corollaries into which the two metaconcerns can be divided, as well as the differential effect produced by the twelve intervening variables on the sessions observed. Regarding *Metaconcern 1* (objectives a-g), this study has allowed us to observe the teachers' and students' direct behaviour in the classroom vis-á-vis the above-mentioned seven main fields of interest.

First of all, it should be mentioned that, bearing in mind the use of a 1–4 Likert scale (1 meaning 'Very much', 2, 'Considerably', 3, 'To some extent', and 4, 'Not at all'), low results are positive, whereas high ones are negative. Since the average marks in most cases are between 1 and 2, they are quite positive. Leaving the first block aside, 24 out of 63 items (more than a third) show positive results, generally speaking.[3]

With regard to the use of English in class at both Primary and Secondary Education levels, the results show that more than half of the teachers have a C1 level (63.2%) according to the Council of Europe's (2001) *Common European Framework of Reference for Languages* (CEFR), whereas nearly a fifth of them have at least a B2 (18.4%) and the rest even a C2 (18.4%) (Fig. 6). Nearly three fourths (71.4%) of teachers use English in class most of the time (between 75 and 100%) (Fig. 7). In line with this idea, more than half of the teachers (56.1%) do not translate from Spanish into English at all and nearly a fifth (19.5%) do so to some extent.

[3]The reason why the protocol number of items mentioned here has been reduced from 71 to 63 is because, contrary to what happens in the rest of blocks, in the first section 'lower' does not mean 'more positive', and 'higher' does not stand for 'more negative', but the other way round, except for the fifth item: 'The students' linguistic competence is appropriate for the year they are studying' (please, see Pérez Cañado, 2016, pp. 27–31).

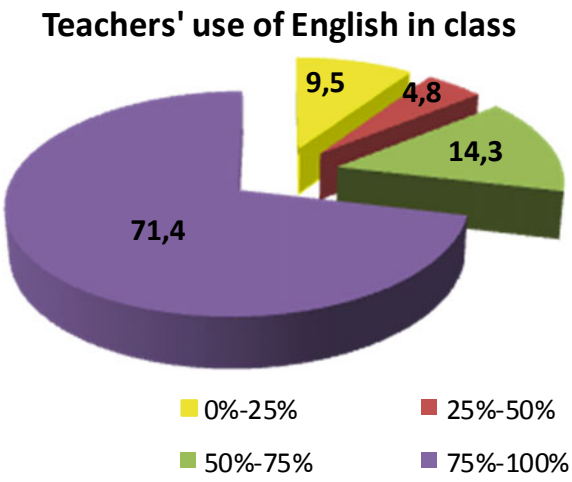


Fig. 7 Breakdown of the overall sample in terms of the teachers' use of English in class

Very similar percentages (although slightly lower) can also be found concerning the teachers' use of code-switching (51.2 and 17.1%, respectively).

Thus, for instance, teachers may, in some cases, just summarise the main ideas in Spanish from time to time or translate very difficult words into Spanish for the sake of comprehension. On some occasions, interestingly, EFL sessions are not communicative and are rather based on drilling, repetition, or translation, and English is even used to a lesser extent than in content classes. Thus, language teaching sometimes seems not to be changing at the same pace as content teaching.

Moreover, the students' level of English is mainly very much (43.9%) or considerably (43.9%) adequate for their year (either 6th grade of Primary Education or 4th grade of Compulsory Secondary Education). Another positive aspect is the fact that students tend to use English in class between 75 and 100% (42.9%) or between 50 and 75% (28.6%) of the time. In any case, students tend to use both Spanish and code-switching as well as to translate from Spanish into English more often than teachers. It can be stated that some students use English for certain activities (e.g. to read aloud, to correct activities when called on, to do role-plays, for question-answer exchanges, for oral presentations, etc.), but speak Spanish for some others. Most teachers encourage their students to use English all the time either explicitly (e.g. S: '*Yo tengo un Kahoot*'; T: 'Sorry?'; S: 'I have a Kahoot') or implicitly. Thus, for example, even though students might ask certain questions in Spanish, teachers answer in English. In addition, for the sake of fluency, sometimes mistakes are not corrected (e.g. *'I am agree') as long as they do not involve a communication breakdown.

The students' level of English is very heterogeneous in some groups. In those cases this pitfall may be due to not having participated in the CLIL programme for a similar amount of time. In fact, on some occasions different ability levels could be detected. Thus, for instance, in the case of the above-mentioned oral presentations,

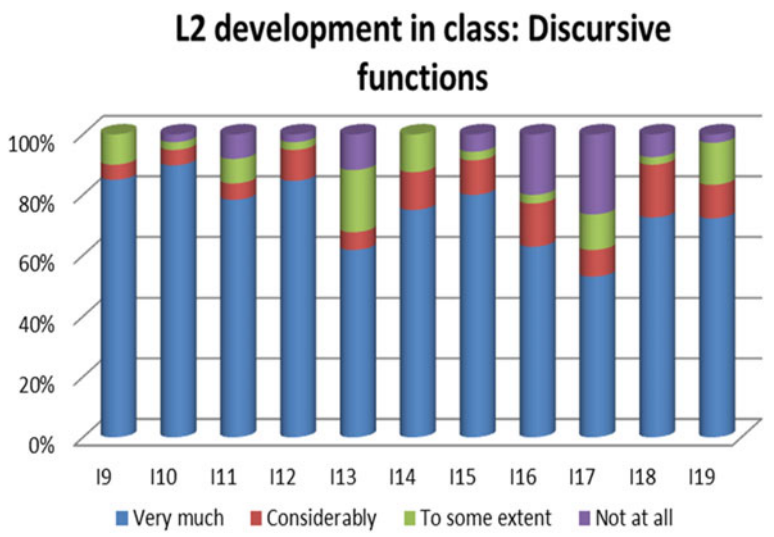


Fig. 8 L2 development in class: Discursive functions

low level students tend to use memorisation, whereas higher level ones tend to rely more on improvisation and speak for much longer in English.

Vis-à-vis discursive functions (Fig. 8), the results are quite positive concerning both teachers and students since all the items except for one (10 out of 11) have a mean between 1 ('Very much') and 2 ('Considerably'), the lowest item being 1.18 (item 10: 'English is used in class to introduce the topic'), and the highest, 2.12 (item 17: 'English is used in class to organise the different types of groupings'). Thus, it can be concluded that English is very often used in class for all discursive functions.

Moreover, generally speaking, the impressions concerning competence development in class are positive, although not as much as in the previous block. The competences which are more often favoured in class include oral comprehension (1.30), oral expression (1.49), oral communicative interaction (listening and speaking) (1.54), written comprehension (1.56), and learning autonomy (1.92).[4] Quite on the contrary, the item with the lowest average is critical capacity (2.62). As for the rest of the competences, the very nature of some subjects makes it easier for teachers to include certain sociocultural aspects (item 30) in the sessions (for example, History or Ethics sessions versus Mathematics).

In addition, the methodology and types of groupings in these sessions are less positively valued than the previous two blocks and only one-third of the items (six out of 18 items) show an average between 1 and 2, including linguistic scaffolding (definitions, examples, paraphrases, repetitions, synonyms, antonyms, etc.), the lexical

[4]Further information about the effects of CLIL on FL grammar and vocabulary as well as on FL skills can be found in the seventh and eighth chapters in this volume (by Navarro-Pablo and Gálvez Gómez, respectively).

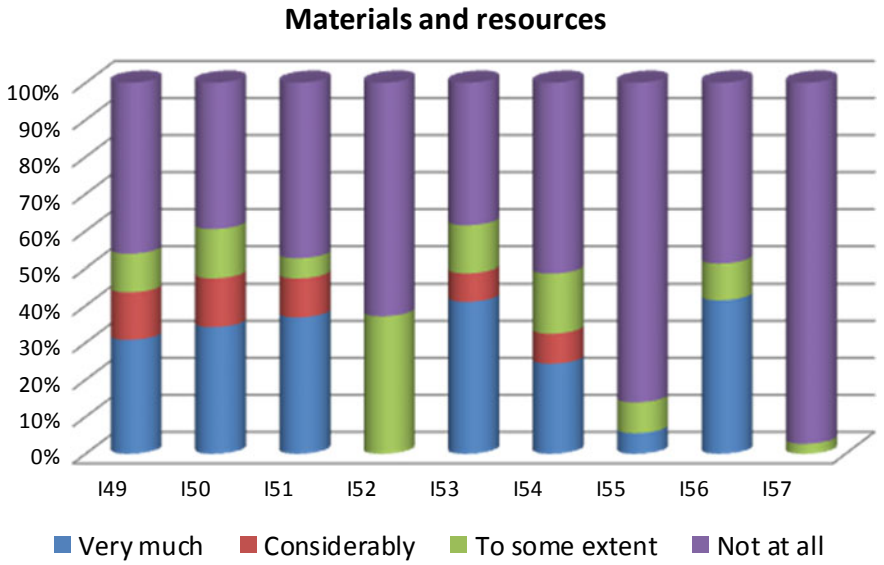


Fig. 9 Materials and resources

dimension, individual work, the audiolingual method, lockstep grouping, and task-based learning. It was also observed that some teachers favour self-correction and even induce students to correct themselves. Fortunately, it can also be added that the item with the highest average is the use of the grammar-translation method in class (3.49) and this means that it is no longer popular either in content classes or in EFL sessions. With regard to groupings, in some sessions teachers use different types of groupings and that does not necessarily involve moving the pieces of furniture in the classroom.

Furthermore, none of the items in the materials and resources section is valued between 1 and 2 (Fig. 9). In fact, the lowest item is 53 ('Multimedia software is used in class') and the highest is 57 ('Computer-mediated communication is used in class') with values of 2.49 and 3.97, respectively. Only one session was observed, but teachers were given the chance to comment on any specific aspect after the direct behaviour observation during the brief interviews held between observers and educators. Thus, for example, it could be checked that some materials and resources which were not introduced in the protocol are nonetheless employed. Such is the case of the blackboard (mainly for the correction of exercises and grammar activities), notebooks (which play a relevant role in some schools, especially in the case of Science classes), and realia (e.g. actual rocks in Science sessions). Nowadays, other gadgets are becoming increasingly popular in education and digital books, iPads,

laptops, and tablets are used in some classes since they allow both teachers and students to get access to other resources such as Kahoot.[5]

It should also be mentioned that in some content sessions the textbook is in Spanish and some NLA teachers highlight the fact that there are no textbooks in English that can be used for their lessons. This, without a doubt, makes lesson planning more time-consuming since they have to take activities from many different sources (e.g. the British Council website) or even create their own teaching materials (e.g. a glossary created by a group of Mathematics teachers).

With regard to coordination and organisation, similarly to the previous block, in this section only 1 item (out of 8) shows a value between 1 and 2 (1.93), therefore being very positive: 'Linguistic learning is supported in content classes'. The item with the highest score (3.33) reads 'The connection among L1, L2, and L3 is underlined'.

Finally, as for evaluation, just a couple of items show values between 1 and 2 ('Formative assessment is practiced' and 'Summative assessment is practiced'), with the highest being 'Assessment takes diversity into account' (2.88). Sometimes rubrics are employed to assess certain activities like oral presentations. Some of them include up to ten aspects to be taken into account. Thus, in one of the schools the rubric focused on presentation, voice, body language, memory, organisation and opinion, interest and originality, use of English, fluency and pronunciation, questions, and others. This is connected with the so-called 'co-evaluation' since on some occasions teachers handed out rubrics for the students to assess their classmates' presentations.

In the brief interviews, some teachers let observers know that certain sessions were slightly different from usual for a number of reasons (e.g. the observation took place before the final exam and some classes were review sessions for students to ask about doubts or any other questions they had).

Turning now to *Metaconcern 2*, statistically significant differences can be detected in terms of the above-mentioned identification variables. They have been found in every single variable, but the one with the highest number is teaching experience in bilingual schools (28 out of 71 items). Quite on the contrary, other variables such as gender, level, type of teacher, bilingual coordinator, and age present statistically significant differences in just 17 or fewer items (2, 8, 8, 16, and 17, respectively).

The different teachers have been classified into two groups according to their experience in bilingual schools (five years or less and more than five years) and the block with the highest number of differences is the one related to discursive functions (Fig. 10). In this respect, they have been identified in seven out of the 11 items, their development of L2 in class being closer only regarding giving instructions, introducing the topic, correcting tasks, and organising the different types of groupings (items 9, 10, 15, and 17, respectively). In the rest of cases, the statistically significant differences in the seven items have to do with teachers with more than five years of teaching experience in bilingual schools using English for the remaining discursive functions more often than teachers with less experience. This could be due to the

[5]Please check the fourth chapter in this volume (by Nieto Moreno de Diezmas) for more details regarding the acquisition of digital competencies in CLIL settings.

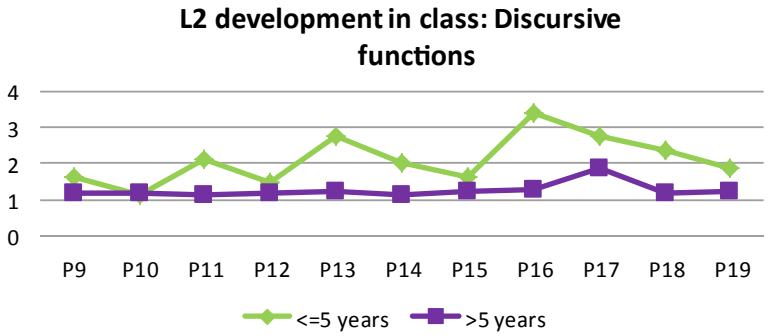


Fig. 10 L2 development in class: Discursive functions according to teaching experience in bilingual schools

fact that, after so many years of experience in the implementation of the APPP, they have become aware of the fact that increasing the students' exposure to the foreign language will improve their proficiency in English.

In turn, the perceptions regarding materials and resources, as well as evaluation, are quite similar, but such is not the case for the other blocks. Teachers with more experience in bilingual schools tend to develop competences in class more often than less experienced ones and they also use a wider range of methodologies and types of groupings in class. With regard to coordination and organisation, it is the only block (together with the two cases of this same section but in relation to the variables of bilingual coordination and overall teaching experience) considering all the blocks and all the variables (that is to say, three out of 49 options) in which a group of teachers shows more positive values (or as positive in one case) than the other in all the items (Figs. 11, 12, 13). In this case, according to the observers, teachers with more experience in bilingual schools show a higher degree of coordination with EFL teachers, TAs, and NLA teachers, include linguistic learning in content sessions more often, and collaborate more on materials preparation and design.

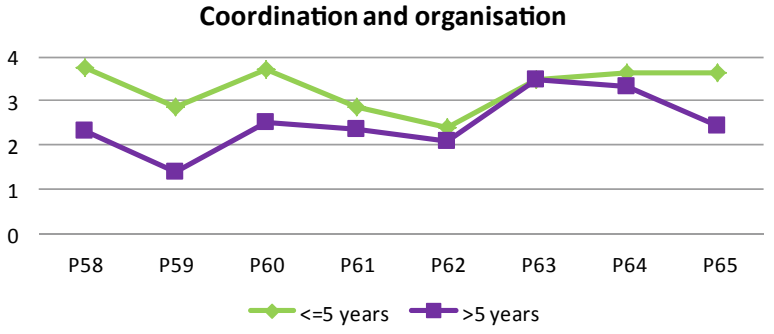


Fig. 11 Coordination and organisation according to teaching experience in bilingual schools

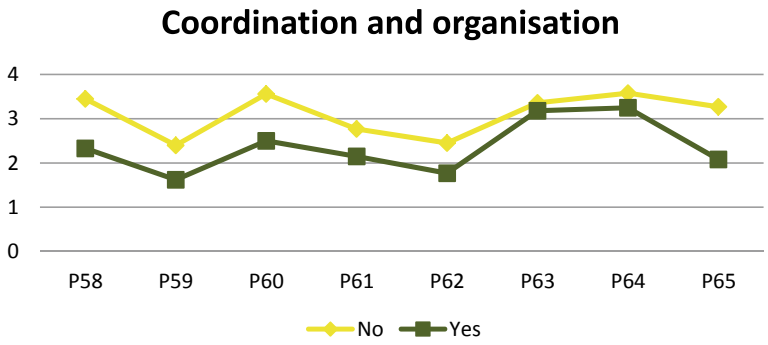


Fig. 12 Coordination and organisation according to bilingual coordinators and non-coordinators

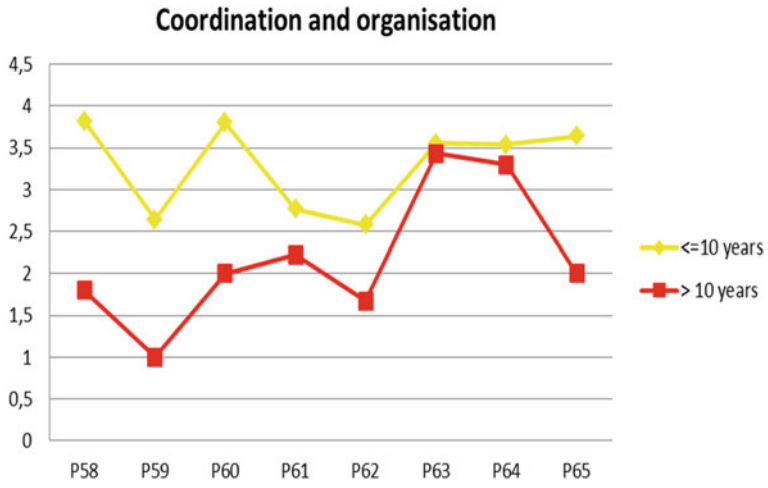


Fig. 13 Coordination and organisation according to overall teaching experience

As far as the variable of bilingual coordination is concerned, within the block of coordination and organisation, the fact that some of the teachers whose sessions have been observed are also bilingual coordinators implies that they have more positive values for the eight items (Fig. 12). The reason for this may be the very crucial role they play in the successful implementation of the APPP in the different Andalusian Primary and Secondary schools. Their very position makes them highlight the relevance of coordination among all the stakeholders involved in these projects, which will prove essential for their educational goals to be reached. However, there are only statistically significant differences in three out of eight items, namely, P58, P60, and P65 ('Coordination among NLA teachers and TA can be verified', 'Coordination among EFL teachers and TA can be verified', and 'There is collaboration for material preparation and design', respectively).

To our surprise, the variable with the second-highest number of statistically significant differences (in 20 out of 71 items) is also related to experience. More specifically, in this case it is connected with the teachers' overall teaching experience, according to which they have been divided into the following two groups: on the one hand, educators with ten or fewer years of experience and, on the other, teachers with more than ten. The latter yield lower (i.e. more positive) averages than the former in more than half of the items (41/63).

The first block (namely, foreign language use in class) is the section in which experienced teachers obtain more positive averages and, together with coordination and organisation, the block comprising the highest number of statistically significant differences. Thus, both educators with more experience and the students attending their sessions spend less time code-switching and translating from Spanish into English in class. In addition, in the section focusing on coordination and organisation (Fig. 13) the same conclusions drawn above according to the teaching experience in bilingual schools variable apply.

Finally, in the rest of blocks it can be identified that more experienced teachers yield a majority of items with favourable averages (blocks 2, 3, 5, 6, and 7), except for the fourth section (methodology and types of groupings), in which there is a balance since each cohort yields lower averages than the other in nine out of the 18 items.

6 Conclusion

We are of the opinion that the way out of the Content and Language Integrated Learning (CLIL) conceptual and pedagogical conundrums evinced in the specialised literature is describing representative practices. Thus, we agree with those authors who have already encouraged this solution to the challenges of the approach posed by the controversies on these two fronts (see Sect. 1). The case study presented here answers the call for conducting extensive classroom observation in a variety of contexts with a view to painting a clearer picture of CLIL. It shows what the approach looks like at the grassroots level in some schools of the Andalusian provinces of Jaén and Granada.

The following findings with regard to objectives a-g of our first metaconcern can be summarised. Vis-á-vis foreign language use in class, the results are quite positive in general, especially on the part of teachers. Students, too, have the adequate level of English expected from them, although they tend to use Spanish more frequently than their educators. Two negative points must be noted, however. First, sometimes English is less used in EFL classes than in the content sessions. Second, the pupils' level of the foreign language is very heterogeneous in a few cases because not all of them have been participating in the CLIL programme for the same amount of time. In connection with the two aspects that were next canvassed, discursive functions and competence development, the outcomes also show a favourable implementation of the approach.

In relation to methodology and types of groupings, on the one hand, and materials and resources, on the other hand, the insights gleaned are not as positive as in the previous blocks. Nonetheless, they are not discouraging either. In a slight majority of the classes observed, most of the pedagogical options examined are followed and approximately half of the tools listed in the protocol are exploited. The room for improvement here has to do with the teachers setting up project work and a wider variety of interaction patterns, being able to count on appropriate NLA textbooks in English that may reduce their burden of lesson planning, and employing materials that take attention to diversity into account.

Pertaining to coordination and organisation, the CLIL contexts assessed again yield mixed results. It can be argued that linguistic learning is supported in content classes and that English as a foreign language (EFL) and non-linguistic area (NLA) teachers get in contact with each other to plan their teaching, but this is not the case between them and the teaching assistants (TAs). We did not find substantial evidence either to assert that interdisciplinary work and connections between the first language (L1) and the second language (L2) are common practices. Finally, CLIL evaluation can be featured in general terms as summative, formative, including an oral component, and giving priority to content over language. The only negative side in this respect is the little attention paid to diversity (which tallies with its scarce presence in materials, as mentioned above).

The main conclusion that can be drawn from the control of a series of variables upon the overall perceptions gathered (*metaconcern 2*) is that experience is a plus for a favourable implementation of CLIL. Not only do practitioners with more than ten years of teaching experience in general develop the approach better, but also those who have participated for over five years in a bilingual programme in particular. In addition, being the person in charge of the programme contributes to a more coordinated and organised model.

The amount of CLIL contexts depicted so far is still insufficient to satisfy the recent surge of demand in the previous literature for recognising the different versions of the approach that can lead to overcoming its vagueness and ambiguity, both conceptually and pedagogically speaking. Thus, we strongly encourage readers to replicate this study in empirically sound conditions and using the observation protocol we followed (or a similar validated instrument). A large number of accounts of what CLIL looks like in its many possible scenarios would ultimately contribute to honing and fine-tuning the requested characterisation of this beneficial and cutting-edge language learning approach.

Acknowledgements This chapter reports on some of the outcomes of the R&D projects FFI2012-32221 and P12-HUM-2348, funded by the *Ministerio de Economía y Competitividad* and the *Junta de Andalucía*, respectively.[6]

[6]Further information about these projects can be found at www.monclil.com.

References

Brown, J. D. (2001). *Using surveys in language programs*. Cambridge: Cambridge University Press.

Bruton, A. (2011a). Is CLIL so beneficial, or just selective? Re-evaluating some of the research. *System, 39*(4), 523–532. https://doi.org/10.1016/j.system.2011.08.002.

Bruton, A. (2011b). Are the differences between CLIL and non-CLIL groups in Andalusia due to CLIL? A reply to Lorenzo, Casal, and Moore (2010). *Applied Linguistics, 32*(2), 236–241. https://doi.org/10.1093/applin/amr007.

Cenoz, J., Genesee, F., & Gorter, D. (2013). Critical analysis of CLIL: Taking stock and looking forward. *Applied Linguistics, 35*(3), 243–262. https://doi.org/10.1093/applin/amt011.

Council of Europe. (2001). *Common European framework of reference for languages: Learning, teaching, assessment*. Strasburg: Cambridge University Press.

Coyle, D. (2007). Content and language integrated learning: Towards a connected research agenda for CLIL pedagogies. *International Journal of Bilingual Education and Bilingualism, 10*(5), 543–562. https://doi.org/10.2167/beb459.0.

Dalton-Puffer, C., Llinares, A., Lorenzo, F., and Nikula, T. (2014). "You can stand under my umbrella". Immersion, CLIL and bilingual education. A response to Cenoz, Genesee, and Gorter (2013). *Applied Linguistics, 35*(2), 213–218. https://doi.org/10.1093/applin/amu010.

Denzin, N. K. (Ed.). (1970). *Sociological methods: A source book*. Chicago, IL: Aldine.

Glaser, B. G., & Strauss, A. L. (1967). *The discovery of Grounded Theory: Strategies for qualitative research*. Chicago, IL: Aldine.

Junta de Andalucía. (2005). *Plan de fomento del plurilingüismo en Andalucía*. Sevilla: Junta de Andalucía. http://www.juntadeandalucia.es/boja/boletines/2005/65/d/5.html. Accessed 26 Feb. 2021.

Lasagabaster, D. (2008). Foreign language competence in content and language integrated courses. *The Open Applied Linguistics Journal,1*, 31–42. http://laslab.org/david. Accessed 26 Feb. 2021.

Paran, A. (2013). Content and language integrated learning: Panacea or policy borrowing myth? *Applied Linguistics Review, 4*(2), 317–342. http://doi.org/10.1515/applirev-2013-0014.

Patton, M. Q. (1987). *How to use qualitative methods in evaluation*. Newbury Park, CA: Sage.

Pérez Cañado, M. L. (2012). CLIL research in Europe: Past, present, and future. *International Journal of Bilingual Education and Bilingualism, 15*(3), 315–341. http://doi.org/10.1080/13670050.2011.630064.

Pérez Cañado, M. L. (2016). From the CLIL craze to the CLIL conundrum: Addressing the current CLIL controversy. *Bellaterra Journal of Teaching & Learning Language & Literature, 9*(1), 9–31. http://doi.org/10.5565/rev/jtl3.667.

Pérez Cañado, M. L. (2017). Stopping the 'pendulum effect' in CLIL research: Finding the balance between Pollyanna and Scrooge. *Applied Linguistics Review, 8*(1), 79–99. http://doi.org/10.1515/applirev-2016-2001.

Pérez Cañado, M. L., & Lancaster, N. K. (2017). The effects of CLIL on oral comprehension and production: A longitudinal study. *Language, Culture and Curriculum, 30*(3), 300–316. http://doi.org/10.1080/07908318.2017.1338717.

Ruiz Gómez, D. A. (2015). A practical approach to CLIL in L2 content-based courses: Methodological guidelines for the Andalusian bilingual classroom. In D. Marsh, M. L. Pérez Cañado, & J. Ráez Padilla (Eds.), *CLIL in action: Voices from the classroom* (pp. 14–30). Newcastle upon Tyne: Cambridge Scholars Publishing.

CLIL and ELF: Friends or Foes?

María Luisa Pérez Cañado

Abstract The characterisation of Content and Language Integrated Learning (CLIL) as an approach to bilingual education has aroused great controversy, particularly in monolingual contexts in Spain. The issue of whether CLIL is bilingual education, of whether the schools where it is implemented are truly bilingual, and whether students who attend them can be expected to attain a balanced bilingualism has sparked off heated debate. This chapter seeks to shed light on this characterisation controversy by setting forth an innovative proposal to examine CLIL in connection to English as a lingua franca (ELF), two acronyms which take centre-stage in the current ELT literature. While the former-CLIL-has burgeoned as a proactive attempt to upgrade European language competence standards, the latter-ELF-favours more realistic language teaching and is willing to forego instruction in certain features which do not hinder intelligibility. However, despite their seemingly contradictory nature, this chapter contends that both approaches are in fact compatible, parallel, and can be taught in connection to each other. After defining CLIL and ELF, the chapter expounds on the main commonalities which can be detected between them and goes on to offer pedagogical implications for their joint incorporation into the ELT classroom. The ultimate aim is to dispel some misconstrued perceptions regarding the characterisation of both approaches and to attune English language teaching to the future needs of current students.

1 Introduction

It is an uncontested fact that we are currently living times of profound change in English language teaching. We are undergoing what Mehisto (2008) terms a period of disjuncture, characterised by the tension between the previous order and a new approach which changes the *status quo*. The demands posed by our increasingly multilingual and multicultural society resonate directly through to the curriculum and

M. L. Pérez Cañado (✉)
University of Jaén, Jaén, Spain
e-mail: mlperez@ujaen.es

M. L. Pérez Cañado (eds.), *Content and Language Integrated Learning in Monolingual Settings*, Multilingual Education 38,
https://doi.org/10.1007/978-3-030-68329-0_3

the need for what Aronin and Hufeisen (2009, p. 105) denominate 'a new linguistic dispensation' arises. In response to this new global order, two acronyms have emerged as timely solutions in harmony with the current needs of ELT: Content and Language Integrated Learning (CLIL) and English as a lingua franca (ELF).

The former-CLIL-is 'a dual-focussed education approach in which an additional language is used for the learning and teaching of both content and language' (Marsh and Langé, 2000, p. 2). It involves a 'two for one' approach (Lightbown and Spada cited in Lyster, 2007, p. 2), where subject matter teaching is used at least some of the time as a means of increased meaningful exposure to the target language. CLIL has burgeoned since the mid-1990s in order to upgrade foreign language proficiency in Europe and to achieve sustainable learning outcomes, in line with the European Commission's 1995 mandate for EU citizens to be proficient in three European languages (the mother tongue + two objective).

In turn, the latter-ELF-has concomitantly been gathering momentum in the past decade in response to the increased use and ownership of English by non-native speakers. As Seidlhofer (2003, pp. 8–9) puts it, ELF occurs 'wherever English is chosen as the preferred option for cross-cultural communication'. It thus emerges as a way to attune ELT (English language teaching) to an increasingly patent reality, favouring 'a more realistic approach' (Siqueira, 2015, p. 242) to it, where perfect communication through a native-like proficiency gives way to competent and intelligible language use (Ur, 2011).

Thus, at first blush, it would seem that these two current trends in language teaching are incompatible and therefore challenging—if not impossible—to reconcile: whereas CLIL is a bid to step up language teaching measures and increase the linguistic proficiency of Europeans, ELF favours a more realistic model of linguistic instruction, where the phonological, grammatical, lexical, and pragmatic features which do not hinder intelligibility are downplayed. Whereas CLIL seeks to upgrade language proficiency, ELF might be held to be conducive to downgrading it, in its move away from the native speaker model.

However, while the overriding impression might be that CLIL and ELF pull ELT in different directions, this chapter contends that, when we delve deeper into their nature and pedagogical implications, the opposite holds true. It unpacks ten main fronts on which both concepts conflate and showcases the view that they are in fact compatible, parallel, and can be incorporated in the language classroom in connection to each other for the benefit of the stakeholders involved. In doing so, it hopes to shed light on the CLIL characterisation controversy (Pérez Cañado, 2016a), by distilling key traits of this approach and debunking false myths which have proliferated around its core features (Pérez Cañado, 2020), while concomitantly and innovatively connecting it to ELF as a way to bridge the gap between the latter el EFL (English as a foreign language).

2 Similarities Between CLIL and ELF

2.1 Their Pragmatic, Proactive, and Controversial Nature

A first commonality shared by CLIL and ELF is their strongly *pragmatic* orientation. CLIL has flourished as a response to Europe's desire to become the most competitive knowledge-based economy in the world (Marsh, 2002). It helps prepare learners to be flexible professionals who can adapt to the varied, unforeseeable, and complex situations they will encounter throughout their personal, social, and professional lives, thus forming successful citizens with a substantial contribution to make to society. Dual-focussed education prepares students for internationalisation and EU integration, for future studies and/or working life, and for lifelong learning. It results in increased employability and better equips individuals for the linguistic and cultural demands of an increasingly integrated and mobile Europe.

In turn, ELF is eminently practical in teaching only those features that are crucial for international intelligibility (Seidlhofer, 2005). Those grammatical (e.g. third-person singular -s), lexical (e.g. idioms and phrasal verbs), or phonological (e.g. different 'th' sounds) aspects which are not conducive to misunderstanding are not accorded attention in class, thereby freeing up valuable teaching time.

In this sense, both CLIL and ELF are crucial to develop a series of soft skills or 'habilidades para la vida' (Moraleja Novillo, 2018, p. 5) which are absolutely pivotal for students in what this same author terms the Fourth Industrial Revolution (cf. Chapter 4 for the impact of CLIL on one of these competencies). With the advent of the latter, we are now faced with a society characterised by its volatility, uncertainty, complexity, and ambiguity (VUCA) (Moraleja Novillo, 2018) which will require what Brende (2019, p. 1) terms 'a reskilling revolution'. According to *The Future of Jobs Report* (World Economic Forum, 2018), 75 million existing jobs will be displaced by 2022 and 133 million new roles will be created. This will necessitate the reskilling and upskilling of 54% of all employees in the next three years (Brende, 2019). The teaching-learning process can thus no longer confine itself to being transmissive and memoristic, but needs to place the onus on the development of the afore-mentioned soft skills, which involve core competencies such as cooperative work, conflict resolution, initiative and leadership, creativity, resilience, proactiveness, decision-making, critical thinking, and, ultimately, lifelong learning. They will favour adaptability to change and to the unforeseen circumstances which graduates will inevitably face in their personal and professional lives and will promote significant and lasting learning. By involving 'a significant change towards language learning for real life' (Köhn, 2015, p. 52), both CLIL and ELF contribute to the development of these pragmatic competencies or soft skills and help 'prepare students for authentic international/intercultural communication' (Vettorel and Corrizzatto, 2016, p. 503).

Thus, they both constitute *proactive* responses coherent with the current linguistic situation: CLIL attempts to redress the deficient foreign language competence levels

in Europe by creating situations which reinforce the continent's levels of multilingualism, while ELF has been fuelled by the historically unique position of English in the world, where non-native users outnumber native speakers (Seidlhofer, 2003). Both approaches to language teaching are increasingly filtering down into the specialised literature and on-the-ground classroom practice, and will undoubtedly contribute to shape the near future of ELT.

Perhaps precisely due to these forward-thinking traits of both approaches, both CLIL and ELF have aroused great *controversy*, a further trait which they share. Indeed, on the one hand, the CLIL controversy hinges on three main fronts: characterisation, implementation, and research (cf. Pérez Cañado, 2016a and Chapter 1 in this volume). In this sense, from the so-called 4 Cs (culture, communication, content, cognition) coined by Coyle (2007), we have moved on to the new 4 Cs (Pérez Cañado, 2018a) in the CLIL arena (craze, critique, conundrum, controversy). Indeed, researchers initially underwent a period of CLIL craze, where research on CLIL (e.g. Lorenzo, et al. 2009; Lorenzo, et al. 2009) only sung its praises and shone an exclusively positive spotlight on its functioning. The problem with these initial studies is that they presented some potentially very serious methodological flaws which could invalidate their results (cf. Pérez Cañado, 2012). From there, the metaphorical pendulum swung violently to the other extreme and researchers entered a phase of intense CLIL critique, where publications dwelled almost exclusively on the negative and pessimistic side of CLIL (cf. Bruton, 2013, 2015, 2019; Paran, 2013). The problem with this second batch of articles is that they were based on unsubstantiated opinions or personal cases, thereby lacking scientific robustness, representativeness, or generalisability. This situation has led to a CLIL conundrum, where confusion as regards the effectiveness of CLIL has predominated, and has caused researchers to be currently positioned within a CLIL controversy, which is playing itself out in some of the major journals in the field, where articles are being tossed back and forth between CLIL advocates and CLIL detractors. Intense debate has been sparked off and contradictory opinions have been harboured vis-à-vis pivotal aspects of CLIL, thereby creating the need to revisit some taken-for-granted issues affecting this approach.

Similarly, ELF has been considered to be 'probably the most radical and controversial approach to emerge in recent years' (Graddol, 2006, p. 87). Parallel to the 4 Cs in CLIL, resistance to ELF can also be summarised with, in this case, a four-pronged acronym: PESTS. It stands for '"Practicality" (not practical to teach different types of English), "Efficiency" (just a waste of time), "Standards" (need to learn standard native English), and "Simplicity" (don't like to be confused, keep things simple)' (Marlina, 2014, p. 9). Indeed, the instability, fluidity, and variability inherent in ELF (Blair, 2020) have led to challenges and obstacles in its application. According to Sifakis (2020), these hurdles are fundamentally two: the negative attitudes towards ELF which teachers tend to harbour due to the strong convictions they normally have about what is standard and appropriate in language teaching, and the lack of a comprehensive instructional framework for ELF, which leads to 'an uncertainty about establishing, applying and evaluating appropriate ELF pedagogy' (Sifakis, 2020, p. 108).

The controversy which surrounds both approaches attests to their exponential uptake as burning issues on the current ELT agenda. While the debate it has spurred on has been very healthy and has blown refreshingly new life into the field, we clearly need to shed light on it in order to bring the metaphorical pendulum to a standstill and ensure that only 'factfulness', as Rosling (2018, p. 1) terms it, that is, the habit of only carrying opinions for which there are strong supporting facts, is our compass for the future.

2.2 The Type of Language Taught

A second trait which CLIL and EIL have in common involves the type of language taught. Both meaningful exposure and language-based focus on form are part and parcel of both approaches. Indeed, on the one hand, both CLIL and ELF target successful communication in uncontrived real-life situations, with a primacy of fluency over accuracy. Recent research (cf. Lancaster, 2018) has evinced that meaningful, communicative exposure to the foreign language through CLIL (with no extra formal instruction) is a superior instructional option to exclusive EFL teaching with extra formal instruction. Similarly, ELF veers towards 'a more communication-oriented and inclusive approach' (Vettorel and Corrizzatto, 2016, p. 505).

However, on the other, a focus on form is also integral to both CLIL and ELF. Lyster (2007) makes a strong case for the inclusion of instructional plans with language objectives in CLIL programmes, involving noticing activities, increase in metalinguistic awareness, and opportunities for production practice. Exposure and authentic communication, he maintains, are not sufficient to push interlanguage development forward in CLIL scenarios. Much the same view is harboured by Ur (2011) for ELF: grammatical accuracy, vocabulary breadth, and error correction all have a substantial place within ELF and need to be provided to the student in order for him/her to become a fully competent speaker of English. Thus, intelligibility and comprehensibility are not all English learners need (Sung, 2013). This would be a reductive approach, as 'a restricted focus on features crucial to intelligibility might result in an impoverished syllabus in ELT' (Sung, 2013, p. 178).

2.3 The Language Level Targeted

Such a full competence, however, does not entail attaining a native-like level in the target language: both CLIL and ELF aim for a functional as opposed to a (near) native-like competence. Indeed, as regards CLIL, 'the term "bilingual" in most parts of Europe does not refer to the mastery of a second language at native-like proficiency, but it describes a partial foreign language competence' (Brüning and Purrmann, 2014, p. 315). As Mearns (2012, p. 176) has put it, 'Unlike the case of EFL approaches, the

main goal of CLIL is to produce competent and confident target language users, while at the same time teaching subject content'. Similarly, 'The priority for students using ELF (…) is to be as intelligible as possible to the people they are communicating with. This does not necessarily mean sounding like a native speaker' (Davies and Patsko, 2013).

Thus, the competent speaker of English in CLIL and international settings may or may not be originally a native speaker, which is why Ur (2011) suggests re-defining Kachru's (1992) circles of English users in terms of their level of competence (limited—competent—fully competent), instead of considering where they live or whether or not they are native speakers. Rampton (1996) also suggests replacing the terms native speaker and non-native speaker with expert and novice, respectively, with the intention of placing the onus on language expertise rather than nativeness in conceptualising language competence. Similarly, Blair (2020) prefers the denomination speakers of the target language to that of native speakers in order to focus on intelligibility and ease of communication. In this sense, the new *CEFR Companion Volume* (Council of Europe, 2018) reflects this all-important shift by updating existing descriptors to include 'speakers of the target language' instead of 'native speakers'.

This more attainable level of the foreign language inherent in both CLIL and ELF involves departing from a standard English model as a convenient starting point in the classroom, but having students develop their own variety as L2 speakers (Sung, 2013). This model is seen as a point of reference (a sort of guide from which variations or deviations are considered acceptable) rather than a target (the standard which the teacher establishes as a goal for students) (Hewings, 2004). This is what Köhn (2015, p. 59) refers to as the creation of 'a pedagogical space for ELF in the English classroom'. It requires a shift from a strong to a weak standard English orientation, which is post-normative, social constructivist, and allows students to develop EFL-specific creativity. In this sense, deviations are tolerated to a greater extent, both in ELF and in CLIL: 'CLIL classrooms clearly require, and indeed practise, a greater tolerance for deviations' (Köhn 2015, p. 54). Both approaches also entail toning down and standardising native varieties (Siqueira, 2015), as fully competent speakers can usually identify in an intuitive way which features of their own speech are (not) internationally standard (Ur, 2011).

2.4 The Teacher and His/Her Training

The foregoing has important implications for the teacher. (S)he need no longer have a native-like competence (Lasagabaster and Sierra 2010). As Ur (2011) underscores, the native—non-native distinction is becoming increasingly irrelevant. To take a case in point, in southern Spain, non-linguistic area teachers participating in CLIL programmes are only required to certify a CEFR B2 level of English, which distinguishes this type of programme from immersion ones, whose teaching staff normally comprises a high percentage of native speakers (Lasagabaster and Sierra, 2010). Instead of simply having a native-like mastery of the English language, CLIL teachers

now need to be well-versed in academic, disciplinary, or subject-related literacies, together with EFL skills. These discipline-specific skills require the development of cognitive discourse functions (Dalton-Puffer, 2013) such as describing, explaining, defining, comparing, or evaluating, together with the development of facts, concepts, procedures, and strategies. This is precisely what is promoted by the pluriliteracies approach (Meyer, et al. 2015). This model makes the integration of content and language more conspicuous and sees it as fundamental to learning. It favours the conflation of BICS (Basic Interpersonal Communication Skills, Cummins, 1999) and CALP (Cognitive academic Language Proficiency, Cummins 1999)[1] and challenges the division of both concepts, strengthening the relationship between the conceptual and communicative continua.

In turn, in ELF teaching, the non-native fully competent speaker becomes an achievable and adequate role model. The 'omniscient' native speaker model (Alptekin, 2002; Rajagopalan, 2004) is now seen as 'conveying a narrow and unfaithful portrait of the multifaceted reality of English, as well as of its uses and users' (Vettorel and Corrizzatto, 2016, p. 489). Thus, what is now important is the instructor's level of linguistic and intercultural competence and his/her teaching ability, and not whether (s)he is a native speaker of the target language.

In line with the foregoing, in addition to language mastery at an attainable level, the CLIL and ELF teachers now have to develop a new skillset that goes well beyond merely linguistic aspects. Teacher education acquires a particularly sharp relief in both cases, as further commonality which can be detected between the two approaches. As Siqueira (2017, p. 391) puts it, 'Concerning English specifically, due to its spectacular spread around the planet, the quality of teacher professional preparation has become a key issue'. The novelty, recency, and crucial implications of CLIL and ELF call for a new teacher development paradigm which comprises 'premises, practices, conceptions, strategies and tools very diverse from the ones teacher educators have been utilising all these years' (Siqueira, 2017, p. 402). This paradigm requires a rethinking of teacher training in order to bridge the gap between implication and application, ensuring such ongoing professional development truly trickles down to on-the-ground practice. As Dewey and Patsko (2017, pp. 1–2) have put it, 'the way ELT has conventionally tended to conceive subject knowledge in teacher education is no longer relevant or appropriate when we take into account the global lingua franca role of English'.

In this sense, great headway has been made in determining the key competencies which CLIL and EFL teachers need to master and in setting forth specific teacher development programmes which sharpen and refine them and which are attuned to actual realities and real teacher training needs. In CLIL, specific projects (e.g. the NALTT[2] Project) have tapped into the main needs which teachers have to step up

[1] Basic Interpersonal Communication Skills (BICS) refers to the everyday, here-and-now language commonly used for social interaction in the classroom and Cognitive Academic Language Proficiency (CALP), to the more abstract, specialised language used for academic purposes.

[2] *Needs Analysis of Language Teacher Training: A European Perspective* (EA2010-0087), funded by the Ministry of Education (Spain).

to the bilingual challenge (cf. Pérez Cañado, 2016b, c). Concrete teacher training proposals (like the *Máster Interuniversitario en Enseñanza Bilingüe y AICLE*[3]) have then been set forth based on these context-specific and, hence, relevant studies (cf. Pérez Cañado, 2015 for its structure and rationale). Similarly, in ELF, as Siqueira has recently underscored (2020, p. 118), 'research findings are slowly gaining ground in regular ELT classrooms, especially due to the inclusion of ELF-related issues in teacher education programmes, both at pre-service and in-service levels'. Indeed, as Vettorel and Corrizzatto (2016) have also highlighted, ELF is now increasingly present in Cambridge ESOL DELTA and pre-service teacher education initiatives.

Which are the teacher competencies which these programmes seek to hone and which have been detected by recent research proposals? Vis-à-vis CLIL, over the past half a decade, conspicuous efforts have been made to map out a clear-cut CLIL teacher profile. Several proposals stand out in this sense: those put forward by Bertaux, et al. (2010), Hansen-Pauly, et al. (2009), and Marsh, et al. (2010) in Europe and those set forth by Lorenzo, et al. (2011), Madrid Manrique and Madrid Fernández (2014), and Pérez Cañado (2018b) in Spain. Seven core CLIL teacher competences transpire in all of them. An initial pivotal one is *linguistic competence*, which also encompasses intercultural aspects and centres on both BICS, CALP, and pluriliteracies (Cummins, 1999; Meyer, et al. 2015). This dovetails with *pedagogical competence*, whereby teachers need to be familiarised with a host of student-centred methodologies, more diversified learning environments and resources (where ICTs acquire a particularly sharp relief), and a more transparent, holistic, and formative type of evaluation. Concomitantly, practitioners must have *scientific knowledge*, which involves familiarity with the theoretical underpinnings of CLIL. This knowledge should be compounded with *organisational competence*, as the vast gamut of groupings and learning modalities that now burgeon within CLIL need to be successfully deployed by teachers, together with classroom management and control strategies. *Interpersonal* and *collaborative competencies* are two further areas that are integral to the CLIL teacher profile and that entail the capacity to create an adequate classroom atmosphere where students receive personalised attention and feel safe and unthreatened to participate and take risks, as well as the capacity to liaise with colleagues to a greater extent, stepping up collaboration and teamwork. All these competencies are fully commensurate with the final one—*reflective and developmental competence*—, which points to the need for lifelong learning and for teachers to be constantly up-to-speed with the latest information and research on CLIL developments.

Again evincing great parallelisms with CLIL, the ELF teacher profile is also increasingly taking shape in recent proposals (Blair, 2020; Dewey and Patsko, 2017; Köhn, 2015; Sifakis 2020; Siquiera, 2017; Sung, 2013; Vettorel and Corrizzatto, 2016). Teachers need to have what Sifakis (2020, p. 106) has termed 'ELF awareness'. As this author highlights, it is necessary to train 'ELF-aware teachers as informed practitioners and reflective and autonomous agents of change for their teaching context' (Sifakis, 2020, p. 113). This awareness hinges on several different

[3]Cf. https://www.ujaen.es/estudios/oferta-academica/masteres/master-interuniversitario-en-ensenanza-bilingue-y-aprendizaje-integrado-de-contenidos-y.

fronts. It is incumbent upon them to be *aware* of the varieties of English they need to master and teach (Pedrazzini and Nava, 2010), of the sociolinguistic changes engendered by the current pluralisation of English and by its lingua franca status (Vettorel and Corrizzatto, 2016), and of ELF use in reality, including the notion of language variation and change in ELF and the role of English in today's world (Sung, 2013). However, awareness-raising is at present simply regarded as an initial impetus or jump-off point (Dewey and Patsko, 2017). Other issues are now acquiring an increased salience on the ELF teacher competence agenda. The ELF teacher also needs to be a fundamentally *critical intercultural educator* (Siqueira, 2017), who can foster reflection on implications for pedagogical approaches (Vettorel and Corrizzatto 2016), on current thinking and practice (Dewey and Patsko, 2017), and on the evaluation of existing materials (Siqueira, 2015). In Siqueira's (2017, p. 400) words, 'in order not to be considered some type of "illiterate of the twenty-first century", language teacher educators should make themselves willing to learn, unlearn and relearn'. They equally need to be *sensitive* the '"pains" felt by future global language learners' (Siqueira, 2017, p. 403), to the potential of ELF in practice (Blair, 2020), and to their own 'glocal' context, which should lead towards pedagogic choices that can be locally tailored to the students' (present and future) needs and contexts of use for English (Vettorel and Corrizzatto, 2016). A further trait which ELF practitioners should exhibit, in harmony with CLIL teachers, is enhanced *collaboration* with their colleagues 'in order to fight for the updating of curricula, courses, programmes and disciplines' (Siqueira, 2017, p. 402). They should also have a *language competence* which allows them to compare their students' pronunciation with the requirements of the Lingua Franca Core (Jenkins, 2000) and with different reference models. Finally, they should develop a *pedagogical competence* which capacitates them to devise ELF-aware activities, tasks, and lesson plans, adapting and suiting them to each concrete learning and teaching context (Sifakis, 2020; Vettorel and Corrizzatto, 2016).

2.5 The Learner

For the learner, the implications of CLIL and ELF are no less significant. The securely established position of ELF for the next 50 years, according to Seidlhofer (2003), means that English will remain firmly entrenched in the educational system and will increase its outreach to more diversified levels of society. ELF will increasingly trickle down to classroom praxis and curricula, much in the same way as CLIL has had an exponential uptake over the past two decades. Numerous authors testify to this rapid and widespread adoption of CLIL in the European arena (Marsh, 2002), assimilating it to a veritable explosion of interest (Coyle, 2007). It has furthermore embedded itself in mainstream education from pre-school to vocational education rather swiftly, no longer being the prerogative of the academic elite. In fact, several authors go as far as to claim that traditional non-CLIL 'drip-feed education' (Vez, 2009, p. 8) involves moving on the slow track to language learning and that 'CLIL is

bilingual education at a time when teaching through one single language is seen as second rate education' (Lorenzo, 2007, p. 35). Thus, socially, both approaches afford all students, regardless of social class and economic consideration, the opportunity to learn additional languages in a meaningful way. They thus promote social inclusion and egalitarianism, as the introduction of these programmes in mainstream education provides a greater range of students with opportunities for linguistic development which they were previously denied.

Indeed, the latest empirically robust research (Madrid and Barrios 2018; Pavón Vázquez 2018; Pérez Cañado 2020; Rascón Moreno and Bretones Callejas, 2018; or Shepherd and Ainsworth 2017) has focussed on the effects of CLIL on L1, L2, and content learning, bearing in mind crucial contextual, cognitive, and affective variables associated to egalitarianism in CLIL (cf. also Chapters 6, 7, and 8 in this volume). The results are unequivocal and point to an increasingly strong emerging pattern: CLIL programmes seem to be cancelling out differences in rural-urban setting, socioeconomic status, and type of school, especially in the long term. Indeed, differences in terms of these variables are particularly pronounced for non-bilingual learners, but they phase out with CLIL groups, a very encouraging circumstance which shoots down elitist claims for CLIL. Bilingual education programmes thus appear to be causing differences ascribed to sociocultural and socioeconomic standing to be levelling out. And what is more, when research (Pérez Cañado, 2020) has examined the potential of CLIL to work even in the most disenfranchised settings (in a public school within a rural context, in a socioeconomically deprived area, with low sociocultural level on the part of parents, and with a majority of students of gypsy ethnicity), it has transpired that CLIL indeed has this potential, as statistically significant differences emerge on absolutely all the linguistic aspects considered in favour of the CLIL group and with extremely high confidence levels. Thus, the outcomes reveal that CLIL 'has the potential to work even in the most deprived settings', where students would not have previously had access to bilingual education had it not firmly embedded itself in public contexts (Pérez Cañado, 2020, p. 11). Similarly, the English as a native language model in the ELT classroom 'is seen to be advantageous for only a very small percentage of the total population of teachers and learners worldwide' (Sung, 2013, p. 174). ELF, on the other hand, is tailored to learners' realistic needs and wants (Sifakis, 2020) and favours a more authentic intercultural communication (Vettorel and Corrizzatto, 2016) which is more accessible for all types of students.

2.6 The Role of Intercultural Awareness

Further pedagogical implications which both ELF and CLIL have affect the role of intercultural awareness. Both acknowledge the importance of developing intercultural understanding and global citizenship issues in order to adapt and function appropriately when interacting with people from other cultures. Blair (2020, p. 9) terms it 'metacultural competence' for English language pedagogy and considers it essential in preparing learners for 'the sociolinguistic reality of the use of English in

today's globalised world'. It entails following a different approach, which places the emphasis on intelligibility, flexibility, and developing intercultural communication skills instead of merely prioritising the acquisition of a prescribed native accent.

As Siquiera (2012, p. 201) has put it, 'because of its de-nationalisation, English has become a language which represents and provides access to several cultures, including those from countries and territories practically unknown or strange to many people'. Intercultural or metacultural competence thus acquires an increased salience due to the fact that English is no longer linked to a single culture, but to diverse languacultures. In ELF, English is used as a vehicle to raise awareness of three types of culture: the home culture, the cultures of those who are able to speak English, and the culture of English-speaking peoples (Ur, 2011). Siqueira (2012), in turn, speaks of three different cultural sources: the source culture (where the student's native culture is represented), the target culture (involving the culture of those countries where English is spoken as first language), and the international target culture (which takes into account the cultures of those countries where English is used, irrespective of whether it has an official status or not).

In turn, the merits of CLIL within the cultural dimension have also been frequently voiced. It builds intercultural knowledge and understanding, develops intercultural communication skills, and promotes intercultural communicative competence. It equally raises awareness of cultures and the global citizenship agenda. In sum, it offers powerful opportunities for intercultural learning which go beyond those provided by traditional language learning settings. It is what Coyle (2007) refers to as the 4 Cs framework: *content* (subject matter), *communication* (language learning and using), *cognition* (learning and thinking processes), and *culture* (intercultural understanding and global citizenship).

In both contexts, it is thus now pivotal that the teacher develop this intercultural awareness. (S)he becomes 'the catalyst for an ever-widening critical cultural competence' (Kramsch, 1993, p. 10) and must be critical instead of subservient: 'a critical teacher makes all the difference. (…) a professional who is fully aware of his/her convictions related to a cultural approach to FL teaching, anchored in an intercultural perspective' (Siqueira, 2012, p. 206).

2.7 Materials

The afore-mentioned intercultural awareness needs to be reflected in the materials employed. This is a key pedagogical aspect related to CLIL and ELF which evinces conspicuous lacunae. Indeed, for both approaches, there is a well-documented scarcity of appropriate materials and instructional resources. Authentic materials are no longer valid options for teaching content and language in an integrated way or English as a lingua franca, and they consequently need to be originally designed or adapted. In CLIL programmes, poor access to materials in English and ICT availability has been highlighted by Czura et al. (2009). Those which *are* available lack quality, practicality, and feasibility (Ruiz Gómez and Nieto García, 2009), so that

offering guidelines for adequate materials design becomes paramount in a CLIL context. Although in recent investigations (cf. Pérez Cañado, 2018c), the materials employed are considered communicative, interesting, and innovative, new technologies are still not prominently used. Interactive whiteboards, multimedia, and online materials are incorporated to a greater extent, but blogs, wikis, webquests, and computer-mediated communication are scarcely employed. It also transpires that materials are primarily elaborated from scratch, collaboratively prepared, or adapted by teachers, as authentic materials are considered very poor. Textbooks are now consequently relegated to a secondary position in CLIL classrooms. Both the elaboration and updating of materials are entailing considerable effort for the practitioners involved. They complain about the time-consuming nature of this endeavour, which is often not acknowledged, and about the scarcity of resources at their reach. Two further deficiencies on this front affect lack of materials to cater to diversity and absence of guidelines in Spanish for parents.

Much the same occurs for ELF: as Seidlhofer (2003, p. 12) stresses, 'EIL is practically non-existent in language teaching curricula and materials …'. As Siqueira (2012, p. 210) underscores, 'The great majority of materials introduce culture as a monolithic, static, and generalised element, with a prevalence, naturally, of cultural references from hegemonic societies'. They do not develop pragmatic competence, a very important aspect for ELF interactions (Vettorel and Corrizzatto, 2016). Textbooks are 'soulless', according to Siqueira (2015, p. 243), with characters and topics that have been 'socially anesthesized and sanitised for commercial purposes'. Through them, the classroom becomes an idealised space, a 'plastic world', a 'fantasy island', a 'pedagogical Disneyland' (Siqueira, 2015, p. 244), instead of reflecting the world which pulses outside. The latter is clearly under-represented in ELT materials and the current global condition of English is demanding changes on this score. Materials should also accommodate the local, reflect local realities, insert local cultures, develop students' intercultural competence, and include topics which are part of the real world so that that 'plastic world of FL textbooks' truly starts to melt (Siqueira, 2015, p. 246). Ur (2011) suggests they should be adapted in terms of content, culture, situations, characters, texts, and language, with a greater inclusion in all cases of international and local sources and fewer inner-circle or authentic instances. And Dewey and Patsko (Davies and Patsko, 2013) propose comparing, e.g. the pronunciation exercises in the coursebook to the pronunciation features identified as important for maintaining intelligibility with ELF, and then ensuring they are attuned to the needs of our students. Existing materials thus need to be analysed, discussed, and updated in more critical and systematic ways so that language pedagogy is realigned with modern times, trends, and needs.

2.8 Methodology

Another pedagogical deficiency affecting ELF and CLIL is methodological: there is currently a rift between the theory and practice of both approaches. While ELF has

been functionally and linguistically acknowledged and described, and there is awareness of its role and salient features, pedagogically, it still remains largely unexplored (Seidlhofer, 2003). Indeed, the special status of English 'has had virtually no effect (so far) on how the language is formulated as a subject in syllabuses and teaching materials' (Seidlhofer, 2011, p. 9). Even though there is awareness of ELF, it is not readily translated into classroom practices (Dewey and Patsko, 2017, p. 3). There is not much sign yet of a movement from awareness to application in teaching (Blair, 2020). Sifakis (2020, p. 111) has in fact highlighted that 'What ELF does not enjoy is a specific way of teaching it that is comprehensively supported by courseware, dictionaries, grammar books and teacher training practices' and Sung (2013, p. 177) even goes as far as to claim that it might be 'inoperable as a pedagogical model'.

Similarly, CLIL is, methodologically speaking, a bid to bring innovation into the classroom, realigning teaching with modern pedagogical practices. It represents an opportunity to move from declarative to procedural knowledge and to minimise the transmission of knowledge model of education, in keeping with the post-methodology era we are currently living. Task-based learning, collaborative work, meaning- and form-focussed processing are all considered pivotal in fostering this transformative pedagogy. However, this considerable pedagogical investment in innovative pedagogical practices which CLIL brings with it may be largely unfamiliar to teachers, not having experienced such practices first-hand as students. In this sense, teacher training actions should be considerably stepped up on the methodological front, as Cabezas Cabello (2010), among others, has underscored.

Thus, a new pedagogical model is called for in order to accommodate both approaches (Sung, 2013). There is a clear need to challenge the pedagogical status quo and re-conceptualise methodological aspects. Current pedagogic practice needs to be rethought in the light of CLIL and ELF (Dewey and Patsko, 2017) and practitioners 'need to abandon their traditional practices and be ready to humbly rethink a lot of what they have been doing over the years in order to indeed prepare intercultural teachers to this more than ever intercultural world' (Siqueira, 2017, p. 400), as earlier orientations have been proven obsolete and outdated.

What should a critically reframed notion of language pedagogy look like in CLIL and ELF? Its traits again run parallel to each other in both approaches. To begin with, this updated methodology should be *student-centered*, *eclectic*, and *plurilingual* in order to promote critical thinking (Dewey and Patsko, 2017; Sifakis, 2020). It should thereby 'empower learners to avoid turning into "teachees" (i.e. passive recipients of what is taught) and grow into active participants in the learning process' (Sifakis, 2020, pp. 111–112). It should equally expose students to a range of native and non-native *varieties* and accents of English, rather than a single or restricted amount (Sung 2013; Köhn, 2015). In order to go beyond static and monolithic representations of the language, authentic materials, such as written texts, audioclips, videoclips, charts, and literary texts can be used to encourage a more dynamic approach to presenting differences in accents and lexical items (Vettorel and Corrizzatto, 2016). In addition to these language varieties, *strategies* now acquire a sharp relief within both CLIL and ELF. Scaffolding, intelligibility, and communication strategies are seen as key elements (Vettorel and Corrizzatto, 2016). Accommodation, interpersonal

and pragmatic competence, and negotiated meaning-making practices (Sifakis, 2020; Siqueira, 2020; Sung, 2013) also come to the fore as essential. These strategies need to be connected to the *local context* of the learner. This is what Sung (2013, p. 184) terms a '*polycentric model* perspective', which 'involves a respect for local variation and a willingness to engage in the shared pursuit of intelligibility and comprehensibility, so that it allows individual variations in ELF use and the expression of local identity, while at the same time enabling the existence of a model or standard'. In this sense, the on-size-fits-all model no longer fits the bill in either approach (Hüttner and Smit, 2014). This entails tailoring language teaching to learners' realistic needs and wants (Sifakis, 2020) and attuning it to *real-life use of English.* Indeed, a shift needs to be operated 'towards language learning for real life' (Köhn, 2015, p. 63). Students should develop an understanding of 'how English language is currently used' (Grazzi and Lopriore, 2020, p. 85) and experience it in real communicative contexts which prepare them for authentic international/intercultural communication (Vettorel and Corrizzatto, 2016) through updated English expressions (Pérez Cañado and Borja-Ojeda, 2018).

Finally, *assessment* also needs to be considerably revamped (Vettorel and Corrizzatto, 2016). A more diversified, ongoing, and holistic evaluation should be set in place and the transition needs to be made towards a more ingrated type of assessment (Otto, 2018), where language-related issues are measured in relation to content objectives. This entails a systemic-functional approach to evaluation, which considers not only if a language form is grammatically correct, but whether it is used appropriately to convey meaning in functional contexts, taking into account communicative intent in terms of language functions rather than linguistic accuracy or grammatical correction. Thus, assessment now needs to consider not only the ability to use linguistic forms correctly, but to use the appropriate form to express meaning in a particular academic context. Attention consequently needs to be paid to how successful communication is achieved, as well as the range and appropriateness of language use, rather than simply whether or not the language produced by the learners conforms to a certain standard (Sung, 2013).

2.9 Research

Much the same needs to happen in the research arena. Perhaps the single most widely consensual affirmation in the specialised literature on both approaches is that research on CLIL and ELF is still at an embryonic stage and consequently needs to be escalated. Seidlhofer (2003, p. 18) makes this point with respect to ELF: '… it is early days yet and no reliable findings based on quantitative investigations can be reported at this stage'. The same conclusion can be reached regarding European CLIL: although the number of studies tapping into the implementation and effects of CLIL has been growing steadily, very few are robust accounts of outcome-oriented research where pertinent variables are factored in and controlled for. The final verdict, thus, is not yet in (Marsh 2002, p. 185): '… there remains insufficient empirical

evidence of the impact of differing types of CLIL/EMILE across Europe'. Exploring the many avenues which are still open to scrutiny in both areas will undoubtedly be the remit of future research.

As in previous areas, parallel pathways for progression can be mapped out for both approaches. To begin with, we stand in need of characterising representative pedagogical practices of both approaches. They are both well-characterised from a theoretical stance, but these traits have as yet had little practical take-up and what they should look like in practice is still a major niche to be filled. In this sense, as CLIL continues to develop over the next decade, its linguistic, methodological, and organisational traits will also need to be honed, sharpened, and fine-tuned in line with the demands of the contexts where it is being applied (Pérez Cañado, 2017). Reflection (e.g. through the *Observatorio de Plurilingüismo* which will be set up in Andalusia), classroom observation (e.g. through specific protocols which have been designed for that purpose—cf. Pérez Cañado, 2016a), and communication (e.g. via job-shadowing or the sharing of good CLIL practices with others) will increasingly characterise representative pedagogical CLIL practices and allow us to make headway in this area. Similarly, in ELF, strides need to be taken in order to determine 'the ways in which it can be integrated into one's classroom context' in order to fill the gap 'that still exists (and that needs urgently to be closed) between ELF research developments and everyday classroom practice' (Siqueira, 2020, p. 138).

In line with the foregoing, replication of existing research on both approaches needs to be carried out within specific contexts in order to base decisions on real needs and ensure teacher preparation is driven by them. As was previously mentioned, the one-size-fits-all model does not fit the bill in CLIL or ELF and we need to maintain a 'context-sensitive stance' on it (Hüttner and Smit 2014, p. 164). To ensure pedagogical decisions are aligned with reliable research outcomes, it thus behooves teachers and practitioners to conduct research at the local level, as this local context of the learner has proved paramount in both approaches.

A third area which should figure prominently on both the CLIL and EFL research agendas is undoubtedly attention to diversity.[4] As Siqueira (2017, p. 395) has underscored, within ELF,

> A denationalised language like English, which today has more non-native speakers than native, brings to visibility several important characteristics such as high levels of hybridity, great diversity of users, and, above all, it makes possible for speakers to engage in more and more meaningful intercultural interactions, having as interlocutors individuals from any part of the world, bearing the most diverse linguacultural backgrounds.

In CLIL, the need to cater for diverse types of learners is no less substantial. This approach has steadily embedded itself in mainstream education and a new CLIL scenario is firmly taking root across the continent, where the move is being made from bilingual sections to fully bilingual schools. Thus, now all learners experience

[4] *Diversity* is understood here as an inherently human trait, based on respect for individual differences and learning styles (Gardner 1993); varying achievement levels, learning paces and intellectual capacity; and diverse motivations and socioeconomic and cultural backgrounds which affect the lives of students (Arnaiz, 2003; Arnaiz and de Haro, 1997; Julius and Madrid, 2017).

foreign language learning both in language-driven and subject content classes and it consequently 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'.

Future research will thus need to address a substantial number of questions which are crucial for the successful development of CLIL and ELF programmes: Do both approaches truly work for all types of learners? Do verbal intelligence, motivation, language level, and academic performance have a bearing on linguistic and content gains? How do other intervening variables such as sociocultural status, type of school, setting, community/country, or extramural exposure interact with the previous ones and impact FL/content competence? What kinds of measures are being set in place to cater to diversity in monolingual communities? How do these compare with those being implemented in northern, central, and eastern Europe? Which measures are working better and why? What can we learn from the best practices of others on attention to diversity in CLIL and ELF in order to improve our own language learning situation and educational system? Very exciting research avenues thus need to be pursued, which largely run parallel to each other in both approaches.

2.10 Pedagogical Implications

Pedagogically speaking, it would also be greatly beneficial to examine CLIL and ELF in connection to each other. In this sense, attaining the plurilingualism targeted by CLIL would be made easier by embracing a realistic approximation to English language teaching which is still accurate, but also pragmatic, and where the competent speaker with a functional linguistic competence becomes the model to follow for both teachers and learners. A change in the criteria for the selection of the language to be taught should be operated, favouring BICS which are more neutral or standardised in terms of language variety. Materials should be adapted in accordance with the criteria of both CLIL and ELF, and intercultural awareness-raising should be worked into the language classroom in response to the demands of both types of programmes. Finally, teacher training actions and research could productively combine both approaches, raising practitioners' awareness of ELF in CLIL Masters and conducting investigations on the effects which CLIL is exerting in countries belonging to Kachru's (1992) expanding circle. In this sense, CLIL could well have the potential to save the divide between ELF and EFL (Köhn, 2015), by following what this author terms a weak native speaker orientation, by understanding language learning as a constructive endonormative process, by targeting a functional competence of the target language, and by accepting deviations if communication is successful in order to focus on comprehensibility. CLIL and ELF can therefore comfortably meet to join forces in their pedagogical endeavours to the benefit of the stakeholders involved.

3 Conclusion

This chapter has set forth an innovative proposal which expounds on the parallelisms between two recent approaches to English language teaching which will undoubtedly continue shaping the direction of this discipline in the future: CLIL and ELF. The main thrust of the discussion has been that, although they might apparently seem antithetical in their goals, both are in fact similar in their origin, nature, type of language and target level, implications for the teacher and learner, methodology and materials employed, and current research panorama. The chapter posits that these commonalities can be fruitfully capitalised upon by incorporating the principles of CLIL and ELF simultaneously into the ELT classroom. Thus, although they have often been pitched against each other, a reconciliation of both approaches is of utmost relevance and deserves the highest priority. As Köhn (2015, p. 65) has put it, 'Hostile brothers who keep a wary eye on each other's views and preferences must shed their suspicion and drop their misgivings to become brothers in arms'.

In doing so, it has characterised CLIL on a multiplicity of levels, fostering the connection to EFL, as a possible way out of the so-called CLIL characterisation controversy, while concomitantly adjusting expectations, teaching, and training to real needs and situations. Taken thus, in conjunction, these seemingly contradictory approaches may well become the lynchpin to boost language learning in EFL contexts with limited extramural exposure while concomitantly making it more realistic and attuned to the future needs of our current language students.

Acknowledgements This article is the result of an extensive review of the literature for the research projects 2018-1-ES01-KA201-050356, funded by the European Union; RTI2018-093390-B-I00, funded by the Spanish *Ministerio de Ciencia, Innovación y Universidades*; and 1263559 and P18-RT-1513, funded by the *Junta de Andalucía.*

References

Alptekin, C. (2002). Towards intercultural communicative competence in ELT. *ELT Journal, 56*(1), 57–64.

Arnaiz, P. (2003). *Educación inclusiva: Una escuela para todos.* Málaga: Editorial Aljibe.

Arnaiz, P., & de Haro, R. (1997). Educación intercultural y atención a la diversidad. In F. Salinas & E. Moreno (Eds.), *Semejanzas, diferencias e intervención educativa* (pp. 141–148). Madrid: CSI-CSIF.

Aronin, L., & Hufeisen, B. (2009). Methods of research in multilingualism studies: Reaching a comprehensive perspective. In L. Aronin & B. Hufeisen (Eds.), *The exploration of multilingualism. Development of research on L3, multilingualism and multiple language acquisition* (pp. 103–120). Amsterdam: John Benjamins.

Bertaux, P., Coonan, C. M., Frigols, M. J., & Mehisto, P. (2010). *The CLIL teacher's competence grid*. Available at http://lendtrento.eu/convegno/files/mehisto.pdf.

Blair, A. (2020). Developing 'ELF competence' in language learners and teachers. *Estudos Linguísticos e Litérarios, 65,* 5–29.

Brende, B. (2019). *We need a reskilling revolution. Here's how to make it happen.* Available at https://www.weforum.org/agenda/2019/04/skills-jobs-investing-in-people-inclusive-growth.

Brüning, C. I., & Purrmann, M. S. (2014). CLIL pedagogy in Europe: CLIL teacher education in Germany. *Utrecht Studies in Language and Communication, 27,* 315–338.

Bruton, A. (2013). CLIL: Some of the reasons why… and why not. *System, 41*(3): 587–897. https://doi.org/10.1016/j.system.2013.07.001.

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.

Bruton, A. (2019). Questions about CLIL which are unfortunately still not outdated: A reply to Pérez-Cañado. *Applied Linguistics Review, 10*(4), 591–602.

Cabezas Cabello, J. M. (2010). A SWOT analysis of the Andalusian Plurilingualism Promotion Plan (APPP). In *Proceedings of the 23rd GRETA Convention*, ed. M. L. Pérez Cañado, 83–91. Jaén: Joxman.

Council of Europe. (2018). *Common European framework of reference for languages: Learning, teaching, assessment. Companion volume with new descriptors*. Strasbourg: Council of Europe.

Coyle, D. (2007). Content and Language Integrated Learning: Towards a connected research agenda for CLIL pedagogies. *International Journal of Bilingual Education and Bilingualism, 10*(5), 543–562.

Cummins, J. (1999). *BICS and CALP: Clarifying the distinction (Report No. ED438551).* Washington, DC: ERIC Clearinghouse on Languages and Linguistics.

Czura, A., Papaja, K., & Urbaniak, M. (2009). Bilingual education and the emergence of CLIL in Poland. In D. Marsh, P. Mehisto, D. Wolff, R. Aliaga, T. Asikainen, M. J. Frigols-Martín, S. Hughes, & G. Langé (Eds.), *CLIL practice: Perspectives from the field* (pp. 172–178). Jyväskylä: University of Jyväskylä.

Dalton-Puffer, C. (2013). A construct of cognitive discourse functions for conceptualising content-language integration in CLIL and multilingual education. *European Journal of Applied Linguistics, 1*(2), 1–38. https://doi.org/10.1515/eujal-2013-0011.

Davies, K. S., & Patsko, L. (2013). *How to teach English as a lingua franca (ELF).* Available at https://www.britishcouncil.org/voices-magazine/how-teach-english-lingua-franca-elf.

Dewey, M., & Patsko, L. (2017). ELF and teacher education. In J. Jenkins, W. Baker, & M. Dewey (Eds.), *The Routledge handbook of English as a lingua franca* (pp. 1–18). London: Routledge.

Durán-Martínez, R., & Beltrán-Llavador, F. (2016). A regional assessment of bilingual programmes in primary and secondary schools: The teachers' views. *Porta Linguarum, 25,* 79–92.

Gardner, H. (1993). *Multiple intelligences: The theory in practice.* New York: Basic Books.

Graddol, D. (2006). *English next: Why global English may mean the end of English as a foreign language.* London: British Council.

Grazzi, E., & Lopriore, L. (2020). ELF awareness for teacher education in Italy: Attitudes and actions. *Estudos Linguísticos e Litérarios, 65,* 69–89.

Hansen-Pauly, M. A., Bentner, G., Jovanovic, V., Zerbato, D., Llinares, A., Dafouz, E., Erades, C., García, M., Carreras, A., Favlli, F., Dale, L., Robinson, P., Hofmannova, M., & Novotna, J. (2009). *Teacher education for CLIL across contexts: From scaffolding framework to teacher portfolio for content and language integrated learning.* Available at http://clil.uni.lu/CLIL/Project_files/CLIL_Project_Final_Report.pdf.

Hewings, M. (2004). *Pronunciation practice activities: A resource book for teaching English pronunciation.* Cambridge: Cambridge University Press.

Hüttner, J., & Smit, U. (2014). CLIL (Content and Language Integrated Learning): The bigger picture. A response to A. Bruton. 2013. CLIL: Some of the reasons why … and why not. *System, 44*(2), 160–167.

Jenkins, J. (2000). *The phonology of English as an international language.* Oxford: Oxford University Press.

Julius, S., & Madrid, D. (2017). Diversity of students in bilingual university programs: A case study. *The International Journal of Diversity in Education, 17*(2), 17–28.

Kachru, B. B. (Ed.). (1992). *The other tongue, English across cultures.* Urbana, IL: University of Illinois Press.

Kohn, K. (2015). A pedagogical space for ELF in the English classroom. In Y. Bayyurt & S. Akcan (Eds.), *Current perspectives on pedagogy for English as a lingua franca* (pp. 51–67). Berlin: Mouton de Gruyter.

Kramsch, C. (1993). *Context and culture in language teaching.* Oxford: Oxford University Press.

Lancaster, N. K. (2018). Extramural exposure and language attainment: The examination of input-related variables in CLIL programs. *Porta Linguarum, 29,* 91–114.

Lasagabaster, D., & Sierra, J. M. (2010). Immersion and CLIL in English: More differences than similarities. *ELT Journal, 64*(4), 367–375.

Lorenzo, F. (2007). The sociolinguistics of CLIL: Language planning and language change in 21st century Europe. *RESLA Vol. Extra, 1,* 27–38.

Lorenzo, F., Casal, S., & Moore, P. (2009). The effects of content and language integrated learning in European education: Key findings from the Andalusian bilingual sections evaluation project. *Applied Linguistics, 31*(3), 418–442.

Lorenzo, F., Casal, S., Moore, P., & Afonso, Y. M. (2009). *Bilingüismo y educación. Situación de la red de centros bilingües en Andalucía.* Sevilla: Fundación Centro de Estudios Andaluces.

Lorenzo, F., Trujillo, F., & Vez, J. M. (2011). *Educación bilingüe: integración de contenidos y segundas lenguas.* Madrid: Síntesis.

Lyster, R. (2007). *Learning and teaching languages through content: A counterbalanced approach.* Amsterdam: John Benjamins.

Madrid, D., & Barrios, M. E. (2018). A comparison of students' educational achievement across programmes and school types with and without CLIL provision. *Porta Linguarum, 29,* 29–50.

Madrid Manrique, M., & Madrid Fernández, D. (2014). *La formación inicial del profesorado para la educación bilingüe.* Granada: Universidad de Granada.

Marlina, R. (2014). The pedagogy of English as an international language (EIL): More reflections and dialogues. In R. Marlina & R. A. Giri (Eds.), *The pedagogy of English as an international language: Perspectives from scholars, teachers and students* (pp. 1–22). Amsterdam: Springer.

Marsh, D., & Langé, G. (Eds.). (2000). *Using languages to learn and learning to use languages.* Finland: University of Jyväskylä.

Marsh, D. (Ed.). (2002). *CLIL/EMILE. The European dimension: Actions, trends, and foresight potential*. Jyväskylä: University of Jyväskylä.

Marsh, D., Mehisto, P., Wolff, D., & Frigols Martín, M. J. (2010). *European framework for CLIL teacher education.* Graz: European Center for Modern Languages.

Mearns, T. L. (2012). Using CLIL to enhance pupils' experience of learning and raise attainment in German and health education: A teacher research project. *The Language Learning Journal, 40*(2), 175–192.

Mehisto, P. (2008). CLIL counterweights: Recognising and decreasing disjuncture in CLIL. *International CLIL Research Journal, 1*(1), 93–119.

Meyer, O., Coyle, D., Halbach, A., Schuck, K., & Ting, T. (2015). A pluriliteracies approach to content and language integrated learning-mapping learner progressions in knowledge construction and meaning making. *Language, Culture and Curriculum, 28*(1), 41–57.

Moraleja Novillo, S. (2018). *Educar el talento: Guía para desarrollar nuevas habilidades en jóvenes y niños.* Barcelona: Amat Editorial.

Otto, A. (2018). Assessing language in CLIL: A review of the literature towards a functional model. *Latin American Journal of Content & Language Integrated Learning, 11*(2), 308–325.

Paran, A. (2013). Content and language integrated learning: Panacea or policy borrowing myth? *Applied Linguistics Review, 4*(2), 317–342.

Pavón Vázquez, V. (2018). Learning outcomes in CLIL programmes: A comparison of results between urban and rural environments. *Porta Linguarum, 29,* 9–28.

Pedrazzini, L., & Nava, A. (2010). The ELF of English language teachers. In C. Gagliardi & A. Maley (Eds.), *EIL, ELF, global English: Teaching and learning issues* (pp. 283–300). Bern: Peter Lang.

Pérez Cañado, M. L. (2012). CLIL research in Europe: Past, present, and future. *International Journal of Bilingual Education and Bilingualism, 15*(3), 315–341.
Pérez Cañado, M. L. (2015). Training teachers for plurilingual education: A Spanish case study. In D. Marsh, M. L. Pérez Cañado, & J. Ráez Padilla (Eds.), *CLIL in action: Voices from the classroom* (pp. 165–187). Newcastle upon Tyne: Cambridge Scholars Publishing.
Pérez Cañado, M. 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.
Pérez Cañado, M. L. (2016b). Teacher training needs for bilingual education: In-service teacher perceptions. *International Journal of Bilingual Education and Bilingualism, 19*(3), 266–295.
Pérez Cañado, M. L. (2016c). Are teachers ready for CLIL? Evidence from a European study. *European Journal of Teacher Education, 39*(2), 202–221.
Pérez Cañado, M. L. (2017). Stopping the "pendulum effect" in CLIL research: Finding the balance between Pollyanna and Scrooge. *Applied Linguistics Review, 8*(1), 79–99.
Pérez Cañado, M. L. (2018a). This issue. *Theory Into Practice, 57*(3), 163–165.
Pérez Cañado, M. L. (2018b). Innovations and challenges in CLIL teacher training. *Theory Into Practice, 57*(3), 1–10.
Pérez Cañado, M. L. (2018c). CLIL and pedagogical innovation: Fact or fiction? *International Journal of Applied Linguistics, 28,* 369–390.
Pérez Cañado, M. L. (2020). CLIL and elitism: Myth or reality? *The Language Learning Journal, 48*(1): 4–17.
Pérez Cañado, M. L., & Ojeda-Pinar, B. (2018). *Communicative classroom language for bilingual education. Teaching "real english" for CLIL*. Frankfurt-am-Main: Peter Lang.
Rajagopalan, K. (2004). The concept of 'world English' and its implications for ELT. *ELT Journal, 58*(2), 111–117.
Rampton, B. (1996). Displacing the "native speaker": Expertise, affiliation, and inheritance. In T. Hedge & N. Whitney (Eds.), *Power, pedagogy and practice* (pp. 9–22). Oxford: Oxford University Press.
Rascón Moreno, D., & Bretones Callejas, C. (2018). Socioeconomic status and its impact on language and content attainment in CLIL contexts. *Porta Linguarum, 29,* 115–136.
Rosling, H. (2018). *Factfulness. Ten reasons we're wrong about the world—And why things are better than you think*. London: Sceptre.
Ruiz Gómez, D. A., & Nieto García, J. M. (2009). Las secciones bilingües en Secundaria y Bachillerato. Marco organizativo. Dificultades y propuestas. In A. Bueno González, J. M. Nieto García, & D. Cobo López (Eds.), *Atención a la diversidad en la enseñanza plurilingüe. I, II y III Jornadas Regionales de Formación del Profesorado* (CD-ROM). Jaén: Delegación Provincial de Educación de Jaén y Universidad de Jaén.
Seidlhofer, B. (2003). A concept of International English and related issues: From 'real English' to 'realistic English'? Available at http://www.coe.int/t/dg4/linguistic/Source/SeidlhoferEN.pdf.
Seidlhofer, B. (2005). English as a lingua franca. *ELT Journal, 59*(4), 339–341.
Seidlhofer, B. (2011). *Understanding English as a lingua franca*. Oxford: Oxford University Press.
Shepherd, E., & Ainsworth, V. (2017). *English impact: An evaluation of English language capability.* Madrid: British Council.
Sifakis, N. C. (2020). Two obstacles to enabling change in ELF-aware teacher education and how to overcome them. *Estudos Linguísticos e Litérarios, 65,* 104–117.
Siqueira, S. (2012). *English language teaching and the place of culture: For a critical intercultural approach.* In D. C. de Lima (Ed.), *Language and its cultural substrate: Perspectives for a globalized world* (pp. 193–216). Campinas: Editora Pontes.
Siqueira, S. (2015). English as a lingua franca and ELT materials: Is the "plastic world" really melting? In Y. Bayyurt & S. Akcan (Eds.), *Current perspectives on pedagogy for English as a lingua franca* (pp. 239–258). Berlin: Mouton de Gruyter.
Siqueira, S. (2017). Intercultural language educators for an intercultural world: Action upon reflection. *Intercultural Education, 28*(4), 390–407.

Siqueira, S. (2020). ELT materials for basic education in Brazil: Is there room for an ELFaware practice? *Estudos Linguísticos e Litérarios, 65,* 118–146.
Sung, C. C. M. (2013). English as a lingua franca and its implications for English language teaching. *JALT Journal, 35*(2), 173–190.
Ur, P. (2011, March 11–13). English as an International Language: What difference does it make? Plenary talk given at the 34th TESOL Spain Convention. Madrid, Spain.
Vettorel, P., & Corrizzatto, S. (2016). Fostering awareness of the pedagogical implications of World Englishes and ELF in teacher education in Italy. *Studies in Second Language Learning and Teaching, 6*(3), 487–511.
Vez, J. M. (2009). Multilingual education in Europe: Policy developments. *Porta Linguarum, 12,* 7–24.
World Economic Forum. (2018). *The future of jobs report*. Geneva: World Economic Forum. http://www3.weforum.org/docs/WEF_Future_of_Jobs_2018.pdf. Accessed 29 April 2019.

Are CLIL Settings More Conducive to the Acquisition of Digital Competences? A Comparative Study in Primary Education

Esther Nieto Moreno de Diezmas

Abstract In contrast to most studies on Content and Language Integrated Learning (CLIL) focussed on language learning outcomes, this chapter is devoted to exploring a different area: the impact of CLIL on the acquisition of digital competence, one of the eight key competences for lifelong learning. In two consecutive years, CLIL and non-CLIL 9-and 10-year olds took two tests in which the two dimensions of digital competence described in the Recommendation of the European Parliament and the Commission on Key Competences for lifelong learning were evaluated: 'communicate and participate in collaborative networks' and 'search, collect and process digital information'. The participants were 1,967 CLIL and 18,303 non-CLIL students for the first test and 2,813 CLIL and 16,518 non-CLIL students for the second one. Results showed that CLIL students scored significantly higher for both dimensions of digital competence and in nine out of the sixteen learning standards assessed. These findings suggest CLIL is more conducive to digital competence acquisition and seem to indicate ICTs are more integrated in CLIL than in non-CLIL settings, possibly as a way to compensate the added difficulty of conveying meaning through a foreign language, and as a consequence of the teaching innovations CLIL entails.

1 Introduction

In an educational panorama in which effective learning of foreign languages is becoming a must to meet the requirements of a globalised knowledge-based society, CLIL (Content and Language Integrated Learning) is gaining more and more momentum and popularity (Doiz & Lasagabaster, 2017; Lasagabaster & López Beloqui, 2015), since it is deemed to be both 'a lever for change and success in language learning' (Pérez Cañado & Ráez Padilla, 2015, p. 1) and 'an innovative form of education' (Cenoz et al., 2013, p. 16).

E. Nieto Moreno de Diezmas (✉)
University of Castilla-La Mancha, Ciudad Real, Spain
e-mail: Esther.Nieto@uclm.es

M. L. Pérez Cañado (eds.), *Content and Language Integrated Learning in Monolingual Settings*, Multilingual Education 38,
https://doi.org/10.1007/978-3-030-68329-0_4

In the last two decades, research on CLIL has mainly focussed on determining the impact it has on language proficiency and, as Lasagabaster and López Beloqui (2015) and Sierra et al. (2011) have highlighted, only to a much lesser extent have other issues been explored, such as content acquisition, effects on the mother tongue, and motivation, mainly.

However, lately, attention has been drawn to a field that has so far been overlooked: the inclusion and acquisition of key competences in CLIL settings. This line of research was put on the CLIL agenda in the policy workshop convened by the British Council in March 2014 in Como (Italy) with the support of European ministers of education and sector specialists, under the rubric 'CLIL policy and practice: Competence-based Education for employability, mobility and growth' (British Council, 2014). As a result of this policy workshop, a document was published promoting the active inclusion of a key-competence-based approach in CLIL to further develop the integrated curriculum. This document was a milestone for the integration of a competence-based approach in CLIL environments, since there is a lack of methodological studies on how to implement the eight key competences in CLIL settings, and in addition, more evidence is needed on the contribution of CLIL to their acquisition.

This is not a minor issue, since one of the most important challenges in today's education is the implementation of competence-based approaches which enable the acquisition of so-called twenty-first-century skills (cf. also Chapter 3 in this volume). This new educational scenario implies a shift from a curricular design oriented to the assimilation of content and memorisation to a curriculum based on the development of competences. This turning point concerns CLIL too, which cannot remain oblivious to these changes.

One of the eight key competences for lifelong learning is digital competence. This competence is considered to be crucial, as it provides students with the necessary tools to interact in an interconnected world and to achieve personal, social, academic and professional fulfilment. However, the potential of combining CLIL and ICT has hardly been studied; only some experiences have been described, and some methodological guidelines to implement ICT in CLIL settings have been reported. In fact, there is a gap in experimental studies on how CLIL methodology may impact on the development of digital competence and, therefore, more research is needed to further look into how the acquisition of digital competence is being carried out in CLIL settings.

2 Theoretical Framework

2.1 The Eight Key Competences and CLIL

It is commonly agreed that traditional approaches based on memorisation, in which education is conceived as a process of transmission of knowledge and students are deemed to be mere containers of information, are no longer satisfactory as they cannot respond to the demands of a constantly evolving world, in which new discoveries call assumed knowledge into question and technological innovations modify education priorities and affect the list of what has to be taught and learnt. In this new context, education is expected to provide students with adaptable and transferable skills for academic, social and personal success, and to equip them with the necessary competences to mobilise their knowledge, skills and attitudes to act and solve real-life situations and to learn throughout life.

Therefore, the Recommendation of 18 December 2006 of the European Parliament and the Council on Key Competences for Lifelong Learning (European Commission, 2006) was published with the main objective of providing a reference framework in order to establish commonly agreed objectives for education; to identify and define the eight key competences necessary for personal fulfilment, active citizenship, social cohesion and employability in a knowledge society; and to recommend their inclusion in the educational systems of all member countries, so that compulsory education could guarantee all citizens are able to develop these key competences throughout their lives.

The eight key competences acknowledged in the aforementioned documents as basic acquisitions for all citizens were: communication in the mother tongue, communication in foreign languages, mathematical competence and basic competences in science and technology, digital competence, learning to learn, social and civic competences, sense of initiative and entrepreneurship, and cultural awareness and expression.

This competence-based approach is underpinned from a sociological, psychological and epistemological perspective (Escamilla, 2008). From a sociological point of view, changes in the demands of a knowledge society require that knowledge not be restricted to the academic environment (Coll, 2007; Pérez Gómez, 2007), but, rather, can be transferred to real situations in social, family and work life. From a psychological point of view, key competencies collaborate for comprehensive development of knowledge, skills and values, and from an epistemological point of view, this approach breaks with the impenetrability inherent to disciplinary teaching in order to create a dynamic knowledge space common to different subjects.

Therefore, the key competences are multifunctional, transferable, inclusive, comprehensive, dynamic and cross-curricular. Key competences are multifunctional because they represent know-how applicable to different contexts. They are also transferable to different situations by means of generalisation, thus representing adaptive knowledge (Pérez Gómez, 2007). They are inclusive, since they include

combinations of three different kinds of components: knowledge, skills and attitudes. In other words, being competent means 'the proven ability to use knowledge, skills and personal, social and/or methodological abilities' (European Commission, 2011, p. 12). This implies that competences are not only solidly based on theoretical knowledge, but are also inspired by principles and values, and are action-oriented (Escamilla, 2008). Another feature of the key competences is their comprehensive nature, since they ensure that all citizens can acquire them at the end of compulsory education. Finally, key competences are cross-curricular, which means there is not a univocal correspondence between competences and school subjects, but, rather, all subjects, including CLIL subjects, should contribute to the development of all key competences and each competence should be acquired by means of all subjects.

The second key competence acknowledged in the European Recommendation of 2006 is communication in foreign languages, the acquisition of which is one of the main rationales behind the implementation of CLIL. Furthermore, in light of the aforementioned European Recommendation, the goal of compulsory education and, consequently, of any educational approach, including CLIL, should be the integrated development of all eight key competences. This may clash with the commonly known definition of CLIL as being 'a dual-focussed educational approach' (Mehisto et al., 2008, p. 9), since CLIL has to focus on the integrated learning of not only language and content, and not just on communication, content, cognition and culture, as in the four C's framework (Coyle, 1999), but, rather on all eight key competences, including digital competence, learning to learn, social and civic competences, and cultural awareness and expression, among others.

If integration of language, content, culture and cognition in CLIL classrooms is a challenge already, then, the introduction of the eight key competences might seem a requirement beyond the capabilities of a CLIL teacher. Including the key competences in teaching practice is a difficult task, and most teachers have not received special training for developing cross-curricular competences. However, CLIL 'is cross-curricular in nature. It teaches some cross-cutting competences explicitly' (Clegg, 2014, p. 84), so that CLIL settings seem to be able to provide a more suitable scenario for the implementation of a competence-based approach. Thus, CLIL methodology is closer to competence-based approaches, as both 'tend towards a more process-oriented view of education, a more hands-on approach' (Ball, 2014, p. 76). Additionally, the tenets of a key competence-based education (i.e. personal fulfillment, social inclusion, and employability) are 'precisely at the heart of what we call CLIL' (Mittendorfer, 2014, p. 96) and, thus, compared to traditional methodologies, 'CLIL is well-placed to develop many of them' (Clegg, 2014, p. 83) and displays 'a pedagogy which raises some competences to the surface of classroom discourse' (Clegg, 2014, p. 84).

This suggests that CLIL methodology may potentially foster the development of key competences, as has been evinced for the acquisition of emotional competence (Nieto Moreno de Diezmas, 2012), learning to learn (Nieto Moreno de Diezmas, 2016), and ICT competence (Nieto Moreno de Diezmas, 2018). All these studies were carried out in Secondary Education settings and conclude that CLIL methodology provides a significant advantage in the development of emotional, learning

to learn and ICT competences. It is time to explore the connection between CLIL methodology and the acquisition of other key competences, particularly in Primary Education, since there is a lack of information about their acquisition in CLIL settings at this education level. The key competence targeted in this study is digital competence, because it is one of the most cross-curricular competences, it is considered one the most important competences to succeed in the knowledge society, and it is an instrument for the acquisition of the rest of key competences.

2.2 Digital Competence and CLIL

Digital competence is the fourth key competence acknowledged by the European Recommendation on 18 December 2006. This Recommendation emphasises the role of digital competence and puts it on a par with basic language skills, literacy and numeracy as regards its contribution to learning, thus considering it as 'an essential foundation for learning' (European Commission, 2006). Furthermore, digital competence is deemed to be essential to function in a knowledge society (Eshet-Alkalai, 2004), and enables full participation in life, by improving not only employability, but also social and personal development.

The European Recommendation on 18 December 2006 defined two core dimensions of digital competence: the use of computers to retrieve, assess, store, produce, present and exchange information (i.e. informational competence), and the use of new technologies to communicate and participate in collaborative networks via the Internet (i.e. communicative competence).

At present, most research that connects CLIL and ICT explores methodological proposals (such as the Content and Language Processing Sequence—CLPS—by Fernández Fontecha, 2012), or describes online platforms and resources for introducing new technologies into a CLIL classroom, and most studies in the field are set in higher education. Gimeno Sanz (2009) and Gimeno et al. (2010) provide pedagogical guidelines on how ICT can be integrated effectively into a CLIL study programme, and present online resources and platforms such as INGENIO to be implemented in CLIL settings, particularly in higher education. In the same vein, Fürstenberg and Kletzenbauer (2013) examined how to use online resources effectively to enhance learning in tertiary CLIL, and Gaballo (2010) showed different ICT-based activities aimed at acquiring both professional expertise and linguistic competence in an introductory level journalism course at university. Some experiences about the implementation of virtual learning environments in CLIL settings have been also reported (Paggiaro, 2013; Paliwoda-Pękosz & Stal, 2015).

The design of ICT materials connected to task-based learning projects has also been studied (Fernández Yubero & Pareja Moreno, 2009; Opp-Beckman, 2002; Reinhardt & Isbell, 2002), by means of webquests (Vlachos, 2009), CLILquests (Fernández Fontecha, 2012) and other resources linked to the concept of gamification, via programmes such as JClic and Atenex (Durán & Cruz, 2013), Hot Potatoes, Scratch and What2Learn (Rodríguez et al., 2014).

Although most studies have been conducted in Tertiary Education, there have also been European Projects promoting the use of web 2.0 tools in pre-Primary, Primary and lower Secondary Education, to enhance language learning in CLIL settings, such as the eCLIL4You project (Rampona, 2013). In addition, Gilleran (2012) encouraged the implementation of Etwinning projects in CLIL environments, and Prentza (2013) and Nieto Moreno de Diezmas and Ortiz Calero (2017) described Etwinning experiences in Primary and pre-Primary CLIL, respectively.

Although this seems to show a mixed picture, research on new technologies and CLIL is restricted to enunciations of methodological principles and descriptions of specific experiences, programmes, platforms or digital resources. As a result, there is no research that evaluates the contribution CLIL makes towards the acquisition of digital competence. In an attempt to bridge this gap, the aim of this paper is to study how CLIL affects the development of digital competence of 9-and 10-year-old Primary School students.

3 Research Questions

This study aims to answer the following questions:

1. How does CLIL contribute towards the acquisition of digital competence in Primary Education?
2. Are there differences between CLIL and non-CLIL students in the acquisition of the dimensions that make up digital competence, i.e. 'communicate and participate in collaborative networks', and 'search, collect and process digital information'?
3. Does CLIL influence the acquisition of particular learning standards of digital competence?

4 Method

4.1 Setting

This research was carried out in Castilla-La Mancha, a monolingual autonomous community located in the centre of Spain, where CLIL programmes were launched in 2005 under the denomination of European Sections.

Additions of schools to the bilingual programme are managed by the local educational authority by means of official calls. To be included in the CLIL programme, schools have to present a project approved by the teaching staff, showing they meet all the requirements and, in particular, that they have enough teachers who can certify they have a B2 CEFR foreign language level to implement the integrated curriculum. At the time of this study, bilingual schools had to provide between 50 and 100% of

the lessons of at least two subjects in a foreign language. The most popular CLIL subjects were Natural Sciences, Social Sciences and Art, although schools were free to choose their CLIL subjects depending on their availability of teachers with the required level of foreign language.

One of the main principles of the bilingual programme is that it aims to be egalitarian, inclusive and non-selective, and therefore the access of students to bilingual schools is organised by applying the annual general admission rules common to all schools bilingual or not (proximity to family home, number of siblings already enrolled in the school, or low income, among others). It must be stressed that any selection based on academic or linguistic performance is explicitly forbidden.

The bilingual programme of Castilla-La Mancha specifically promotes the implementation of CLIL methodology in the classroom and the use of the foreign language in the whole school establishment. It also encourages the organisation of complementary activities connected to the target language and the participation in European programmes such as Erasmus + and exchange programmes. eTwinning projects, which, precisely, are closely linked to online collaboration and communication, are specifically fostered.

4.2 Participants

The data analysed in this study was collected during a Diagnostic Assessment of the Educational System of Castilla-La Mancha. The study involved all schools in the autonomous community and the census of students enrolled in the 4th year of Primary Education (9–10-year olds). Two tests evaluating digital competence were conducted over a two-year period. The control group (non-CLIL students) was made up of 18,303 students for the first test and 16,518 for the second one, and the experimental group (CLIL students) was comprised of 1,967 for the first test, and 2,813 for the second one.

4.3 Instruments

Following the aforementioned definition of digital competence provided by the European Commission (2006), two dimensions were analysed. The first year of the study, a test entitled 'Virtual classroom', was conducted to assess the dimension 'communicate and participate in collaborative networks', while the second year, the dimension evaluated was 'search, collect and process digital information' by means of the test 'My school'.

Competences integrate knowledge, skills and attitudes, but as Ball (2014, p. 78) points out, 'a student might possess all three but still be "incompetent". In order to demonstrate competence, a student had to perform', and to do so, 'competence needs a situation (as authentic as possible)' (Ball, 2014, p. 78). In keeping with Ball's ideas

Table 1 Assessed dimensions and learning standards for digital competence

Dimension 1. Communicate and participate in collaborative networks. 'Virtual classroom'	Dimension 2. Search, collect and process digital information.'My school'
Respect the rules of participation in virtual networks	Use the Internet as a source of information
Use Internet as a source of information	Create a backup copy
Understand the risks of sharing personal data	Print the document
Send e-mails	Create folders and files
Handle network communication tools	Edit with word processor
Manage files and folders	Select information critically
	Edit images
	Manage folders
	Copy the file to share

on competence assessment, the tests were made up of two main parts: a realistic scenario, which provided an authentic situation for the mobilisation of knowledge, skills and attitudes, and a system of tasks connected to a set of learning standards. The 'Virtual classroom' scenario was the use of a blog to organise a collaborative project for the Social Sciences subject. During testing, students had to use communication tools, make comments on the blog, send e-mails, use the Internet and manage files and folders. In turn, the scenario of 'My school' was a work project in which students had to research the name of their school, search the net, edit information and carry out other actions, such as create a list of favourites, print a document and make a backup copy. The completion of these tasks was awarded with a maximum of one, two, or three points depending on the number of activities students had to carry out to accomplish them. Regarding the two tests, they contained three one-point tasks, eleven two-point tasks and two three-point tasks. Every task was connected to the acquisition of one learning standard. In Table 1, the learning standards of every dimension are shown.

4.4 Procedure

The tests were carried out in the computer lab of schools, where each computer came equipped with a word processor, image editor, web browser, Internet and printer connections. All information in the tests was written in Spanish, the mother tongue of the students, who had 60 min to complete the tasks for each test.

4.5 Data Analysis

To analyse data, the Statistical Package for Social Science, SPSS, was used. Cronbach alpha for the test 'My school' was 0.834, and for the 'Virtual classroom', 0.725, which meant that the internal consistency and reliability of the tests were high. The distribution was normal (Kolmogorov–Smirnov Test), and t-tests (to compare independent samples and determine if differences between CLIL and non-CLIL students were significant) were run.

5 Results

5.1 Overall Results in Digital Competence

Results showed that CLIL students performed significantly higher in both dimensions of digital competence: 'communicate and participate in collaborative networks' (dimension 1) and 'search, collect and process digital information' (dimension 2). For both CLIL and non-CLIL groups, higher scores were recorded in communicating and participating in collaborative networks than in searching, collecting and processing digital information (Table 2).

5.2 Learning Standards of Dimension 1: Communicate and Participate in Collaborative Networks

Regarding the first dimension assessed, 'communicate and participate in collaborative networks', as shown in Table 3, CLIL students scored higher in all its learning standards, and differences were statistically significant in 'send e-mails' ($p = 0.000$), 'handle network communication tools' ($p = 0.001$) and 'manage files and folders' ($p = 0.000$).

Table 2 Results of the dimensions of the digital competence

Dimension	Group	Mean	Std. deviation	Std. error mean	Sig. (2-tailed)
Dimension 1 Digital information	CLIL	.6987	.21174	.00495	.000
	NON-CLIL	.6739	.22409	.00169	.000
Dimension 2 Digital communication	CLIL	.6189	.22175	.00440	.000
	NON-CLIL	.5980	.23333	.00191	.000

Table 3 Results of the learning standards of the dimension 'communicate and participate in collaborative networks'

Learning standards	Group	Mean	Std. deviation	Std. error mean	Sig. (2-tailed)
Respect the rules of participation	CLIL	.96	.762	.017	.145
in virtual networks	NON-CLIL	.93	.765	.006	.143
Use the Internet as a source of information	CLIL	1.67	.658	.015	.305
	NON-CLIL	1.65	.677	.005	.294
Understand the risks of sharing personal data	CLIL	1.40	.823	.018	.477
	NON-CLIL	1.39	.837	.006	.471
Send e-mails	CLIL	1.83	.445	.010	.000
	NON-CLIL	1.79	.478	.003	.000
Handle network communication tools	CLIL	1.62	1.120	.026	.001
	NON-CLIL	1.54	1.100	.008	.001
Manage files and folders	CLIL	1.54	.675	.015	.000
	NON-CLIL	1.45	.706	.005	.000

5.3 Learning Standards of Dimension 2: Search, Collect and Process Digital Information

As far as the second dimension of digital competence is concerned, 'search, collect and process digital information', CLIL students outstripped their counterparts in all its standards, except for two: 'create a list of favourites', a standard in which mainstream students recorded higher scores than CLIL students, and the difference was significant; and 'copy a file to share', in which the difference between the two groups was not significant. In contrast, CLIL students significantly outperformed their non-CLIL peers in six standards: 'create a backup copy' ($p = 0.000$), 'print a document' ($p = 0.003$), 'create folders and files' (0.001), 'edit with word processor' ($p = 0.000$), 'select information critically' ($p = 0.000$) and 'manage folders' ($p = 0.003$), as shown in Table 4. CLIL students scored higher but the difference was not significant in 'use the Internet as a source of information' ($p = 0.171$).

Table 4 Results of the learning standards of the dimension 'search, collect and process digital information'

Standard	Group	Mean	Std. Deviation	Std. Error Mean	Sig. (2-tailed)
Use Internet as a source	CLIL	2.15	1.073	.020	.171
of information	NON-CLIL	2.12	1.070	.008	.172
Create a backup copy	CLIL	.56	.537	.010	.000
	NON-CLIL	.51	.538	.004	.000
Print a document	CLIL	1.22	.747	.014	.003
	NON-CLIL	1.18	.796	.006	.002
Create folders and files	CLIL	1.57	.678	.012	.001
	NON-CLIL	1.53	.696	.005	.001
Edit with word processor	CLIL	1.35	.807	.015	.000
	NON-CLIL	1.25	.827	.006	.000
Select information critically	CLIL	.71	.548	.010	.000
	NON-CLIL	.66	.532	.004	.000
Edit images	CLIL	1.17	.856	.016	.180
	NON-CLIL	1.15	.833	.006	.188
Manage folders	CLIL	1.58	.686	.013	.003
	NON-CLIL	1.54	.704	.005	.003
Copy the file to share	CLIL	.95	.901	.017	.315
	NON-CLIL	.97	.892	.007	.319
Create a list of favourites	CLIL	.40	.491	.010	.000
	NON-CLIL	.45	.498	.004	.000

6 Discussion

The results of this study show the following main findings: first, CLIL significantly contributes to the acquisition of digital competence; second, CLIL positively impacts both dimensions of digital competence; and third, there are affected and unaffected aspects regarding development of digital competence by CLIL methodology, as had already been observed in the research on the acquisition of areas and skills in the foreign language (Dalton-Puffer, 2011; Pérez Cañado, 2012; Ruiz de Zarobe, 2011).

6.1 CLIL Significantly Contributes to the Acquisition of Digital Competence

Results indicated CLIL methodology had a positive effect on the acquisition of digital competence for 9–10-year olds enrolled in the 4th year of Primary School Education. These outcomes bring out a new benefit CLIL has: its contribution to the development of digital competence. They are also in keeping with the literature which claims that 'CLIL may indirectly help create favourable conditions for ICT integration' (Fernández Fontecha, 2012, p. 320), for the implementation of a competence-based approach (Ball, 2014; Clegg, 2014; Mittendorfer, 2014) and for the acquisition of transferable cross-curricular competences (Ball, 2014; Clegg, 2014; Mittendorfer, 2014).

These findings can be explained, among other reasons, because the difficulty CLIL teachers have to convey meaning when the means of instruction is a foreign language, demands an extra effort which 'inevitably leads to a widening of their teaching repertoires, and to a heightening of their methodological awareness' (Ball, 2014, p. 77). Thus, CLIL seems to act as a catalyst for educational change, due to its potential for including more holistic perspectives, connected to competence-based approaches, and particularly, to successful practice of implementation of new technologies in the classroom.

6.2 CLIL Positively Impacts Both Dimensions of Digital Competence

CLIL students significantly outstripped their non-CLIL counterparts in both dimensions of digital competence: 'communicate and participate in collaborative networks', and 'search, collect and process digital information'. Scores for both groups were higher in 'communicating and participating in collaborative networks' than in 'searching, collecting and processing digital information'. This might have been due not so much to the students using this information for academic purposes, but rather, to the presence of new technologies for social communication in their everyday lives, as suggested in previous studies that explored how 10-and11-year olds used technology for communication purposes both at school and outside it (Nieto Moreno de Diezmas & Dondarza Manzano, 2016).

Significantly higher scores of CLIL students in 'communicate and participate in collaborative networks' were probably due to the fact that CLIL methodology is based on active participation, and creates student-centred learning environments in which communication, collaboration and cooperative learning play an essential role (Ting, 2014, p. 105), and as a result, it may be concluded that CLIL provides a setting more conducive to the introduction of digital tools to foster communication and participation. In this vein, Gimeno Sanz (2009) underlines the potential CLIL has for the implementation of new technologies, claiming that 'technology in education

is better exploited when an environment favouring student participation is developed' (Gimeno Sanz, 2009, p. 80).

In turn, the positive impact of CLIL on the dimension 'search, collect and process digital information' seems to indicate that the use of new technologies as learning tools to access and express knowledge is more generalised in CLIL than in non-CLIL settings, and, in this sense, CLIL seems to help successfully integrate new technologies in the learning process.

6.3 There are Affected and Unaffected Aspects of CLIL Regarding Development of Digital Competence

Regarding particular digital skills, CLIL students scored higher in 14 of the 16 learning standards used to assess digital competence, and significant differences favoured the CLIL group in 9 of them: 'send e-mails', 'handle network communication tools', 'manage files and folders', 'create a backup copy', 'print a document', 'create folders and files', 'edit with word processor', 'select information critically' and 'manage folders'. Differences between the CLIL and non-CLIL groups were not significant in 'respect the rules of participation in virtual networks', 'use the Internet as a source of information', 'understand the risks of sharing personal data', 'edit images' and 'copy a file to share'. Although CLIL students scored higher in nearly all the learning standards assessed and significantly higher in most of them, they recorded significantly lower scores in 'create a list of favourites'.

These results seem to indicate that CLIL positively affects the development of particular digital skills, whereas there are other areas unaffected by the type of methodology whether this be CLIL/non-CLIL. For example, CLIL students proved to have significantly better skills for sending e-mails and handling network communication tools, which are abilities connected to communication and participation, which, in turn, are key concepts in CLIL methodology. However, CLIL and non-CLIL students had a similar understanding of the risks of sharing personal data, possibly as a result of increasing social concern on Internet safety that could have had implications for the students' and teachers' awareness of this issue, regardless of the type of instruction.

As for the use of the Internet as a source of information, there were also not any significant differences between the CLIL and non-CLIL students, and the results in this learning standard were consistent in both stages of the study, since it was evaluated twice: the first year as part of the dimension 'communicate and participate in collaborative networks', and the second year, as a component of the second dimension of digital competence 'search, collect and process digital information'. In both cases, no significant differences were found. This indicates that both CLIL and non-CLIL programmes made similar contributions in the acquisition of this skill, maybe because students usually carry out this kind of activity outside school, i.e. searching the net with different aims (academic, leisure…) and this may have

levelled differences between the groups. However, CLIL did significantly impact the critical selection of online information, which meant that CLIL students showed more awareness of the need to contrast online resources. This ability is deemed to be a higher order thinking skill, since it entails critical evaluation of the reliability of online sources, and one of the benefits typically purported of CLIL is precisely its potential to promote critical thinking (Coyle et al., 2010; Quiroga, 2013). Accordingly, the results of the CLIL students seem to confirm that these cognitive strategies of evaluation and critical thinking can be applied to different contexts, including digital environments.

Additionally, results showed CLIL students were more familiar than their counterparts with fundamental computer applications, such as word processing, creating and handling files and folders, print a document and create a backup copy, which might suggest that the use of new technologies to process, manage and store information is more integrated in the classroom routines in CLIL than in non-CLIL settings. However, there were no significant differences between both groups in editing images, which might have been due to recreational uses related to the image editing that students carry out in other settings, so that both groups acquired similar proficiency in this skill.

All in all, it can be concluded that CLIL students in the 4th year of Primary School education showed better acquisition of digital competence than their non-CLIL peers, and this fact cannot be explained by claiming that students were selected to participate in CLIL programmes. In a couple of studies, Bruton (2011a, b) shed doubt on the promising results recorded for CLIL students in research studies on second language proficiency, mainly arguing that the admission of students to CLIL programmes under research was selective. However, this was not the case for the students enrolled in the CLIL programmes of Castilla-La Mancha, especially in Primary School Education, since the local educational authority explicitly forbids any entry exam or any selection based on academic or linguistic grounds. What is more, all students in Primary Schools that are implementing CLIL participate in the programme, which means there are not different classes for students that follow the programme or not, as this CLIL programme is inclusive and encompasses all students enrolled in a CLIL school, regardless of their academic or linguistic performance.

7 Conclusions

In this study, the digital competence of CLIL and non-CLIL 9- and 10-year olds enrolled in the 4th year of Primary Education was assessed. CLIL students obtained significantly higher scores in both dimensions of digital competence described by the European Commission (2006): 'communicate and participate in collaborative networks' and 'search, collect and process digital information'. CLIL students scored higher in all the learning standards for both dimensions except two, and differences were significant in nine of the sixteen skills evaluated.

The results seem to support the hypothesis put forward by several authors (Ball, 2014; Ball & Kelly, 2014; Clegg, 2014, among others), according to which CLIL provides learning environments more suitable for the integration not only of content and language, but also of all key competences (cf. also Chapter 2 in this volume), including digital competence. In this vein, the results of this study suggest CLIL students had more practice with digital devices, in light of their significantly better skills in the management of the main digital applications and operations, such as creating and saving files and folders, creating a backup copy, printing a document and editing with word processors, for example, which implies ICT is more integrated in CLIL settings. Additionally, these outcomes may be due to the fact that some of the features of CLIL methodology are conducive to the acquisition of skills that are transferable to digital environments. Thus, CLIL focus on the development of communication abilities, critical thinking, high order skills, participation and collaborative learning is related to some of the learning standards in which CLIL students scored significantly higher, such as 'send e-mails', 'handle network communication tools' and 'select information critically'. However, there were some skills unaffected by CLIL methodology, such as 'use the Internet as a source of information', 'edit images' and 'understand the risks of sharing personal data', for example. Regarding the first two skills mentioned, typical recreational uses of the Internet (for example, to retrieve different kind of information and to edit images for fun) that all students, regardless of their type of instruction, probably carried out outside school settings, could have levelled differences between CLIL and non-CLIL students. The absence of differences in 'understand the risks of sharing personal data' could be explained by the increase of social concern on the issue of Internet safety.

As a final remark, the success of CLIL in developing this crucial competence to interact in a knowledge society in the digital era can have an implication for mainstream education. Some of the ingredients of CLIL methodology could be taken as a model to improve the acquisition of digital skills in non-CLIL settings, so that all students could benefit from a more active, participative, collaborative, student-centred and competence-based approach in which ICTs are more integrated in the teaching–learning process.

Acknowledgements This study is part of the research project ADiBE (Ref.: RTI2018-093390-B-I00), funded by the Ministry of Economy, Industry and Competitiveness (MINECO), 2019–2022. I would also like to thank the Castilla-La Mancha Office of Evaluation for their invaluable help.

References

Ball, P. (2014). CLIL and competences: Assessment. In *CLIL policy and practice: Competence-based education for employability, mobility and growth* (pp. 76–80). British Council.

Ball, P., & Kelly, K. (2014). Competence-based learning and skills for employability and mobility. In *CLIL policy and practice: Competence-based education for employability, mobility and growth* (pp. 81–82). British Council.

British Council. (2014). Regional policy dialogues 2013–2014. https://www.britishcouncil.org/sites/default/files/bc_regionalpolicydialogues_baja.pdf. Accessed 2 Feb 2018.

Bruton, A. (2011). Is CLIL so beneficial, or just selective? Re-Evaluating Some of the Research. *System, 39*(4), 523–532. https://doi.org/10.1016/j.system.2011.08.002.

Bruton, A. (2011). Are the differences between CLIL and non-CLIL groups in Andalusia due to CLIL? A reply to Lorenzo, Casal, and Moore (2010). *Applied Linguistics, 32*(2), 236–241. https://doi.org/10.1093/applin/amr007.

Cenoz, J., Genesee, F., & Gorter, D. (2013). Critical analysis of CLIL: Taking stock and looking forward. *Applied Linguistics, 35*, 243–262. Oxford University Press (OUP). https://doi.org/10.1093/applin/amt011.

Clegg, J. (2014). The role of CLIL in developing language and cognitive skills in the Curriculum. In *CLIL policy and practice: Competence-based education for employability, mobility and growth* (pp. 83–94). British Council.

Coll, C. (2007). Las competencias en la educación escolar: Algo más que una moda y mucho menos que un remedio. *Aula De Innovación Educativa, 161,* 34–39.

Coyle, D. (1999). Theory and planning for effective classrooms: Supporting students in content and language integrated learning contexts. In J. Mash (Ed.), *Learning through a foreign language* (pp. 46–62). London: CILT.

Coyle, D., Hood, P., & Marsh, D. (2010). *CLIL: Content and language integrated learning*. Cambridge: Cambridge University Press.

Dalton-Puffer, C. (2011). Content-and-language integrated learning: From practice to principles? *Annual Review of Applied Linguistics, 31,* 182–204.

Doiz, A., & Lasagabaster, D. (2017). Management teams and teaching staff: Do they share the same beliefs about obligatory CLIL programmes and the use of the L1? *Language and Education, 31*, 93–109. Informa UK Limited.https://doi.org/10.1080/09500782.2017.1290102

Durán, A., & Cruz, B. (2013). How to integrate stories and ICT in content-based units of work for English learning. *Porta Linguarum, 19,* 219–237.

Escamilla, A. (2008). *Las competencias básicas: Claves y propuestas para su desarrollo en los centros*. Barcelona: Graó.

Eshet-Alkalai, Y. (2004). Digital literacy. A conceptual framework for survival skills in the digital era. *Journal of Educational Multimedia & Hypermedia, 13,* 93–106.

European Commission. (2006). Recommendation of 18 December 2006 of the European parliament and the council on key competences for lifelong learning. *Official Journal of the European Union L394*.

European Commission. (2011). *Languages for jobs report—Providing multilingual communication skills for the labour market report*. https://ec.europa.eu/dgs/education_culture/repository/languages/policy/strategic-framework/documents/languages-for-jobs-report_en.pdf. Accessed 2 Feb 2016.

Fernández Fontecha, A. (2012). CLIL in the foreign language classroom: Proposal of a framework for ICT materials designing language-oriented versions of Content and Language Integrated Learning. *Alicante Journal of English Studies, 25,* 317–334.

Fernández Yubero, M. A., & Pareja Moreno, M. I. (2009). CLIL and ICT: Expanding our borders. In E. Dafouz & M. C. Guerrini (Eds.), *CLIL across educational levels* (pp. 45–58). Madrid: Richmond Publishing.

Fürstenberg, U., & Kletzenbauer, P. (2013). CLIL: From online sources to learning resources. In *Proceedings of International Conference, ICT for Language Learning* (5th ed.). Padova: Libreriauniversitaria. it Edizioni.

Gaballo, V. (2010). Integrating content and language in specialized language teaching and learning with the help of ICT. In *Proceedings of International Conference, ICT for Language Learning* (3rd ed.). Padova: Libreriauniversitaria. it Edizioni.

Gilleran, A. (2012). eTwinning and CLIL. *CLIL Magazine, 1,* 9–10.

Gimeno Sanz, A. (2009). *How can CLIL benefit from the integration of information and communications technologies?* In L. Carrió (Ed.), *Content and language integrated learning: Cultural diversity* (pp. 77–102). Series: Linguistic Insights (Vol. 92). Frankfurt: Peter Lang.

Gimeno, A., Seiz, R., de Siquera, J. M., & Martínez, A. (2010). Content and language integrated learning in higher technical education using inGenio online multimedia authoring tool. *Procedia—Social and Behavioral Sciences, 2,* 3170–3174. https://doi.org/10.1016/j.sbspro.2010.03.484.

Lasagabaster, D., & López Beloqui, R. (2015). The impact of type of approach (CLIL versus EFL) and methodology (book-based versus project work) on motivation. *Porta Linguarum, 23,* 41–57.

Mehisto, P., Marsh, D., & Frigols, M. J. (2008). *Uncovering CLIL, content and language integrated learning in bilingual and multilingual education.* Oxford: MacMillan.

Mittendorfer, F. (2014). Competence matters: Signposting the road to learner empowerment. In *CLIL policy and practice: Competence-based education for employability, mobility and growth* (pp. 95–97). British Council.

Nieto Moreno de Diezmas, E. (2012). CLIL and development of emotional competence. *Miscelanea: A Journal of English and American Studies, 45*, 53-74.

Nieto Moreno de Diezmas, E. (2016). The impact of CLIL on the acquisition of the learning to learn competence in secondary school education in the bilingual programmes of Castilla-La Mancha. *Porta Linguarum, 25*, 21–34.

Nieto Moreno de Diezmas, E. (2018). Exploring CLIL contribution towards the acquisition of cross-curricular competences: A comparative study on digital competence development in CLIL. *Revista de lingüística y lengua aplicadas, 13*, 75–85.

Nieto Moreno de Diezmas, E., & Dondarza Manzano, P. (2016). PleS in primary school: The learners' experience in the Piplep project. *Digital Education Review, 29*: 45–61.

Nieto Moreno de Diezmas, E., & Ortiz Calero, C. (2017). CLIL e interculturalidad en educación infantil: Aprendiendo por proyectos. *Multárea, Revista de Didáctica, 9.*

Opp-Beckman, L. (2002). Africa online: A web—and content-based English language teaching course. *TESOL Journal, 11,* 4–8.

Paggiaro, L. (2013). Planning, implementing and evaluating a CLIL online course. In *Proceedings of International Conference, ICT for Language Learning* (6th ed.). Padova: Libreriauniversitaria. it Edizioni.

Paliwoda-Pękosz, G., & Stal, J. (2015). ICT in supporting content and language integrated learning: Experience from Poland. *Information Technology for Development, 21,* 1–23. https://doi.org/10.1080/02681102.2014.1003521.

Pérez Cañado, M. L. (2012). CLIL research in Europe: Past, present, and future. *International Journal of Bilingual Education and Bilingualism, 15*, 315–341. Informa UK Limited. https://doi.org/10.1080/13670050.2011.630064.

Pérez Cañado, M. L., & Ráez Padilla, J. (2015). Introduction and overview. In D. Marsh, M. L. Pérez Cañado, & J. Ráez Padilla (Eds.), *CLIL in action: Voices from the classroom* (pp. 1–12). Newcastle upon Tyne: Cambridge.

Pérez Gómez, A. I. (2007). *La naturaleza de las competencias básicas y sus aplicaciones pedagógicas: Cuadernos de Educación 1.* Santander: Consejería de Educación de Cantabria.

Prentza, A. (2013). CLIL & ICT in English foreign language learning: The eTwinning experience of a primary school of intercultural education. In *Proceedings of International Conference, ICT for Language Learning* (6th ed.). Padova: Libreriauniversitaria. it Edizioni.

Quiroga, M. A. (2013). *The development of critical thinking through CLIL.* Lap Lambert Academic Publishing.

Rampona, S. (2013). eCLIL4You: A European foreign Comenius project to promote language learning through Web 2.0 tools from pre-school to lower secondary education. In *Proceedings of International Conference, ICT for Language Learning* (6th ed.). Padova: Libreriauniversitaria. it Edizioni.

Reinhardt, J., & Isbell, K. (2002). The recycling research documentary: A technology-integrated project. *TESOL, 11,* 41–48.

Rodríguez, R., Blázquez, M., López, B., Castro, M., San Cristobal, E., & Martín, S. (2014). Educational games for improving the teaching-learning process of a CLIL subject: Physics and Chemistry in secondary education. In *Proceedings of IEEE Frontiers in Education Conference (FIE)* (pp. 1–8).

Ruiz de Zarobe, Y. (2011). Which language competencies benefit from CLIL? An insight into applied linguistic research. In Y. Ruiz de Zarobe, J. Sierra, & F. Gallardo del Puerto (Eds.), *Content and foreign language integrated learning* (pp. 129–153). Bern: Peter Lang.

Sierra, J. M., Gallardo del Puerto, F., & Ruiz de Zarobe, Y. (2011). Good practice and future actions on CLIL: Learning and pedagogy. In Y. Ruiz de Zarobe, J. Sierra, & F. Gallardo del Puerto (Eds.), *Content and foreign language integrated learning* (pp. 317–338). Bern: Peter Lang.

Ting, Y. L. T. (2014). A concrete directive towards literacy. In *CLIL policy and practice: Competence-based education for employability, mobility and growth* (pp. 105–113). British Council.

Vlachos, K. (2009). The potential of information communication technologies (ICT) in content and language integrated learning (CLIL): The case of English as a second/foreign language. In D. Marsh, P. Mehisto, D. Wolff, R. Aliaga, T. Asikainen, M. J. Frigols-Martin, S. Hughes, & G. Langé (Eds.), *CLIL practice: Perspectives from the field* (pp. 189–198). University of Jyväskylä.

Investigating the Effects of CLIL on Language Attainment: Instrument Design and Validation

Daniel Madrid Fernández, Antonio Bueno González, and Juan Ráez Padilla

Abstract This chapter presents the English language tests which have been designed and validated in a longitudinal quantitative study in order to determine the effects of Content and Language Integrated Learning (CLIL) on English language competence. The sources used to formulate the tests, the steps followed for their design and the procedures employed for their validation are all expounded on. The ways in which their reliability, content, construct, and face validity have been guaranteed are outlined. Moreover, the statistical operations performed to guarantee them are showcased (Cronbach α, Kuder-Richardson reliability coefficient or item difficulty and discrimination). The actual tests are then presented in a format which can be directly applied in any CLIL classroom in order to determine the effects of CLIL on language competence. We are in dire need of empirically grounded data in this area and it can only ensue from the employment of empirically valid and reliable instruments such as those presented herein. The latter data will be crucial to fine-tune, reengineer or revamp CLIL implementation in order to keep it on track.

1 Introduction

If we want to evaluate the linguistic and communicative competence of second language learners, it is necessary to define what those competences entail and what a learner's proficiency involves in order to design adequate tests which provide suitable data. Once a clear-cut concept of proficiency is adopted, it is important to design tests according to specific objectives, the characteristics of language learners and the

D. Madrid Fernández (✉)
University of Granada, Granada, Spain
e-mail: dmadrid@ugr.es

A. Bueno González · J. Ráez Padilla
University of Jaén, Jaén, Spain
e-mail: abueno@ujaen.es

J. Ráez Padilla
e-mail: jraez@ujaen.es

M. L. Pérez Cañado (eds.), *Content and Language Integrated Learning in Monolingual Settings*, Multilingual Education 38,
https://doi.org/10.1007/978-3-030-68329-0_5

type of data we want to obtain. As will be indicated below, tests must be designed so that they fulfil certain qualities; they must be reliable and valid. Their criteria must be authentic and their social impact should not be very high. In addition, test tasks should be as authentic as possible and their difficulty and discrimination potential should be adequate. All these characteristics will be defined and explained briefly in the following sections.

One of the main problems education governing bodies face when adopting different linguistic policies is the lack of empirical data that may sustain those decisions. As for CLIL, in words by David Marsh, "there remains insufficient empirical evidence of the impact of differing types of CLIL/EMILE across Europe" (2002, p. 185). With this empirical drive in mind, we designed and validated two tests (see Appendices 1 and 2) assessing use of English, vocabulary, reading, listening, writing and speaking, corresponding to 6th grade of Primary Education and 4th grade of Compulsory Secondary Education. The instruments encompass these six aspects to assess linguistic competence, following the *Common European Framework of Reference for Languages* (henceforth, CEFRL), the national Decrees and the regional Orders which establish the official curriculum for these educational stages, and which all present contents, descriptors and evaluation strategies focusing on grammatical, lexical and skill-based aspects.

After the tests were designed following the aforementioned documents, they were subjected to a doublefold pilot process for their validation. They were initially scrutinised by a minimum of five external experts, who critically assessed their length, adequacy for each level, difficulty, variety of testing facets, types of inputs, clarity of rubrics and layout. After introducing the recommended modifications, they were applied to a representative sample of students from the same levels and of similar characteristics as our target participants in order to determine the internal consistency of their parts, their reliability, and the index of difficulty and discrimination index of the items included within them. Again, necessary readjustments were incorporated after this operational stage. In this phase, we also guaranteed that our instruments met the main testing requirements acknowledged in the specialised literature (content, construct, face and ecological validity, reliability and practicality). The procedures used in this preliminary design phase, together with the ensuing statistical procedures to validate the tests (see Sect. 6) both endeavour to fill the gap pointed out by Marsh above when it comes to the design and validation of testing instruments which may shed light, in the form of contrasted empirical data, upon the linguistic competence of students taking part in CLIL programmes. This will ultimately offer relevant objective data which may inform other processes and decisions concerning CLIL implementation.

In line with the foregoing, the overall rationale of this chapter is to offer instruments which provide empirically grounded data on the present state of CLIL. Only by doing so will we be able to faithfully diagnose the present state of the art in bilingual education. The results will hopefully help educational authorities base the decisions which need to be made for such an ambitious and forward-thinking initiative on empirical data obtained in objective, controlled conditions, such as those pervading

in the design and implementation of the tests below. The ultimate goal is to reinforce the aspects which prove to be yielding positive results and take remedial action where other possible weaknesses and threats are unveiled.

2 The Concept of Language Proficiency

Language proficiency is a term that refers to people's competence or ability to use a language for a specific purpose: the degree of skill with which they are able to comprehend and employ that language. Thus, our level of second language proficiency may range from zero to native-like proficiency in oral and written communication.

Traditionally, the concept of proficiency has been associated with *linguistic competence*. However, this concept has been severely criticised for its excessive emphasis on formal aspects in the second language. Relatively recently, this concept has been completed by incorporating sociolinguistic, pragmatic, cultural and strategic competence (Bachman, 1990; Canale, 1983; Council of Europe, 2001, 2020) and has been termed *communicative competence*.

In general, language speakers know intuitively the forms and rules governing the language and have some knowledge of the linguistic, cognitive, affective and sociocultural aspects which can be expressed with language forms. They have also developed a competence that allows them to communicate and interact with other people in a variety of contexts and situations. That is, they have a sociolinguistic and pragmatic knowledge of the language that allows them to communicate a variety of linguistic functions and use the language in a creative way. In addition, they have developed an intuitive system that enables them to understand and/or produce a considerable variety of messages thanks to the receptive and productive strategies and procedures they have developed, together with their linguistic and communicative competence. As will be seen, all these aspects are included in the concept of language proficiency that will be briefly discussed in this chapter.

Several authors have dealt with language proficiency and have defined its scope and the competences that characterise it (Bachman, 1990; Bachman & Palmer, 1996; Davies, 1990; Harley et al., 1990; Oller, 1979). This construct can be analysed through the language performance of individuals, such as when they listen, speak, read or write the language and it can be explained in different ways, depending on the approaches on which we base our analysis.

Firstly, proficiency has been defined as a linguistic or grammatical competence (Canale, 1983; Council of Europe, 2001, 2020), which focuses on the knowledge and use of phonological, lexical, grammatical, orthographic and orthoepic aspects of language. This competence includes the knowledge and skills required to understand and produce a variety of oral and written messages. Consequently, it is also associated with the mastery of the four skills: listening, speaking, reading and writing.

Secondly, proficiency has been defined as a set of intralingual and crosslingual (mediating) skills. Within the first group, we can include the four skills: (a) audiolingual receptive skills (aural reception, listening activities), (b) audiolingual productive skills (speaking and interactive events), (c) graphic or visual receptive skills (reading tasks) and (d) graphic productive skills (writing activities). In the second group, we can include (a) audiolingual mediating skills (interpreting oral messages) and (b) graphic mediating skills (translating messages) (Council of Europe, 2001, 2020).

Proficiency has also been explained as an integration of language skills and linguistic components (Carroll, 1968). For this author, when analysing proficiency, attention must be paid to oral communication skills (receptive skills: listening and productive skills: speaking) and written communication skills (receptive skills: reading and productive skills: writing). These productive and receptive skills involve the mastery of several linguistic components: phonology, orthography, lexicon and grammar (morphology and syntax).

For Oller (1979), the nature of second language proficiency is unitary and depends on the learner's *pragmatic expectancy grammar.* Communicative interaction is a process in which the interlocutors anticipate part of the information; they then compare the message received with the information expected and react according to the feedback received. In fact, *expectancy* is the key concept for Oller's unitary proficiency theory. In order to evaluate the learner's capacity to interpret, understand and produce messages, he proposes the use of *pragmatic or integrative texts*, which measure the learner's pragmatic competence. The most common procedures are as follows: dictations, combined cloze and dictation, oral cloze procedure, oral interviews, composition or essay writing, narrations and translation (Madrid & Pérez Cañado, 2004).

*C*ummins (1979, 1984) proposes a binary system of basic interpersonal and communicative skills (BICS) and cognitive academic language proficiency (CALP). The former refers to the ability to maintain a simple conversation, for example in a shop or in the street. This competence may be acquired in a relatively short space of time and often occurs in informal and face-to-face situations. The latter refers to the ability to cope with study programmes in academic contexts. It is developed at school and in more formal contexts.

Another approach to the interpretation of language proficiency was developed by the interlanguage theory (Selinker, 1972), which considered the learner's language level as a built-in syllabus that reflects their transitional competence. Error analysis was considered essential to understand the degree of competence that learners had developed (Corder, 1967).

The concept derives from a mentalist interpretation of L2, contemplated as a creative process of construction of a system in which the subjects elaborate their own hypotheses about the target language on the basis of different sources: the target language itself, the mother tongue, the concept of language in general, its communicative function, the conception of the world and of the human beings in it, and the linguistic input we receive from others.

Selinker (1969) introduced the term *interlanguage*, which is very close to that of Nemser's (1971) *approximative system*, and Corder's (1973) *idiosyncratic dialect*. It can be located in what Vygotsky (1978) called *Zone of Proximal Development* and, likewise, can be associated to Krashen's (1985) *i + 1* concept.

This linguistic system is temporary and dynamic (a *continuum*) in the sense that it increasingly acquires more difficulty and complexity. In this respect, it must be considered as an essential part of the learning process, although it may be erroneous if we compare it to the formal system of the target language. Being aware of this concept is crucial to deal with errors, for the development of remedial strategies and techniques and for curricular design so that the syllabus may be compatible with the students' stage of development.

In the seventies a new term was coined that was greatly influenced by sociolinguistic and pragmatic language theories. It was the *communicative competence* concept, which proposed a broader notion of the concept by incorporating into it new sociolinguistic and cultural elements, such as feasibility, contextual appropriateness or communicative efficiency. Nonetheless, the most widely accepted paradigm of communicative competence in Spain put forward by Canale (1983), possibly due to the fact that it was officially adopted as a theoretical framework in the Spanish FL curricular designs (MEC, 1990). Subsequently, other authors have suggested further modifications in the terminology and conceptualisation of this construct (see also Kohonen et al., 1985).

Canale (1983) includes several subcompetences in the communicative competence construct: *grammatical competence* (which includes: phonology, orthography, vocabulary, word formation and sentence formation), *sociolinguistic competence, discourse competence* (cohesion, coherence) and *strategic competence* (which includes strategies needed to resolve grammatical, sociolinguistic, discourse and performance difficulties).

Later on, Bachman (1990), divided communicative competence into *organizational* and *pragmatic* competence. The former includes *grammatical competence* (vocabulary, morphology, syntax, phonetics and graphology) and *textual competence* (cohesion and rhetorical organisation). Pragmatic competence includes *illocutionary competence* (ideational, manipulative, heuristic and imaginative functions) and *sociolinguistic competence* (sensitivity to dialects and variety, differences in register and naturalness, cultural references and figures of speech).

A very widespread model of communicative language competence in Spain has been proposed by the CEFRL (Council of Europe, 2001, 2020), which includes *linguistic competences* (lexical, grammatical, semantic, phonological, orthographic and orthoepic competence), *sociolinguistic competences* (markers of social relations, politeness conventions, expressions of folk wisdom, register differences, dialect and accent) and *pragmatic competences* (discourse competence and functional competence) (Madrid & Pérez Cañado, 2004).

More recently, language proficiency has been described as a set of proficiency levels described by rating scales with specific rubrics and scales (Council of Europe, 2001, 2020). A very popular example is the common reference levels established by the CEFRL: level A (basic user): divided into A1 and A2; B (independent user): divided into B1 and B2; and C (proficient user): divided into C1 and C2. These levels are presented in (1) a global scale in six levels (A1-C2), (2) a self-assessment grid for learners to self-evaluate their proficiency level in listening comprehension, reading, oral interaction, speech production and writing and (3) specific scales to evaluate the learner's proficiency level in aural reception (listening), oral production (speaking), spoken interaction, visual reception (reading), written production (writing) and mediation, by putting a great emphasis on the strategies involved in each process.

In conclusion, given the complexity of language, we believe that a multidimensional concept of proficiency that includes the learning of language content (knowing about language) and procedures (skills), the development of strategies and attitudes, and the learner's intercultural competence can be adequate to design tests which evaluate the outcomes of teaching and learning processes.

3 Evaluating Proficiency Through Achievement Tests

Our purpose has been to evaluate the students' final proficiency level of English as a second language in grade 6 of Primary Education and grade 4 of Compulsory Secondary Education in English as a Foreign Language (EFL) and CLIL programmes. Therefore, our approach has centred on the students' final learning outcomes; that is, on the final product obtained at the end of these grades. We have performed an external evaluation, which has been carried out by external personnel not involved in the teaching process. However, the instruments used for testing, which are based on Madrid and Hughes (2011), had been previously validated by CLIL experts and the teachers involved in the teaching and learning process.

As will be seen below, the tests which have been designed and administered for this study fulfil the basic requirements of reliability, validity, authenticity, interaction, washback, practicality, difficulty and discrimination potential (Bachman, 1990; Bachman & Palmer, 1996; Hughes, 1989; Madrid & Pérez Cañado, 2004).

In our study, the tests proposed are specifically attuned to the levels of the two groups, 6th grade of Primary Education and 4th grade of Secondary Education in the Spanish Education system. Therefore, they are tailor-made tests, as opposed to standardised tests, such as those provided by institutions such as the British Council and Cambridge Examinations, which evaluate the same language components (use of English and vocabulary) and communicative skills (listening, speaking, reading and writing).

In the case of the British Council, information about different aspects of the tests it offers (Aptis and IELTS) can be found at the electronic links provided:

General information: https://www.britishcouncil.org/exam;
IELTS scoring system: https://www.ielts.org/ielts-for-organisations/ielts-scoring-in-detail;
General information about IELTS design: https://www.ielts.org/about-the-test/ensuring-quality-and-fairness.
Aptis scoring system:
https://www.britishcouncil.es/sites/default/files/aptis_scoring_system_layout_final.pdf.
As for Cambridge Examinations, the following link provides useful input about the different Cambridge tests, also including IELTS: https://www.cambridgeenglish.org/exams-and-tests/.

4 Stages of the Validation Process

The validation process of the language tests started with a revision of the literature about the effects of CLIL on the students' language competences in Spain (Lasagabaster, 2008; Lorenzo et al., 2009; Madrid & Hughes, 2011; Pérez Cañado, 2012; Ruiz de Zarobe, 2010, 2011). Then, the language tests validated and used by Madrid and Hughes (2011) were adapted and validated for the evaluation of the new EFL and CLIL population.

4.1 Objectives

The general objective was to evaluate the effects of EFL and CLIL on the Primary (6th grade) and Secondary (4th grade) students' English language competence. This objective involved the evaluation of the sample's linguistic and communicative competences at the end of these grades by using tests that cover the main areas of language proficiency and that are designed according to the principles of language testing as described in Sect. 4 [for the effects of CLIL on EFL competence, please see Chapters "The Impact of CLIL on FL Grammar and Vocabulary" and "The Effects of CLIL on FL Learning: A Longitudinal Study", where the instruments described in the present chapter were applied].

4.2 Test Orientation and Structure

The instruments used to test the students' language competence in the English language in Primary and Secondary Education are presented in Appendices 1–2. These tests were based on Madrid and Hughes (2011) and were revised and enhanced by following Bachman and Palmer's (1996) qualities of language tests: the characteristics they propose for test tasks and the requirements of language ability. Subsequently, they were validated by experts on testing EFL and CLIL students and piloted with Primary and Secondary EFL and CLIL students in their final year of study at each level (Pérez Cañado, 2016). The resulting version of the tests, which was obtained as a product of the piloting phase and an editing process that involved the expert ratings approach, is presented in Appendices 1–2.

4.2.1 Primary English Test

This test (see Appendix 1) aimed at measuring the students' linguistic and communicative competence through a total of 100 items. The first 25 items tested the students' competence with (a) the use of English. The following 15 items focused on (b) the students' lexical competence. Next, (c) their listening comprehension skill was measured through 16 items. Afterwards, (d) the reading skill was evaluated with 15 items. Subsequently, (e) their writing skill was measured through 14 items, which included two tasks: writing an e-mail to a friend (items 72–78) and writing some habitual actions performed by their family at the weekend (items 79–85). Finally, the speaking skill was tested by means of 15 items, ten of them inquiring into the students' personal life (items 86–95), and the other five asking students to describe what is happening in a park (items 96–100). The tasks included in each of the test components were the following:

Use of English (25 points):

A. Writing five questions for given answers (five items: 1–5).
B. Describing the location of several objects and people as illustrated in a picture (five items: 6–10).
C. Matching basic communicative functions with their corresponding language exponents (five items: 11–15).
D. Expressing possession by filling in gapped sentences with the corresponding possessive forms (four items: 16–19).
E. Expressing actions with present and past forms by completing gapped sentences with the corresponding verb forms (six items: 20–25).

Vocabulary (15 points):

F. Expressing habitual actions by completing gapped sentences with the given verbs (four items: 26–29).

G. Writing some parts of the body with the help of the corresponding illustrations (six items: 30–35).
H. Completing a brief gapped dialogue with the missing lexical items (five items: 36–40).

Listening comprehension (16 points):

I. Listening to a short dialogue about summer holidays and matching what each child did with the corresponding illustration (six items: 41–46).
J. Listening to a short conversation about a party and completing a table with the required information (five items: 47–51).
K. Listening to a short dialogue about Christmas presents and matching what each child got with the corresponding illustration (five items: 52–56).

Reading (15 points):

L. Reading a short text (72 words) about pets, graded in terms of difficulty, and answering four comprehension questions (four items: 57–60).
M. Identifying structures, morphological features of words and the vocabulary studied in a written text about the weather in Britain (six items: 61–66).
N. Completing a brief text based on a nursery rhyme by using the missing elements (five items: 67–71).

Writing (14 points):

O. Writing a brief e-mail message about holiday activities (seven items/points: 72–78).
P. Describing a member of the family by following a model (seven items/points: 79–85).

Speaking (15 points):

Q. Ten questions about the students' personal life (ten items: 86–95).
R. Describing what is happening in a park (five items: 96–100).

Once all interviews were conducted, each question was scored by two researchers on a scale from 0 to 1 by taking into account the use of grammar, lexical range, fluency and interaction, pronunciation and task fulfilment. In the evaluation of written production, assessment was based on grammatical, semantic, pragmatic and orthographic accuracy and on elements of cohesion and coherence.

4.2.2 Secondary English Test

The Secondary English Test (see Appendix 2) was also based on Madrid and Hughes (2011). It also included 100 items. The tasks included in each test component were the following:

Use of English (26 points):

A. Rewriting questions by using the correct word order (four items: 1–4).
B. Changing sentences from the passive into the active voice (four items: 5–8).
C. Completing a brief gapped text (52 words) with verbs in the past tense (ten items, 0.5 points each: 9–13).
D. Completing gapped sentences with the right form of given verbs and other words (ten items, 0.5 points each: 14–18).
E. Expressing hypothetical situations (conditionals) by choosing the right verb forms (eight items, 0.5 points each: 19–22).
F. Completing a brief gapped text (94 words) with the past or the present perfect verb forms of the verbs given in brackets (eight items, 0.5 points each: 23–26).

Vocabulary (15 points):

G. Completing a brief gapped text (79 words) with the given words and expressions (seven items: 27–33).
H. Matching illustrated symbols/icons with their corresponding meaning/text (eight items: 34–41).

Listening comprehension (14 points):

I. Listening to the recording of two presenters and choosing the correct multiple choice options as they correspond to specific data mentioned in the recording (seven items, 2 points each: 42–55).

Reading (14 points):

J. Reading a text of 413 words that is divided into five paragraphs and choosing the correct multiple choice option according to the text's content (14 items/points: 56–69).

Writing (15 points):

K. Writing a brief e-mail reply to a message (15 items/points: 70–84).

Speaking (16 points):

L. Questions about the students' personal life (eight items: 85–92).
M. Taking part in a two-way dialogue about one of the following three scenarios: (1) an imaginary trip to New York (the things the interlocutors are going to bring and what they are going to do), (2) organising a surprise birthday party for a friend (the things they are going to buy and what they are going to do) or (3) a school project on animals (the animals they are going to write about and what they are going to do for the project) (four items: 93–96).
N. Taking part in a three-way dialogue about two topics: Block A: The importance of English and the Internet or Block B: The importance of school and mobile phones (four items: 97–100).

The criteria used to grade the students' speaking and writing competence were the same as the ones used in Primary Education.

Table 1 Population used to validate the tests in Primary and Secondary Education

Primary Education ($N = 828$)	
Provinces	Cádiz: 14.5%, Córdoba: 3.6%, Badajoz: 14.4%, Cáceres: 5.2%, Granada: 13%, Jaén: 11%, Málaga: 15%, Sevilla: 23%
Setting	rural: 33.7%, urban: 66.3%
Type of school	public: 84.7%, private: 3.5%, charter: 11.8%
Programme	bilingual: 42.4%, non-bilingual: 57.5%
Gender	males: 49.2%, females: 50.7%
Father/mother studies	no studies: 8.5%, primary: 25%, secondary: 35%, university: 31.5%
Secondary Education ($N = 1{,}196$)	
Provinces	Almería: 9%, Cádiz: 34%, Córdoba: 2.3%, Badajoz: 9%, Cáceres: 4.2%, Granada: 12%, Jaén: 11%, Málaga: 5.5%, Sevilla: 13%
Setting	rural: 39.3%, urban: 60.7%
Type of school	public: 69.1%, private: 6.4%, charter: 24.5%
Programme	bilingual: 53%, non-bilingual: 47%
Gender	males: 50.6%, females: 49.4%
Father/mother studies	no studies: 8%, primary: 32.5%, secondary: 34.5%, university: 25%

5 Population Used for Validating the Tests

The characteristics of the population used for validating the texts in Primary and Secondary Education are shown in Table 1.

6 Procedures Used to Validate the Tests

In the process of planning, designing and validating the achievement tests that were used in this study, the following procedures were taken into account and applied:

Reliability. The tests that we designed and administered showed a considerable internal consistency of results, as proved by the Cronbach alpha calculations (Table 2).

For the marking of open-end items in writing and speaking, rubrics designed and validated following the CEFRL were used, together with CAF analysis (Complexity, Accuracy, Fluency). Moreover, in the marking of speaking, inter-rater reliability was fostered by specific instructions provided to markers in which it was stated that two researchers-examiners were, if possible, to be present in the oral interview in pairs (10 mins), one of them conducting it and the other one observing and taking notes. Students' answers were recorded with their consent, and it was recommended to have a girl-boy pair in each interview, so that their answers could be easily differentiated in the recording. For the same reason, researchers-examiners were also recommended

Table 2 Reliability of the language test in Primary and Secondary Education

Primary Education ($N = 828$)			Secondary Education ($N = 1116$)		
Components	Cronbach alpha	Items (N)	Components	Cronbach alpha	Items (N)
Whole test	.94	71	Whole test	.96	69
Part 1: Use of English	.89	25	Part 1: Use of English	.94	26
Part 2: Vocabulary	.86	15	Part 2: Vocabulary	.85	15
Part 3: Listening	.80	16	Part 3: Listening	.62	14
Part 4: Reading	.81	15	Part 4: Reading	.68	14

to address students with certain frequency by their names so as to ease male–female identification in the pair.

Content and construct validity. All the tests were designed so that they measured what they were intended to measure. Their validity was put into effect in relation to their content (content validity) and the linguistic and communicative constructs defined previously (construct validity).

Authenticity. The test items in Primary Education are less authentic than in Secondary Education, since many of them are pre-communicative, but the degree of authenticity increases in Secondary Education, where most of the proposed activities are more related to real-life tasks.

Interactive characteristics. The extent to which the students rely on their individual characteristics to solve the questions is intensive since there is a solid representation of male and female students with a variety of linguistic, communicative and existential competences due to the variety of EFL and CLIL programmes that they followed.

Social and educational impact. The social and educational impact on the task-takers was minimal, since they were informed that the tests' main purpose was to measure the levels of English proficiency obtained at the end of the grade they were studying but that the results would in no way influence their marks.

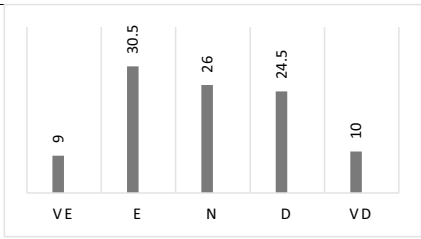
Practicality. Tests were designed in order to make their application viable and feasible so that they could be completed in approximately one hour.

Item difficulty. The item difficulty index (DI) shows how easy or difficult items have been in a test. It is obtained by calculating the percentage of correct answers for each item. The criteria used to interpret the DI and the results obtained with the tests in Appendices 1–2 are shown in Table 3 (Lafourcade, 1977; Madrid & Pérez Cañado, 2004):

Item discrimination. The discrimination index shows the extent to which a test item separates high achievers from low achievers. Good test items must be solved successfully by good learners and unsuccessfully by those with less ability. If that

Table 3 Results of the English test items difficulty indexes (DI) in Primary and Secondary Education with corresponding graphic representation

PRIMARY EDUCATION	
Criteria and DI	%
Very easy: ≥ 0.75	9
Easy: 0.55-0.74	30.5
Normal: 0.45-0.54	26
Difficult: 0.25-0.44	24.5
Very difficult: < 0.25	10



SECONDARY EDUCATION	
Criteria and DI	%
Very easy: ≥ 0.75	21
Easy: 0.55-0.74	36
Normal: 0.45-0.54	14
Difficult: 0.25-0.44	29
Very difficult: < 0.25	0

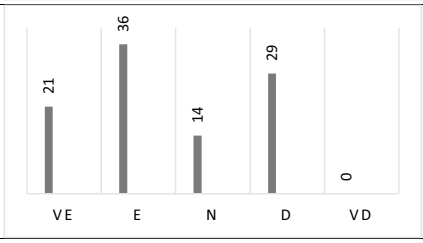
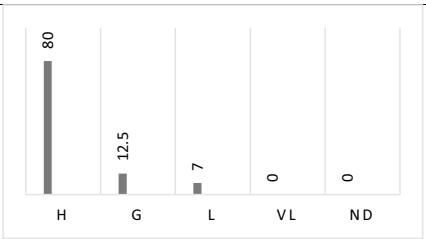
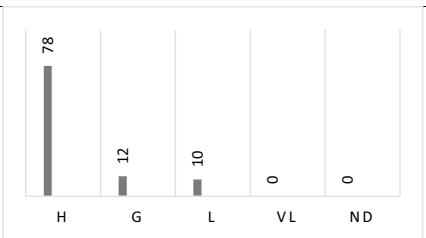


Table 4 Results of the English test items discrimination indexes in Primary and Secondary Education with corresponding graphic representation

PRIMARY EDUCATION	
Criteria and discrimination indexes	%
High: ≥ 0,40	80.5
Good: 0,30-0,39	12.5
Little: 0,20-0,29	7
Very little: 0,10-0.19	0
No discrimination: < 0,10	0



SECONDARY EDUCATION	
Criteria and discrimination indexes	%
High: ≥ 0,40	78
Good: 0,30-0,39	12
Little: 0,20-0,29	10
Very little: 0,10-0.19	0
No discrimination: < 0,10	0



is not the case, the item cannot be said to "discriminate" among students correctly (Heaton,1975; Madrid & Pérez Cañado, 2004) (Table 4).

7 Conclusions

The language test characteristics and results presented in Sect. 6 allow us to conclude that the sample tests presented in the appendices are reliable and valid to test the students' linguistic and communicative competence. Their item difficulty is adequate and their discrimination potential is high (80.5% in Primary Education and 78% in Secondary Education). This responds to the need of designing classroom tests bearing in mind test requirements such us validity (content validity, construct validity and face validity), reliability (internal and external), difficulty, practicality and discrimination, and ensuring they are fulfilled by using statistical procedures. In this respect, piloting them is also a pre-requisite.

We hope that the teachers involved in CLIL and non-CLIL programmes find these empirically grounded tests useful to control the results obtained in both programmes and to compare them in order to evaluate to what extent the introduction of CLIL schemes in the education system is worthwhile. Likewise, we encourage reflective practitioners to design their own tests (thus responding to their students' needs and wants) and to validate them both empirically (by using them in class) and statistically (by applying the appropriate statistical measures). This will also contribute to strengthen the image of *the teacher as researcher*, as a clear example of *action research.*

The pedagogical implications derived from results will undoubtedly improve their own teaching and testing practice and will be useful for other teachers and test designers, and for students in both CLIL and non-CLIL contexts. We hope this research study will provide some insights into test design and validation.

Acknowledgements This work was supported by the Spanish Ministry of Economy and Competitiveness [research grant number FFI2012-32221] and the Government of Andalusia [research grant number HUM 2348] (Project: *The Effects of Content and Language Integrated Learning in Monolingual Communities: A Large-Scale Evaluation*).

Appendix 1: English Test, Primary Education, 6th Grade

/100

ENGLISH TEST
PRIMARY EDUCATION, 6TH GRADE

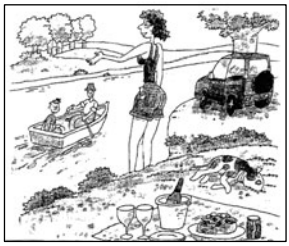
Centro: _____; Curso: ____; Fecha: ___ /___ / 20___
Apellidos, Nombre: _____; Sexo: H ☐ / M ☐

Use of English: /25

A. Write questions for these answers:

(1) My name's Bert: _____?
(2) I live in Madrid: _____?
(3) Scotland's in the north of Great Britain: _____?
(4) I'm fine, thanks!: _____?
(5) My computer's grey: _____?

B. Look at the picture on the right. Read and complete.
Use: ON, IN, BY, WITH, UNDER



(6) The man's _____ the boat.
(7) The woman's _____ the river.
(8) The boy's _____ the man.
(9) The car's _____ the tree.
(10) The dog's _____ the grass.

C. Match the numbers with the letters:

(11)	1. Expressing ability	A. Let's play hide & seek
(12)	2. Narrating what is happening	B. Hi, Anne! How are you?
(13)	3. Greeting people	C. This is Pablo.
(14)	4. Introducing a friend	D. Can you play the guitar?
(15)	5. Suggesting a game	E. The plane's taking off.

1: _____; 2: _____; 3: _____; 4: _____; 5: _____

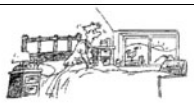
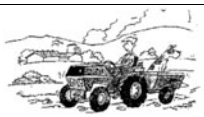
D. Complete using: MY, ITS, HIS; THEIR
(16) John and Peter are doctors. This is _____ hospital.
(17) This is a policeman. That's _____ car.
(18) Look at the fox. _____ tail isn't short. It's long.
(19) I'm a clerk. This is _____ office.

E. Complete using: DO, DOES, DID, AREN'T, ISN'T, IS.

(20) Is your teacher Spanish? No, s/he _____.
(21) _____ John study English yesterday?
(22) Your friends _____ sleeping. They are reading English.
(23) _____ your teacher reading a book? Yes, s/he is.
(24) _____ he live in London?
(25) _____ they live in Africa?

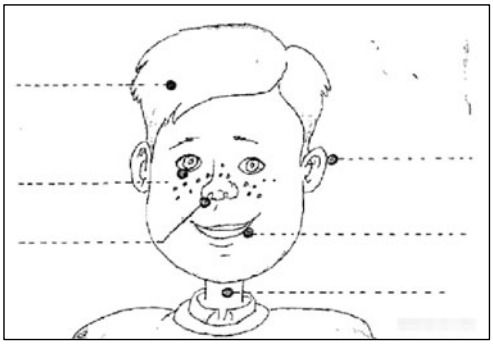
Vocabulary: /15

F. Complete using: FEEDS, CHASES, DRIVES, GETS UP

			
(26) In summer he _____ early.	**(27)** He _____ the pigs.	**(28)** Sometimes he _____ the chickens.	**(29)** And sometimes he _____ a tractor.

G. Write the parts of the body:

(30) _____
(32) _____
(33) _____



(31) _____
(34) _____
(35) _____

H. (36-40) Read and complete. Use: A GIRL, RIGHT, HERE, LOVELY, NAME

Listening: /16

I. (41-46) Mary and Rex are talking about their summer holidays. What did they do? Listen and match, using arrows, with the corresponding picture.

Mary	Rex

J. (47-51) Amy and Clive are having a party. Write in the table below what they have got and what they haven't got:

They've got…	
They haven't got any…	

K. (52-56) Listen to a short dialogue about Christmas presents and match, using arrows, what each child got with the corresponding illustration.

Bess	Sam	Vic





Reading: /15

L. Read this brief text and then answer the questions below:

Have you got any pets?

Hi there, friends! These are our pets. We've got a white goat, a grey rabbit and a small brown, white pony. There's also a hen and woolly sheep. Ah, and a noisy dog that barks in the evening! All of them eat grass but they also like the food we buy in the supermarket! We often play with them in the garden and have fun!

(57) What's the text about? _____
_____.

(58) What happens with the dog? _____
_____.

(59) What's the pony like? _____
_____.

(60) What do they eat? _____
_____.

M. What's the weather like in Britain? Read this text about the weather in Britain and underline the correct words:

The weather in Britain

In Britain it rains a lot and it is very **(61)** *want/wind/windy/with* too. So the countryside is very **(62)** *sweet/green/well/water* and there are lots of bushes, trees and hedges in the **(63)** *sea/fields/sky/car*. The weather is very changeable: very **(64)** *good/often/well/nice* you may have a cloudy and foggy **(65)** *morning/fish/pencil/bike*, a sunny and windy afternoon, and a **(66)** *with/wet/well/left* evening.

N. Read the following poem:

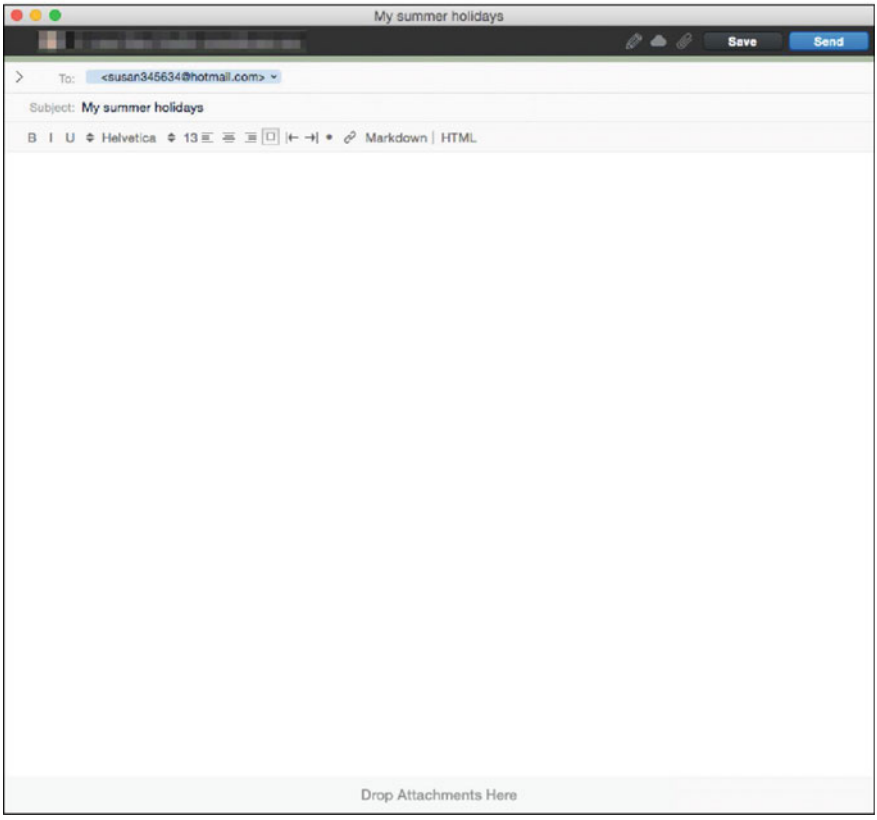
Jack and Jill went up the hill
To fetch a pail of water
Jack fell down and broke his crown
And Jill came tumbling after!

Now tell the story by using the following words: BROKE, WALKED UP, HURT, WAS, CAME DOWN.

Yesterday **(67)** _____ a terrible day for Jack and Jill. They **(68)** _____ the hill to fetch a bucket of water and they fell down. Jack **(69)** _____ his head and Jill **(70)** _____ after him and **(71)** _____ her arms and knees.

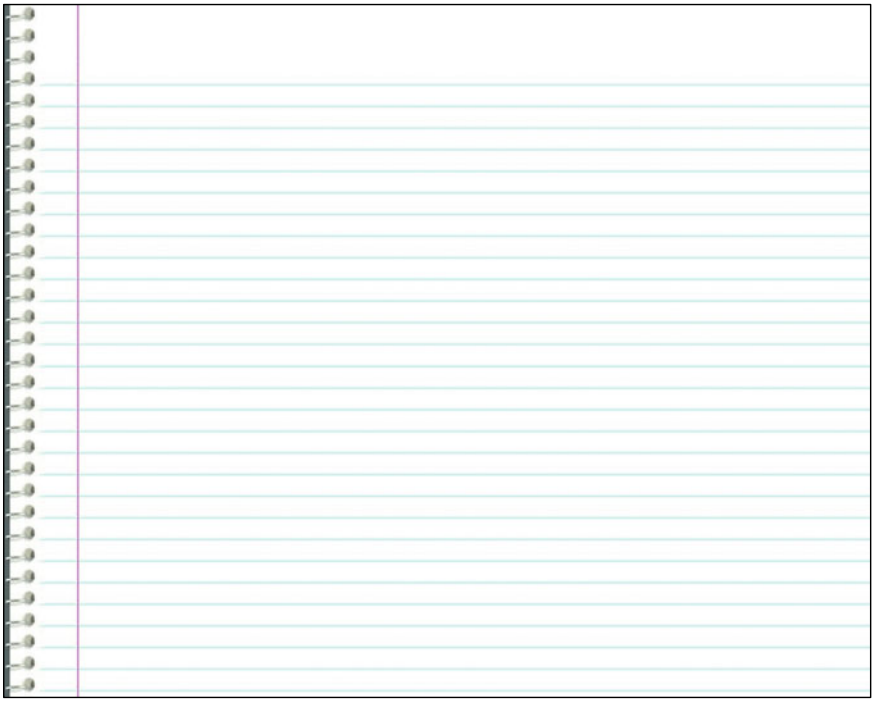
Writing: /14

0. (72-78) Write an e-mail to a friend telling him/her what you do on your summer holidays.



P. (79-85) Write about your family: your brother, sister, mother or father. What do they do at the weekend? Use the following example as a guide:

My brother Pedro gets up late at the weekend. He has breakfast and then helps my parents in the bar. He rides his bike and plays soccer with his friends, too.

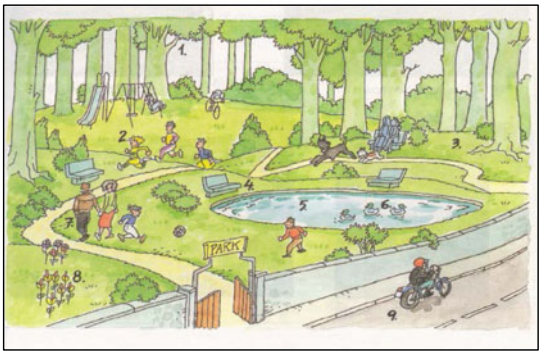


Speaking: /15

Q. In pairs. Answer the questions below about your personal life.

(86) Hi, what's your name? And what's your English teacher's name? ☐
(87) How old are you? ☐
(88) Where do you live? ☐
(89) What's your mother's first language? ☐
(90) And your father's first language? ☐
(91) What language do you speak at home? ☐
(92) How many brothers or sisters have you got? ☐
(93) What are your hobbies? What do you like? ☐
(94) What did you do last weekend?☐
(95) What are your plans for this summer? What will you do? ☐

R. In pairs. Now let's talk about this park:



(96) Are there any people in this park? What people can you see? ☐
(97) Are there any animals? What animals can you see? ☐
(98) Are there any swings? Where are they? ☐
(99) What's the boy in picture 9 doing? ☐
(100) Is there a park near your house or in your town? What is it like? What is there in that park? ☐

Appendix 2: English Test, Secondary Education, 4th Grade

/100

ENGLISH TEST

SECONDARY EDUCATION, 4TH GRADE

Centro: _____; Curso: ____; Fecha: ___ /___ / 20___

Apellidos, Nombre: _____; Sexo: H ☐ / M ☐

Use of English: /26

A. Rewrite these questions using the correct word order:

(1) long the how programme does last?

_____?

(2) to radio Sunday did listen the last you?

_____?

(3) of do like kind which programme you?

_____?

(4) loud too radio the is why?

_____?

B. Change these sentences into the active. Example: Pictures are transmitted by television → *Television transmits pictures*.

(5) The transmission can be interrupted by high mountains.

High mountains _____.

(6) How can this problem be solved?

How can we _____?

(7) Other transmitters can be built on the top of the hills.

We can _____.

(8) The waves are changed into sound by the aerial.

The aerial _____.

C. (9-13) Complete the text with the verbs in the past tense:

Yesterday Aunt Betty _____ (lose) her glasses. She _____ (can) find them. "Help me!", she _____ (ask). "O.K.", I _____ (reply). "Remember where you _____ (put) them". "Yes, now what _____ (do) I do? I _____ (come) into the house. I _____ (take) off my coat. Tibby _____ (want) some food. I _____ (give) him some". At last she found her glasses.

D. (14-18) Complete these sentences with: *anything/something*; *tell/say* (in the right form); *bring/take* (in the right form):

- Do you mean to _____ me that you're going out with Celo AGAIN??!
- Can you _____ a bottle of wine to my party, please? Thanks.
- Did you speak to him? Yes, but I didn't understand a word he _____!
- I'm so nervous! I've got an exam tomorrow, and I can't remember _____!
- If you go to Scotland, _____ some warm clothing with you!
- Come here: there's _____ I want to _____ you.
- Last month he _____ a break for a week.
- Do you understand _____ about car engines?
- Yesterday he _____ to me: "I'm not going to drive tomorrow. I'll walk".

E. Underline the right words in brackets:

(19) If I (*lived / am living*) in London, (*I'd go / I'll go*) to the theatre every week.
(20) If we (*went / go*) to Madrid, (*we'll visit / we visit*) you.
(21) If my father (*has / had*) a car, he (*would travel / travelling*) a lot.
(22) If I (*am / were*) you, I (*wouldn't / will*) eat so much.

F. (23-26) Complete this text with the correct forms –the PAST (for example, *went*) or the PRESENT PERFECT (for example, *has gone*)– of the verbs in brackets:

I'm feeling terribly hungry, because I _____ (not, have) my lunch yet. I _____ (have) breakfast four hours ago, at 8 o'clock, and I _____ (not, have) anything since then. So you can imagine how I feel. I have a lot of work to do these days: yesterday I _____ (stay) at the office until half past seven! I _____ (never, do) that before, but my boss _____ (want) me to finish the report I was writing. Luckily I _____ (finish) it on time, so I _____ (be) able to get home in time to see the football on TV.

Vocabulary: /15

G. Complete this text using one of the following words and expressions: ON YOUR OWN, OVERCOME, QUIT, CUT DOWN, IMPROVE, HEALTH, WILL POWER, HARM.

If you want to keep your **(27)** _____, it is important to **(28)** _____ smoking or at least to **(29)** _____: let's say from twenty to ten or five cigarettes a day. Even if you only reduce by three a day, your health will **(30)** _____. It isn't easy, of course; you need a lot of **(31)** _____, and you will need to **(32)** _____ the temptation to start smoking again. But keep at it! Smoking does you tremendous **(33)** _____: so quit now!

H. (34-41) Match symbol and text:

1		A. Hot water to all washbasins
2		B. Tents admitted (with number of pitches and rates)
3		C. Shaver points
4		D. Dogs admitted on lead
5		E. Sailing from site
6		F. Riding/pony trekking from site
7		G. Indoor heated swimming pool on site
8		H. Fishing at site

1: _____; 2: _____; 3: _____; 4: _____; 5: _____; 6: _____; 7: _____; 8: _____

Listening: /14

I. You will hear a news presenter. You must listen and circle the correct answer. You will hear the news report twice.

(42-43) The news report takes place on

A. Thursday 8th February B. Tuesday 8th February
C. Thursday 18th February D. Thursday 18th January

(44-45) Which sentence is true?
A. Babies are fatter than before B. Five-year-old children are fatter than before
C. Babies are thinner than before D. Five-year-old children are thinner than before

(46-47) What have children stopped doing?

A	B	C	D
			

(48-49) How much money will Peter possibly earn in total?

A. 50 000 dollars
B. 150 000 dollars
C. 165 000 dollars
D. 200 000 dollars

(50-51) What instrument does Peter play now?

A	B	C	D
			

(52-53) What is Kirstie's profession?

A. A model
B. An actress
C. A hockey player
D. A flight attendant

(54-55) Where does Kirstie work?

A. In Australia
B. In New Zealand
C. In London
D. In Manchester

Reading: /14

J. Read the following text and circle the correct answer:

* * *

BODY TALK

A smile, a frown, a handshake or a kiss. All of these actions are part of our everyday communication and make up what is commonly known as "body language". Even though we might not know it, we are constantly sending messages with our bodies without speaking. Some researchers believe that our body language makes up to sixty percent of all of our face-to-face communication and it has been shown that it plays an important role in key areas of our lives, including our success in job interviews, understanding what people are saying, making friends and falling in love. But what are the keys to body language? Here are some examples.

Paragraph A

Hitting someone in the face is a very obvious example of aggressive body language; but there are more subtle forms of showing your aggression. If you look at someone directly in the eyes, frown, and lean forward, you are showing the other person that you do not agree with them and that you probably do not like what they are saying. If you cross your arms or you legs, you are showing a defensive posture. Similarly, if you avoid eye contact with someone, you might be unconsciously telling them that you do not want to tell them the truth.

Paragraph B

It might seem unfair, but the success of a job interview often depends on the first few seconds of the interview and during these initial moments your body is giving many important signals. In an interview situation, apart from dressing well, it is important to move with confidence, not too fast and not too slow. You should also have a pleasant facial expression and try to show positive signs like interest, not boredom or nervousness.

Paragraph C

For many people, one of the most important aspects of body language is the communication of your feelings towards another person. Common signs of romantic love can be observed when two people smile at each other, sit or stand close together and look at each other for longer periods of time than usual.

Paragraph D

So what can we do to improve communication with our bodies? Firstly, it is important to be conscious of our own body language and, in certain situations, to try to control it. We should also try to observe the other person as we pay attention to what they are saying in order to understand not only the words, but also the feelings that they are expressing.

* * *

(56-59) Which title is best for each of the paragraphs?

First impressions	A	B	C	D
Attack or defence	A	B	C	D
How to use body language	A	B	C	D
I love you	A	B	C	D

(60-61) According to the text, in which situation is body language useful?

A. When you are speaking to another person on the phone
B. When you are writing a job application
C. When you are speaking in the presence of another person
D. When you are playing a role

(62-63) If a person does not look at you, it is possible that

A. S/he is lying to you
B. S/he is in love with you
C. S/he is confident
D. S/he is being aggressive

(64-65) In an interview,

A. What you wear in not important
B. You might get the job depending on your body language
C. You should cross your legs
D. You should have a neutral expression

(66-67) According to the text, looking at another person for a long while can be a sign of

A. Indifference
B. Affection
C. Self-defence
D. Boredom

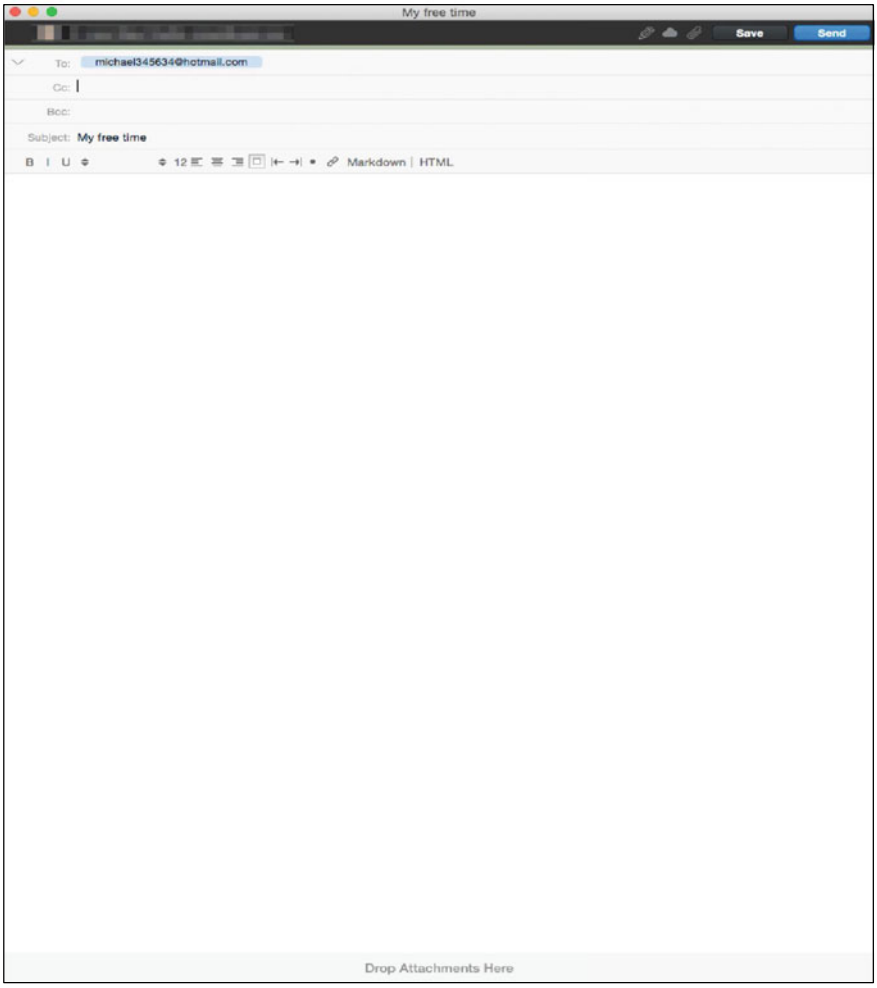
(68-69) When using body language, it is useful to

A. Avoid eye contact
B. Show that you are bored
C. Hit the other person
D. Look and listen to the other person

Writing: /15

K. (70-84) You receive an e-mail from an English friend, Michael. Here is part of the e-mail: *I have just started to do karate. It's great! I go every Tuesday and Thursday. What about you? What do you do in your free time?*

Answer Michael's e-mail in the space provided below:



Speaking: /16

L. In pairs. Answer the questions below about your personal life.

(85) Hi, what's your name? And surname(s)? ☐
(86) How do you spell that [referring to name or surname(s)? ☐
(87) Where do you live? ☐
(88) Can you describe your house? ☐
(89) Who do you live with? ☐
(90) What do you like to do in your free time? ☐
(91) Have you ever travelled to another country [if not, city]? What did you do there? ☐
(92) What would you like to do in the future? ☐

M. In pairs. Spoken interaction: two-way dialogue. Use ONLY one scenario.

- **Scenario 1**

Examiner: Now, I want you to imagine that you are going on a trip to New York. Both of you have to plan the trip. Look at the examples in the pictures below and decide:

- **(93-94)** What things you are going to bring with you
- **(95-96)** What you are going to do in New York

* * *

- **Scenario 2**

Examiner: Now, I want you to imagine that you are going to organise a surprise birthday party for a friend. Both of you have to plan the party. Look at the examples in the pictures below and decide:

- **(93-94)** What you are going to buy for the party
- **(95-96)** What you are going to do at the party

* * *

- **Scenario 3**

Examiner: Now, I want you to imagine that you are going to do a school project on animals. Both of you have to plan the project. Look at the examples in the pictures below and decide:

- **(93-94)** What animal or animals you are going to write about
- **(95-96)** What you are going to do for the project

C. Three-way dialogue.

Examiner: now the three of us are going to talk about two topics. Here we have to give our opinion on each thing. We can ask questions, debate, agree or disagree [examiner chooses only from Block A or Block B].

Block A	Block B
The importance of English **(97-98)**	The importance of school **(97-98)**
Internet **(99-100)**	Mobile phones **(99-100)**

References

Bachman, L. F. (1990). *Fundamental considerations in language testing*. Oxford: Oxford University Press.

Bachman, L. F., & Palmer, A. S. (1996). *Language testing in practice.* Oxford: Oxford University Press.

Canale, M. (1983). From communicative competence to communicative language pedagogy. In J. C. Richards & R. W. Schmidt (Eds.), *Language and communication* (pp. 2–27). London: Longman.

Carroll, J. B. (1968). The psychology of language testing. In A. Davies (Ed.), *Language testing symposium: A psycholinguistic perspective* (pp. 46–69). Oxford: Oxford University Press.

Corder, S. P. (1967). The significance of learners' errors. *International Review of Applied Linguistics, 5*(4), 161–170.

Corder, S. P. (1973). *Introducing applied linguistics.* Harmondsworth: Penguin.

Cummins, J. (1979). Linguistic interdependence and the educational development of bilingual children. *Review of Educational Research, 49,* 222–251.

Cummins, J. (1984). Wanted: A theoretical framework for relating language proficiency to academic achievement among bilingual students. In R. Charlene (Ed.), *Language proficiency and academic achievement* (pp. 2–19). Clevedon: Multilingual Matters.

Council of Europe. (2001). *Common European framework of reference for languages: Learning, teaching, assessment*. Cambridge: Cambridge University Press.

Council of Europe. (2020). *Common European framework of reference for languages: Learning, teaching, assessment (Companion Volume).* Council of Europe.

Davies, A. (1990). *Principles of language testing*. Oxford: Basil Blackwell.

Harley, B., Allen, P., Cummins, J., & Swain, M. (1990). *Development of second language proficiency*. Cambridge: Cambridge University Press.

Heaton, J. B. (1975). *Writing English language tests.* London: Longman.

Hughes, A. (1989). *Testing for language teachers.* Cambridge: Cambridge University Press.

Kohonen, V., von Essen, H., & Klein-Bradley, C. (Eds.). (1985). *Language testing in school*. Tampere, Finland: Finnish Association for Applied Linguistics.

Krashen, S. D. (1985). *The input hypothesis: Issues and implications.* London and New York: Longman.

Lafourcade, P. (1977). *Evaluación de los aprendizajes.* Madrid: Cincel.

Lasagabaster, D. (2008). Foreign language competence in content and language integrated courses. *The Open Applied Linguistics Journal, 1,* 31–42.

Lorenzo, F., Casal, S., & Moore, P. (2009). The effects of content and language integrated learning in European education: Key findings from the Andalusian bilingual sections evaluation project. *Applied Linguistics, 31*(3), 418–442.

Marsh, D. (Ed.). (2002). *CLIL/EMILE. The European dimension. Actions, trends, and foresight potential*. Jyväskylä: University of Jyväskylä.

MEC. (1990). *Diseño curricular base.* Madrid: Ministerio de Educación y Ciencia.
Oller, J. W. (1979). *Language tests at school: A pragmatic approach.* London: Longman.
Madrid, D., & Hughes, S. (Eds.). (2011). *Studies in bilingual education.* Bern: Peter Lang.
Madrid, D., & Pérez Cañado, M. L. (2004). Evaluation. In D. Madrid & N. McLaren (Eds.), *TEFL in Primary Education* (pp. 442–480). Granada: Editorial Universidad de Granada.
Nemser, W. (1971). Approximative systems of foreign language learners. *International Journal of Applied Linguistics, 9,* 115–125.
Pérez Cañado, M. L. (2012). CLIL research in Europe: Past, present, and future. *International Journal of Bilingual Education and Bilingualism, 15*(3), 315–341. https://doi.org/10.1080/13670050.2011.630064.
Pérez Cañado, M. L. (2016). 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.
Ruiz de Zarobe, Y. (2010). Written production and CLIL: An empirical study. In C. Dalton-Puffer, T. Nikula, & U. Smit (Eds.), *Language use and language learning in CLIL classrooms* (pp. 191–209). Amsterdam: John Benjamins.
Ruiz de Zarobe, Y. (2011). Which language competencies benefit from CLIL? An insight into applied linguistic research. In Y. Ruiz de Zarobe, J. Sierra, & F. Gallardo del Puerto (Eds.), *Content and Foreign Language Integrated Learning* (pp. 129–153). Bern: Peter Lang.
Selinker, L. (1969). Language transfer. *General Linguistics, 9*(2), 67–92.
Selinker, L. (1972). Interlanguage. *International Review of Applied Linguistics, 10,* 209–231.
Vygotsky, L. S. (1978). *Mind in society: The development of higher psychological processes.* Cambridge, MA: Harvard University Press.

CLIL and L1 Competence Development

Elvira Barrios

Abstract One of the major concerns regarding Content and Language Integrated Learning (CLIL) education is the detrimental effect it may have on first language development and achievement. In light of contradictory evidence, the present investigation set out to investigate this controversial issue in a specific monolingual context (Andalusia, Spain); additionally it attempted to fill some voids in the literature concerning the impact of intervening variables (gender, setting (urban vs. rural), parental education level, extramural exposure to the foreign language (English), verbal intelligence and academic motivational factors) on L1 attainment. In order to guarantee the homogeneity and, hence, the comparability of the groups, participants were previously matched in terms of verbal intelligence and academic motivation. The study sample consisted of 720 school students. Of these, 247 (34.3%) were 6th Grade Primary school students aged 11–12 and 473 (65.7%), 4th Grade Compulsory Secondary Education school students aged 15–16. Results indicate that the curricular competence in Spanish (L1) of CLIL students was not negatively affected when L1 school grades are compared to those of their peers in regular classes. Additionally, intervening variables such as setting, gender, parents' education level, extramural exposure to English seemed to have a dissimilar effect on the grades of CLIL and non-CLIL groups. Furthermore, the effects of the variable verbal intelligence and the four motivational factors considered by the study (desire to work and self-esteem, exam anxiety, lack of interest in studying and realistic personal self-demand) on school grades in Spanish were more pronounced in the Secondary Education CLIL group.

1 Introduction

One of the major concerns associated with Content and Language Integrated Learning (henceforth CLIL) programmes is that they might entail detrimental effects on first

E. Barrios (✉)
University of Málaga, Málaga, Spain
e-mail: elvira.barrios@uma.es

M. L. Pérez Cañado (eds.), *Content and Language Integrated Learning in Monolingual Settings*, Multilingual Education 38,
https://doi.org/10.1007/978-3-030-68329-0_6

language (L1) development and proficiency (Baetens Beardsmore, 2002; Halbach, 2009; Merisuo-Storm, 2006; Pladevall-Ballester, 2015; Wolff, 2005). Several studies have compared the performance of CLIL and non-CLIL groups on different measures of L1 competence (e.g. Merisuo-Storm, 2006, 2007; Merisuo-Storm & Soininen, 2014; Seikkula-Leino, 2007). Most of them conclude that the learners' L1 development is not hindered by CLIL instruction; however, some conflicting results are also found in the literature (Hämäläinen, 1998; Koivumäki & Stara, 1994). These inconsistent findings call for more research attention into an issue—that of L1 attainment of students participating in CLIL programmes—with potentially far-reaching consequences. Additionally, although the study of the effects of CLIL programmes on the mother tongue is one of the areas around which CLIL research has been conducted (Wolff, 2005), there are still gaps in the literature which need to be addressed, such as the effect of possible intervening variables on L1 achievement in CLIL—in comparison with non-CLIL—programmes. Lastly, in order to draw firm(er) conclusions, the relationship between CLIL instruction and L1 development and achievement across a range of contexts (geographical, educational, etc.) should be explored.

In light of this analysis, this chapter is an attempt to contribute to fill this void in the literature by investigating the effects of certain variables on L1 attainment in CLIL and non-CLIL students of Primary and Secondary Education in two provinces in the south of Andalusia (Spain). Specifically, this study tried to answer the following research questions:

1. Is the curricular competence in Spanish (L1) of CLIL-educated students negatively affected in comparison to that of their conventionally schooled peers?
2. What is the effect of intervening variables (gender, setting—urban vs. rural—, parents' educational level, extramural exposure to English, verbal intelligence and motivational factors) on the level of attainment in Spanish of CLIL-educated students, both in itself and compared to their conventionally schooled peers?

The study aimed to explore the effect of CLIL on L1 competence within the framework of the two government-funded research projects referred to in the Introduction to this book. It did so by addressing the methodological flaw in CLIL research discussed in the literature concerning the comparability of CLIL and non-CLIL groups (e.g. Lyster, 2007; Madrid Fernández, 2006; Ruiz de Zarobe & Lasagabaster, 2010).

2 Literature Review

As previously mentioned, although most research has attested to the beneficial effect of CLIL instruction on L1 development and achievement, some contradictory results are also found in the literature (Hämäläinen, 1998; Koivumäki & Stara, 1994). In addition, methodological flaws in the study design that have been relatively common in CLIL research may have been responsible for erroneous attributions of gains in L1 observed in CLIL groups to CLIL education solely (e.g. Bonnet, 2012; Bruton, 2011a, b, 2013, 2015; Dallinger et al., 2016; Paran, 2013; Pérez-Cañado, 2012).

Finland stands out as the country in which research into the association between CLIL instruction and L1 learning has been most extensively studied. Seikkula-Leino (2007) refers to research conducted in this country (Hämäläinen, 1998; Koivumäki & Stara, 1994; Merisuo-Storm, 2000, 2002; Rahman, 2001) that provides mixed evidence as to the positive impact of CLIL on L1 development. In the first of these studies, Koivumäki and Stara (1994) focused on the first and second grades of comprehensive school and concluded that, although the CLIL group outperformed the so-called normal group in the reading speed test, the reverse was the case with mechanical writing skills, reading comprehension and vocabulary. Hämäläinen's (1998) outcomes confirmed these study results with reference to vocabulary skills from the second to fifth grades. It must be noted that the test consisted mainly of vocabulary items related to the fields of nature and the environment. However, results obtained by Merisuo-Storm (2000, 2002) would seem to suggest otherwise; in her first study (Merisuo-Storm, 2000), the development of pupils' reading and writing skills during the first school year between the learners taught in Finnish and those taught in a foreign language (henceforth FL) were comparable. Similar results were obtained in a further study (Merisuo-Storm, 2002) that confirmed that pupils partly taught (20% of their teaching) in an FL in the first two years of school developed better reading skills than those who were taught in Finnish. And, finally, Rahman's (2001) research also provides support to the claim that CLIL does not exert a negative effect on the pupils' mother tongue. Furthermore, the learning of L1 reading and writing skills of CLIL and non-CLIL classes from grades 5 and 6 did not differ significantly in Seikkula-Leino's (2007) study, which led the researcher to conclude the following:

> … this study shows that as regards Finnish as a mother tongue, the pupils' learning results were similar in both Finnish and CLIL classes. CLIL pupils overachieved even more strongly than those in Finnish teaching, even though the difference was not remarkable … Hence, the results of this study strongly support the idea that the mother tongue skills of pupils learning in a foreign language were not weaker than the skills of those learning in their mother tongue. (p. 336)

The long-term effects of participating in an immersion programme on L1 writing were assessed by Bergroth (2006) in the context of the Finnish university entrance examination. Participants were involved in an immersion programme that started when they were 3–6 years old and ended by the end of elementary school (9th grade). Results indicated that a) most immersion students demonstrate a level of competence in written composition comparable to the average level in Finland among Finnish-speaking students; and b) the fact that "There is a lower number of low results in mother tongue texts amongst the immersion students than expected by the matriculation board" (p. 128) could be indicative that the programme might support L1 acquisition among weaker students. Similarly, Merisuo-Storm (2006, 2007) found no negative effect of CLIL teaching—20% of the instruction was delivered in English—on Primary learners in a longitudinal study in Finland, with the CLIL group performing particularly well in reading comprehension skills. The researcher concludes: "The results of the study support the view that when most of the teaching is carried out in the pupils' first language, CLIL does not affect the development of first language literacy

skills negatively" (p. 233). This supports the analysis adopted by Wolff (2005) who, while recognising language competence in the L1 as one of the controversial issues concerning CLIL, argues that, provided only a limited number of subjects are taught through the FL, "An impoverishment of first language competence can be regarded as rather exceptional" (p. 18). More recently, another longitudinal study investigated the effects of CLIL on the development of Primary students' L1 literacy skills during the first six school years (Merisuo-Storm & Soininen, 2014). The study concluded that CLIL classes had a statistically significant advantage over regular classes in terms of L1 reading and writing skills, and they also showed more positive ideas towards reading, writing and FL learning.

In a Secondary school context in The Netherlands, no negative effect was found on the school leaving exams for Dutch when a group of students receiving bilingual education was compared with a regular Secondary group (Admiraal et al., 2006). As pointed out by Pérez-Cañado (2012), the cohorts were not matched with respect to baseline L1 scores; additionally, the authors also admit that results may be influenced by the fact that bilingual schools were part of an experimental programme that may have benefitted from the extra support allotted to them and the knowledge and experience in bilingual education they had accrued over the years. Further evidence that the L1 is not negatively affected by CLIL education was obtained by Gebauer et al. (2012) in Germany. They compared the development of CLIL and traditionally educated students' L1 reading fluency and orthographic skills during elementary school. Cognitive abilities and socio-economic background were used as covariates. The study also detected that CLIL students' reading fluency progressed faster.

It has even been suggested that the linguistic advantage of CLIL-educated students may extend beyond their L1 development; the higher performance in mathematics of a group of CLIL students in Belgium when compared with their conventionally schooled peers led Surmont et al. (2016) to conjecture that the increased metalinguistic awareness found in bilingual students evidenced in a better understanding of and insight into linguistic structures could also lie behind the better performance in the understanding of the abstract language of maths shown by CLIL Secondary students.

Studies conducted in Spain also seem to attest to the absence of detrimental consequences of CLIL programmes on the students' mother tongue. Egiguren (2006 cited in Ruiz de Zarobe & Lasagabaster, 2010) concluded that CLIL did not negatively affect the acquisition of two co-official languages (Basque and Spanish) in Primary school students. Several studies have also corroborated the positive effect of CLIL on attitudes towards languages in general (e.g. Lasagabaster & Sierra, 2009) and towards trilingualism (Lasagabaster, 2009; Ruiz de Zarobe & Lasagabaster, 2010). As Ruiz de Zarobe and Lasagabaster (2010) concluded from their study with a group of 3rd and 4th year students of Compulsory Secondary Education, "CLIL can be influential in producing more positive attitudes towards languages and the language learning process" (p. 23).

In the Andalusian (monolingual) context, Ramos et al. (2011) researched the effect of CLIL on L1 achievement in three school types (private bilingual, public bilingual and monolingual and charter (semiprivate) monolingual), thus adding the type of

school as an intervening variable in CLIL impact research. Unlike the present study, the authors constructed an *ad-hoc* Spanish (L1) test to gauge attainment of L1 knowledge and skills. The test was aimed at evaluating the official curricular objectives and competences at the time of the study. As a whole, CLIL groups achieved higher levels of attainment than non-CLIL groups. The public school CLIL sample outperformed the public school non-CLIL sample with statistical significance both in Primary and in Secondary Education; additionally, the private school CLIL group obtained significantly higher scores than the public school non-CLIL sample in Primary Education, although the difference between these two samples was small in Secondary Education. The private school CLIL group also obtained significantly higher scores than the charter school one in Secondary Education, whereas no significant difference was found in Primary Education.

As previously mentioned, methodological weaknesses in CLIL research studies compromise the validity of results concerning the attribution of the L1 attainment advantage of CLIL over non-CLIL students to CLIL education. In order to address methodological flaws found in other studies on CLIL impact on L1 learning, the participants in the bilingual and the monolingual samples in Pérez-Cañado's (2018) study had previously been matched in terms of verbal intelligence and motivation so as to guarantee the homogeneity of the treatment and comparison groups. They were Primary and Secondary Education students from three monolingual communities in Spain: Andalusia, Extremadura and the Canary Islands. Furthermore, she explored the effect on the Spanish competence of the following intervening variables: type of school (public, private, charter), setting (rural vs. urban) and socio-economic status. The study concluded that the L1 (Spanish) academic competence was not negatively affected by CLIL instruction; in fact, the CLIL sample outstripped the monolingual one at the end of both educational levels. As to the effects caused by moderating variables, it was substantial for type of school and socio-economic status, but not for rural vs. urban setting. Public and private bilingual school groups performed significantly higher than the public non-bilingual school counterpart. Socio-economic status, on the other hand, caused statistically significant differences for CLIL and non-CLIL strands both in Primary and in Secondary Education.

3 The Study

3.1 Participants

The study sample consisted of 720 school students from 13 public (n = 551, 76.5%) one private (n = 42, 5.8%), and three charter schools (n = 127, 17.6%). Of these, 247 (34.3%) were 6th Grade Primary Education students aged 11–12 and 473 (65.7%), 4th Grade Compulsory Secondary Education students aged 15–16. The schools were located in the provinces of Cádiz and Málaga (southern Andalusia, Spain).

Table 1 Information on the primary education sample

School type		Programme		Gender	
Public n (%)	Charter n (%)	CLIL n (%)	Non-CLIL n (%)	Men n (%)	Women n (%)
182 (73.7)	65 (26.3)	81 (67.2)	166 (32.8)	111 (44.9)	136 (55.1)

Table 2 Information on the secondary education sample

School type			Programme		Gender	
Public n (%)	Charter n (%)	Private n (%)	CLIL n (%)	Non-CLIL n (%)	Men n (%)	Women n (%)
369 (78.0)	62 (13.1)	42 (8.9)	275 (58.1)	198 (41.9)	253 (53.5)	220 (46.5)

The majority of students attended urban schools (n = 594, 82.5%) (see Tables 1 and 2 for further information on the sample).

3.2 Instruments and Measures

At a preliminary stage of the investigation, the CLIL and the non-CLIL groups were matched for verbal intelligence through the subtest for verbal aptitude in the EFAI (*Evaluación factorial de las aptitudes intelectuales*) (Santamaría et al. 2014). The level 2 verbal aptitude subtest was used for Primary students and the level 4 subtest was employed for Secondary students. This subtest has acceptable levels of difficulty and discrimination indices, and good internal reliability (coefficient $\alpha = 0.75$ for Level 2 subtest and $\alpha = 0.70$ for Level 4 subtest). The authors also report adequate measures of validity for the overall test. The subtest for verbal aptitude in the EFAI has a multiple-choice format, and each question has four possible answers. The level 2 subtest contains 26 questions, and the level 2 subtest, 23.

In order to guarantee the homogeneity and, hence, the comparability of the groups in terms of motivation, Pelechano's (1994) MA test was used. This self-report questionnaire consists of 36 statements, each of which has a dichotomous (yes/no) response format and identifies four motivational factors: (i) desire to work and self-esteem (10 items); (ii) exam anxiety (9 items); (iii) lack of interest in studying (with a potential negative (inhibitory) effect (9 items); and (iv) realistic personal self-demand (7 elements). The score in L1 (Spanish) was used as a measure of academic competence in Spanish (L1). Spain uses a 10-point grading scale for Primary and Secondary Education divided into categories where 9.0–10 is "outstanding" (*sobresaliente*), 7–8.9 "very good" (*notable*), 6–6.9 "good" (*bien*), 5–5.9, "sufficient" (*suficiente*), and below that, "fail" (*insuficiente*).

A questionnaire based on an instrument developed by Sundqvist and Sylvén (2014) was employed to obtain information on length of time of extramural exposure to English. Extramural English is a term coined by Sundqvist (2009 cited in Sundqvist & Sylvén, 2014) to refer to "all types of English-related activities that learners come in contact with or are engaged in outside the walls of the English classroom, generally on a voluntary basis" (Sundqvist & Sylvén, 2014, p. 4). The parents' highest level of education together with demographic data was obtained through an initial questionnaire administered to the student.

Tests and questionnaires were distributed in all the participating schools during class time under the researchers' supervision. The test for verbal aptitude and the motivation questionnaires were administered in the same session in February–March 2015. The schools provided the researchers with the end-of-year scores in Spanish in June 2015. Previous to the data gathering stage of the study, the prescriptive permissions had been obtained from both the Education Administration and the participating schools.

3.3 Data Analysis

For statistical data analyses, chi square tests were used in order to compare the CLIL and the non-CLIL samples' attainment in L1, and the effect of gender, setting (urban vs. rural), parents' educational level, and extramural exposure to English on the CLIL sample's L1 competence. The school grades "fail", "sufficient" and "good" were collapsed into one category, and those of "very good" and "outstanding" into another. The education levels of the mother and of the father were distributed into non-university and university level. As to the time spent on extramural activities in English, two categories were used: up to 9 h and more than 9 h of weekly exposure to extramural English.

ANOVA tests were conducted in order to detect the effect of verbal intelligence and motivational factors (desire to work and self-esteem, exam anxiety, lack of interest in studying and realistic personal self-demand) on the L1 competence. In this case, three categories of grades were used: fail, the collapsed category of "sufficient" and "good", and that of "very good" and "outstanding".

4 Results and Discussion

The chi square test results indicate that there are no statistically significant differences in the school grades in Spanish (L1) between the CLIL and the non-CLIL sample either at Primary ($\chi^2 = 0.26$, df $= 1$, p $= 0.79$, $V = 0.05$) or Secondary Education ($\chi^2 = 0.032$, df $= 1$, p $= 0.96$, $V = 0.012$). This would lend support to previous findings in that CLIL instruction does not have a damaging effect on the students'

L1 competence (e.g. Gebauer et al., 2012; Merisuo-Storm, 2000, 2002, 2006,2007; Merisuo-Storm & Soininen, 2014; Seikkula-Leino, 2007).

When Secondary urban and rural CLIL students are compared in terms of attainment in Spanish, no statistically significant difference is found ($\chi^2 = 0.26$, df = 1, p = 0.74, $V = 0.05$). This result is consistent with that obtained by Pérez-Cañado (2018), also in connection with L1 performance. In the case of English as a FL, however, Alejo and Piquer-Píriz (2016) found that urban CLIL students reached a higher level of attainment than their urban counterparts. Contrarily, a statistically significant difference is observed in the case of non-CLIL students ($\chi^2 = 9.54$, df = 1, p = 0.004, $V = 0.28$). Rural non-CLIL students performed significantly lower than their urban counterparts; whereas 22.22% of the former obtained grades between "very good" and "outstanding" in Spanish, more than double (56.04%) of the latter obtained such grades.

The variable gender only had a significant effect on the Spanish competence of non-CLIL secondary students ($\chi^2 = 5.68$, df = 1, p = 0.03, $V = 0.22$). In this group, 59.65% of the female students obtained grades between "very good" and "outstanding" (this value was 52.54% in the CLIL sample), whereas 37.70% of the male students achieved such grades (41.19% in the CLIL sample). This result seems to corroborate Merisuo-Storm and Soininen's (2014) findings that led them to conclude that "especially the boys seemed to have benefited from bilingual education" (p. 72) as, while the girls' literacy skills were significantly better than the boy' in regular classes in their study, this was not the case in CLIL classes. Additionally, in light of this result it can be conjectured that CLIL instruction may contribute to narrow the gap between girls and boys in literacy performance, though this should be confirmed in further studies. It should be noted that in PISA 2015 girls achieved better scores than boys in reading in all OECD countries (OECD, 2016).

Studies have shown that parental level of education influences their children's academic performance as this level impacts on the beliefs and attitudes that promote learning; besides, a strong connection has been found between mothers' educational level and children's academic achievement and cognitive development (Davis-Kean, 2005; Sirin 2005). In our study, the mother's education level seemed to have a considerable impact on the participants' school achievement in L1. Students in all groups whose mothers have a higher educational level (that is, a university level) obtain better grades in Spanish. This effect achieves a statistically significant level in the case of the Primary non-CLIL sample ($\chi^2 = 8.32$, df = 1, p = 0.008, $V = 0.32$), the Secondary CLIL sample ($\chi^2 = 5.84$, df = 1, p = 0.03, $V = 0.22$) and the Secondary non-CLIL sample ($\chi^2 = 10.48$, df = 1, p = 0.003, $V = 0.30$). In contrast, the impact of the father's education level did not seem to be so evident as it only reached statistical significance in the case of the Secondary CLIL sample ($\chi^2 = 4.81$, df = 1, p = 0.05, $V = 0.20$).

Quite unexpectedly, the chi square result indicates a relationship between the extramural exposure to English and Secondary non-CLIL students' level of attainment in Spanish ($\chi^2 = 6.69$, df = 1, p = 0.03, $V = 0.22$). In this group, 40% of those exposed to English up to 9 h a week obtained grades between "very good" and "outstanding"; of the students exposed to English more than 9 h 62.79% achieved

those grades. It may be hypothesised that exposure to language—to any language—may have an impact on the person's development in other language(s) of his or her repertoire; this hypothesis, however, needs confirmation.

The results from the descriptive statistics and the ANOVA tests aimed to detect the effect of verbal intelligence and motivational factors (desire to work and self-esteem, exam anxiety, lack of interest in studying and realistic personal self-demand) on L1 attainment are shown in Tables 3 (Primary Education) and 4 (Secondary Education). As expected, the descriptive statistical results indicate that, overall, the higher the performance in the verbal intelligence test, the higher the school grade in Spanish L1. The ANOVA test revealed that, except in the case of the CLIL Primary group, there are statistical significant differences in Spanish performance between all the different attainment groups.

As to the motivational factors in Pelechano's (1994) MA test, results were, on the whole, not surprising either in CLIL or in non-CLIL groups. Without exception, the higher the grades obtained in Spanish, the higher the values in the desire to work and self-esteem factors. Differences between attainment groups achieved significant levels only in Secondary Education. Exam anxiety exhibited a somewhat erratic behaviour, though, and did not yield any significant difference between groups. When it comes to the potentially debilitating lack of interest in studying factor, higher values in it also corresponded to lower grades; additionally, differences between levels of attainment in Spanish regarding this factor reached significant level at Secondary Education both in the CLIL and the non-CLIL group, although the effect size in the latter was considerably smaller. Finally, the relationship between self-demand and the school grades in Spanish followed the expected trend as higher values in this motivational factor corresponded to higher grades; however, the difference in this factor between attainment groups in Spanish was statistically significant only in the Primary non-CLIL group and in the Secondary CLIL group, where the eta partial squared value indicated a large effect size.

In general, as Table 4 shows, the effects of the variable verbal intelligence and the four motivational factors in the MA test on school grades in Spanish seem to be more pronounced in the Secondary Education CLIL group, as both the p and the effect values indicate. A different effect of enrolment in a CLIL class on L1, FL and subject content learning outcomes was observed by Madrid and Barrios (2018).

5 Conclusion

It is reasonable to suppose that the L1 competence of students in CLIL programmes—who are consequently exposed to fewer hours of instruction in Spanish—may be negatively affected as a result of their participation in such a programme. In view of this possibility, this study aimed at investigating whether CLIL instruction had a detrimental effect on Spanish (L1) competence in a specific monolingual context

Table 3 Mean, standard deviation and variation in verbal intelligence and motivational factors according to school grade levels in Spanish (Primary Education)

Non-CLIL						
Variable	Scores					
	Insufficient (I) (n=13)	Sufficient + good (SG) (n=29)	Very good + outstanding (VO) (n=54)			
	M (SD)	M (SD)	M (SD)	*F* (2, 93) (*p*)	ηp2	Comments
Verbal intelligence	8.69 (2.87)	12.83 (2.70)	15.00 (2.93)	26.73 (< .001)	.37	VO>SG>I**
Desire to work and self-esteem	3.77 (1.96)	4.79 (1.98)	5.06 (1.71)	2.60 (.080)	.053	
Exam anxiety	7.08 (1.115)	6.28 (1.579)	6.30 (1.609)	1.46 (.24)	.030	
Lack of interest in studying	4.08 (1.44)	3.55 (1.76)	2.74 (1.81)	4.04 (.080)	.080	
Self-demand	1.23 (1.42)	1.55 (1.27)	2.41 (2.02)	3.65 (.030)	.073	VO>I*
CLIL						
	Scores					
	Insufficient (I) (n=1)	Sufficient + good (SG) (n=9)	Very good + outstanding (VO) (n=10)			
	M (SD)	M (SD)	M (SD)	*F* (2, 17) (*p*)	ηp2	Comments
Verbal intelligence	9.00	9.33 (3.32)	11.30 (4.27)	.68 (.52)	.074	–
Desire to work and self-esteem	3.00	3.22 (1.20)	5.70 (1.49)	8.40 (.003)	.50	–
Exam anxiety	6.00	6.33 (1.23)	5.90 (1.45)	.25 (.78)	.028	–
Lack of interest in studying	2.00	2.89 (2.03)	2.40 (1.17)	.28 (.76)	.032	–
Self-demand	2.00	1.89 (1.05)	3.10 (1.73)	1.71 (.21)	.17	–

* *Note* Significant at a level of $p < 0.05$
** *Note* Significant at a level of $p < 0.001$

Table 4 Mean, standard deviation and variation in verbal intelligence and motivational factors according to school grade levels in Spanish (Secondary Education)

Non-CLIL						
Variable	Scores					
	Insufficient (I) (n=23)	Sufficient + good (SG) (n=38)	Very good + outstanding (VO) (n=57)			
	M (SD)	M (SD)	M (SD)	*F* (2, 215) (*p*)	ηp2	**Comments**
Verbal intelligence	9.35 (2.95)	10.03 (3.27)	11.02 (2.66)	3.066 (.050)	.051	
Desire to work and self-esteem	3.43 (1.70)	4.05 (1.94)	4.72 (1.98)	3.985 (.021)	.065	VO>I*
Exam anxiety	5.65 (2.10)	5.89 (1.81)	5.96 (1.96)	.213 (.808)	.004	
Lack of interest in studying	5.26 (1.51)	4.97 (2.11)	4.18 (1.66)	3.950 (.022)	.064	VO>I*
Self-demand	1.04 (1.15)	1.37 (1.50)	1.74 (1.49)	2.091 (.128)	.035	
CLIL						
	Scores					
	Insufficient (I) (n=15)	Sufficient + good (SG) (n=50)	Very good + outstanding (VO) (n=58)			
	M (SD)	M (SD)	M (SD)	*F* (2, 120) (*p*)	ηp2	**Comments**
Verbal intelligence	9.67 (3.70)	9.64 (3.32)	11.69 (2.96)	6.248 (.003)	.094	VO>I*
Desire to work and self-esteem	3.87 (1.73)	3.90 (1.91)	4.90 (2.12)	3.901 (.023)	.061	VO>SG* VO>I*
Exam anxiety	5.53 (2.48)	6.34 (2.00)	6.12 (1.84)	.956 .387)	.016	
Lack of interest in studying	5.93 (2.40)	4.94 (2.39)	3.14 (1.70)	15.705 (< .001)	.207	VO>I** VO>SG**
Self-demand	.67 (.72)	1.08 (.92)	1.98 (1.40)	12.191 (< .001)	.169	VO>I** VO>SG**

* *Note* Significant at a level of $p < 0.05$
** *Note* Significant at a level of $p < 0.001$

(Andalusia, Spain). Additionally, the study sought to determine the effect of moderating variables (gender, setting (urban vs. rural), parental education level, extramural exposure to the foreign language (English), verbal intelligence and academic motivational factors) on level of Spanish competence attainment in Spanish.

Results indicate that the L1 competence is not jeopardised by participating in the CLIL programme, thus supporting the view that a limited number of subjects taught through the medium of the FL does not adversely impinge on the students' competence in L1. Secondary urban and rural CLIL students do not significantly differ in their school competence in Spanish. However, the attainment in Spanish (L1) in the CLIL group seems to be affected differently by other variables depending on the educational programme. In this sense, gender has a statistically significant effect on the attainment of non-CLIL secondary students only. Also, while in Secondary Education the mother's education level significantly impacts on the Spanish competence in both the CLIL and the ordinary educational programmes, the father's education level only has a significant impact on the Spanish competence of the CLIL group. Additionally, more than 9 h of weekly exposure to English has a significant impact on L1 attainment only in the Secondary non-CLIL group. Concerning the effect of verbal intelligence and the four academic motivational factors included in the study on L1 competence, our study concludes that it is more pronounced in the Secondary CLIL group.

These results must be taken with caution, though, as our study clearly has some limitations. Firstly, given both the sample size and the distinctive implementation of CLIL in Andalusia, results may not be transferable to other contexts in which CLIL is being implemented. Secondly, only studies with a pretest–posttest design which control for potentially confounding and intervening variables can determine whether findings are attributable to the educational programme (although in our study participants were matched for verbal intelligence and academic motivation, no baseline data were available). Thirdly, since the end-of-year scores were used as a proxy for attainment in Spanish there is no guarantee that such scores are reliable indicators of the different levels of attainment in Spanish since there is the danger that each school sets its own standards.

Further studies in this field are therefore required that use pretest–posttest design with a control group. Additionally, standardised instruments are needed to measure L1 attainment in different schools and programmes (e.g. Ramos et al., 2011). Finally, although our findings are promising, they should be validated by a larger sample size. Notwithstanding these limitations, this study provides further evidence as to the learning impact of participating in a CLIL programme and explores a research area—that of the effect of intervening variables on learning in this programme in comparison with the regular educational programme—that has not been thoroughly investigated so far.

Acknowledgements This work was supported by the Spanish Ministry of Economy and Competitiveness [research grant number FFI2012-32221] and the Government of Andalusia [research grant number HUM 2348] (Project: *The Effects of Content and Language Integrated Learning in Monolingual Communities: A Large-Scale Evaluation*).

References

Alejo, R., & Piquer-Píriz, A. (2016). Urban vs. rural CLIL: An analysis of input-related variables, motivation and language attainment. *Language, Culture and Curriculum, 29*(3), 245–262.

Admiraal, W., Westhoff, G., & de Bot, K. (2006). Evaluation of bilingual secondary education in the Netherlands: Students' language proficiency in English. *Educational Research and Evaluation, 12*(1), 75–93. https://doi.org/10.1080/13803610500392160.

Baetens Beardsmore, H. (2002). The significance of CLIL/EMILE. In D. Marsh (Ed.), *CLILE/EMILE—The European dimension: Actions, trends and foresight potential* (pp. 24–26). Jyväskylä: UniCOM Continuing Education Centre.

Bergroth, M. (2006). Immersion students in the matriculation examination: Three years after immersion. In S. Björklund, K. Mård-Miettinen, M. Bergström, & M. Södergård (Eds.), *Exploring dual-focussed education: Integrating language and content for individual and societal needs* (pp. 123–134). Vaasan: Vaasan Yliopiston Julkaisuja. https://www.uwasa.fi/materiaali/pdf/isbn_952-476-149-1.pdf. Accessed 30 June 2017.

Bonnet, A. (2012). Towards an evidence base for CLIL: How to integrate qualitative and quantitative as well as process, product and participant perspectives in CLIL research. *International CLIL Research Journal, 4*(1), 66–78. https://www.icrj.eu/14/article7.html. Accessed 30 June 2017.

Bruton, A. (2011a). Is CLIL so beneficial, or just selective? Re-Evaluating Some of the Research. *System, 39*(4), 523–532. https://doi.org/10.1016/j.system.2011.08.002.

Bruton, A. (2011b). Are the differences between CLIL and non-CLIL groups in Andalusia due to CLIL? A reply to Lorenzo, Casal, and Moore (2010). *Applied Linguistics, 32*(2), 236–241. https://doi.org/10.1093/applin/amr007.

Bruton, A. (2013). CLIL: Some of the reasons why... and why not. *System, 41*(3), 587–897. https://doi.org/10.1016/j.system.2013.07.001.

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.

Dallinger, S., Jonkmann, K. & Hollm, J. (2016). Selectivity of content and language integrated learning programmes in German secondary schools. *International Journal of Bilingual Education and Bilingualism*, 1-12. https://doi.org/10.1080/13670050.2015.1130015.

Davis-Kean, P. E. (2005). The influence of parent education and family income on child achievement: The indirect role of parental expectation and the home environment. *Journal of Family Psychology, 19*(2), 294–304. https://doi.org/10.1037/0893-3200.19.2.294.

Gebauer, S. K., Zaunbauer, A. C. M., & Möller, J. (2012). Erstsprachliche Leistungsentwicklung im Immersionsunterricht: Vorteile trotz Unterrichts in einer Fremdsprache? *Zeitschrift Für Pädagogische Psychologie, 3,* 183–196. https://doi.org/10.1024/1010-0652/a000071.

Halbach, A. (2009). The primary school teacher and the challenges of bilingual education. In E. Dafouz & M. Guerrini (Eds.), *CLIL across educational levels* (pp. 19–26). London: Richmond.

Hämäläinen, M. (1998). *Äidinkielen kehittyminen vieraskielisessä opetuksessa: Sanavarasto-tutkimus.* Unpublished EdMa thesis, University of Turku, Teacher Training School, Turku.

Koivumäki, K., & Stara, P. (1994). *Puolalan koulun kieliluokkien opetuksessa käytettävän englannin kielen määrä ja tilanteet sekä oppilaiden äidinkielen taidot*. Unpublished EdMa thesis, University of Turku, Teacher Training School, Turku.

Lasagabaster, D. (2009). The implementation of CLIL and attitudes towards trilingualism. *ITL, International Journal of Applied Linguistics, 157,* 23–45.

Lasagabaster, D., & Sierra, J. M. (2009). Language attitudes in CLIL and traditional EFL classes. *International CLIL Research Journal, 1,* 4–17.

Lyster, R. (2007). *Learning and teaching languages through content: A counterbalanced approach.* Amsterdam: John Benjamins.

Madrid Fernández, D. (2006). Bilingual and plurilingual education in the European and Andalusian context. *International Journal of Learning, 12*(4), 177–185.

Madrid, D., & Barrios, E. (2018). CLIL across diverse educational settings: Comparing programmes and school types. *Porta Linguarum, 29,* 29–50.

Merisuo-Storm, T. (2000). *Ensiluokkalaisen ¨aidinkielen lukemisen ja kirjoittamisen taitojen kehittyminen kielirikasteisella luokalla.* Unpublished EdMa thesis, University of Turku, Teacher Training School, Turku.

Merisuo-Storm, T. (2002). *Oppilaan ¨aidinkielen lukemisen ja kirjoittamisen taitojen kehittyminen kaksikielisess¨a alkuopetuksessa (Publication Series C: 185).* Turku: Painosalama.

Merisuo-Storm, T. (2006). Development of boys' and girls' literacy skills and learning attitudes in CLIL education. In S. Björklund, K. Mård-Miettinen, M. Bergström, & M. Södergård (Eds.), *Exploring dual-focussed education: Integrating language and content for individual and societal needs* (pp. 176–188). Vaasan: Vaasan Yliopiston Julkaisuja. https://www.uwasa.fi/materiaali/pdf/isbn_952-476-149-1.pdf. Accessed 30 June 2017.

Merisuo-Storm, T. (2007). Pupils' attitudes towards foreign-language learning and the development of literacy skills in bilingual education. *Teaching and Teacher Education, 23*(2), 226–235. https://doi.org/10.1016/j.tate.2006.04.024.

Merisuo-Storm, T., & Soininen, M. (2014). Students' first language skills after six years in bilingual education. *Mediterranean Journal of Social Sciences, 5*(22), 72–81. https://doi.org/10.5901/mjss.2014.v5n22p72.

OECD. (2016). *PISA 2015 results (Volume I): Excellence and equity in education.* Paris: PISA, OECD Publishing. https://doi.org/10.1787/9789264266490-en.

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.

Pelechano, V. (1994). Prueba MA. *Análisis Y Modificación De La Conducta, 20,* 71–72.

Pérez-Cañado, M. L. (2012). CLIL research in Europe: Past, present, and future. *International Journal of Bilingual Education and Bilingualism, 15*(3), 315–341. https://doi.org/10.1080/13670050.2011.630064.

Pérez-Cañado, M. L. (2018). The effects of CLIL on L1 and content learning: Updated empirical evidence from monolingual contexts. *Leaning and Instruction, 57,* 18–33. https://doi.org/10.1016/j.learninstruc.2017.12.002.

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.

Rahman, H. (2001). *Kaksikielisen (suomi-englanti) ja suomenkielisen kuudennen luokan äidinkielen kirjoitelmien vertailua yhdyssanojen oikeinkirjoituksen osalta.* Unpublished EdMa thesis, University of Helsinki, Teacher Training School, Helsinki.

Ramos García, A. M., Ortega Martín, J. L., & Madrid, D. (2011). Bilingualism and competence in the mother tongue. In D. Madrid & S. Hughes (Eds.), *Studies in bilingual education* (pp. 135–155). Bern: Peter Lang.

Ruiz de Zarobe, Y., & Lasagabaster, D. (2010). CLIL in a bilingual community: The Basque autonomous community. In D. Lasagabaster & Y. Ruiz de Zarobe (Eds.), *CLIL in Spain: Implementation, results and teacher training* (pp. 12–29). Newcastle: Cambridge Scholars Publishing.

Santamaría, P., Arribas, D., Pereña, J., & Seisdedos, N. (2014). *EFAI: Evaluación factorial de las aptitudes intelectuales.* Madrid: TEA.

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.

Sirin, S. R. (2005). Socioeconomic status and academic achievement: A meta-analytic review of research. *Review of Educational Research, 75,* 417–453. https://doi.org/10.3102/00346543075003417.

Sundqvist, P., & Sylvén, L. K. (2014). Language-related computer use: Focus on young L2 English learners in Sweden. *ReCALL, 26*(1), 3–20. https://doi.org/10.1017/S0958344013000232.

Surmont, J., Struys, E., Van Den Noort, M., & Van De Craen, P. (2016). The effects of CLIL on mathematical content learning: A longitudinal study. *Studies in Second Language Learning and Teaching, 6*(2), 319–337. https://doi.org/10.14746/ssllt.2016.6.2.7.

Wolff, D. (2005). Approaching CLIL. In D. Marsh (Ed.), *The CLIL quality matrix. Central workshop report 6/2005* (pp. 10–25). https://www.ecml.at/mtp2/CLILmatrix/pdf/wsrepD3E2005_6.pdf. Accessed June 30 2017.

The Impact of CLIL on FL Grammar and Vocabulary

Macarena Navarro-Pablo

Abstract This chapter examines the impact of Content and Language Integrated Learning (CLIL) on 351 Primary and Secondary Education students' English language attainment regarding grammar and vocabulary in seven bilingual public schools and one non-bilingual charter school in the province of Seville. Intergroup comparisons are carried out across the different schools and discriminant analyses are performed with all the intervening variables of the study (setting, socio-economic status, verbal intelligence, motivational variables and extramural exposition) in order to determine whether CLIL is truly responsible for the differences ascertained or whether other variables account for a greater proportion of the variance.

1 Introduction

CLIL is nowadays the methodological approach adopted by most European countries in order to meet the challenges posed by the European Union's multilingualism policy (European Union, 2008), which establishes that European citizens should be able to communicate in two languages other than their mother tongue. The implementation of this measure implies the need for greater levels of foreign language (FL) proficiency, which therefore calls for a revision and, probably in many countries, a thorough reform of the approaches to FL teaching and learning. Dalton-Puffer (2014) points out that CLIL is a methodological revolution, not only in the context of FL teaching, but also in the teaching of non-linguistic subjects. All over Europe, CLIL is considered 'a new learning and teaching environment' (Coonan, 2007, p. 625). In the last few decades, interest in CLIL, defined as the 'teaching and learning through a foreign language' (Marsh, 2002, p. 54), has gained momentum. In many countries, this methodology has been adopted in and adapted to different school settings. Pérez-Cañado (2012), who offers an overview of the literature on CLIL since the term was coined in Finland in the 1990s, has pointed out that in

M. Navarro-Pablo (✉)
University of Sevilla, Seville, Spain
e-mail: mnp@us.es

M. L. Pérez Cañado (eds.), *Content and Language Integrated Learning in Monolingual Settings*, Multilingual Education 38,
https://doi.org/10.1007/978-3-030-68329-0_7

recent years new studies on CLIL have started to emerge, offering sometimes contradictory and conflicting views on this methodological approach. One controversial issue pertains to the categorisation of CLIL. For example, not all researchers agree that this is an altogether new methodological approach, but rather consider it as having evolved from different communicative methodologies already in use in FL teaching (Content-Based Language Teaching, Communicative Language Teaching or the Natural Approach L+1 Hypothesis) (Cenoz et al., 2014).

However, CLIL was not originally intended as a definitive break from preceding methodologies. Rather, as several researchers contend (Lasagabaster & Sierra, 2009; Nikula, 2007), CLIL is used as an 'an umbrella term covering a dozen or more educational approaches' (Mehisto et al., 2008, p. 12). Furthermore, a commonality between CLIL and preceding methods for FL teaching, or the fact that CLIL may be considered a continuation thereof, is not necessarily a negative aspect of CLIL. As suggested by Dueñas (2004), CLIL is 'a flexible operational framework for language instruction' and this is precisely what allows this approach to be adapted to a great variety of contexts. In this line, Turner (2013, p. 397) indicates that 'this broad definition of CLIL allows for programmes that existed in different European countries to be redefined as CLIL programmes'. At the same time, there are also authors that identify some distinguishing features in CLIL. For example, Mehisto et al. (2008, p. 12) state that 'What is new about CLIL is that it synthesizes and provides a flexible way of applying the knowledge learnt from these various approaches'. In this line, Coyle (2008, p. 97) establishes that the distinctiveness of CLIL lies in its 'integrated approach, where both language and content are conceptualised on a continuum without an implied preference for either'. CLIL gives teachers the opportunity to introduce cross-curricular connections, meaningful interactions, cognitive skills training and a variety of cultural contexts in the subject content classroom, something that also goes beyond traditional non-linguistic content instruction. All these aspects that make CLIL different from other types of instruction need to be considered when analysing CLIL's potential language benefits for learners.

2 Prior Research on CLIL

Research on CLIL began in different European countries in the first decade of the twenty-first century and initially focused on comparative studies of CLIL and non-CLIL learners. More recently, there has been an increasing number of studies concerning the implementation and the effects of CLIL programmes. Research on the assessment of language learning and teaching programmes in CLIL contexts entails certain difficulties due to the variety of factors that influence final learning outcomes, and to the different ways in which this methodology has been implemented in different countries. According to Lasagabaster (2015, p. 19), 'as CLIL syllabuses are usually developed to meet local needs, there is huge variation in its implementation'. Also, Nikula (2007, p. 208) draws attention to the 'great deal of variation

in CLIL'. In the Finnish context, Nikula (2007) show how forms of implementation may vary in terms of both depth and breadth.

Table 1 sums up and classifies some of the literature produced on CLIL and allows us to narrow down and select the most relevant sources for the present study. The first section includes research focused on theoretical assumptions concerning CLIL conceptualisation. These studies provide a conceptual framework for bilingual and plurilingual education, bilingual programme design, policy issues from the different countries where CLIL programmes are implemented, and the main challenges to CLIL implementation (see Research type 1 in Table 1). From a more practical standpoint, several studies gather data about the implementation of CLIL programmes

Table 1 Types of research

Research type	Author/s, year of publication
Research type 1 (RT1): CLIL conceptualisation	Marsh (2002) Dueñas (2004) Nikula (2007) and Nikula et al. (2016) Coyle (2008) Mehisto et al. (2008) Turner (2013) Cenoz et al. (2014) Dalton-Puffer (2011, 2014) Pérez-Cañado (2012) Ruiz de Zarobe (2017)
Research type 2 (RT2): CLIL Implementation. Stakeholders' beliefs and opinions	Coonan (2007) Dalton-Puffer (2011) Pérez-Cañado (2014)
Research type 3 (RT3): CLIL and English learning outcomes	Admiraal et al. (2006) Ruiz de Zarobe (2008) Agustín-Llach (2009) Dalton-Puffer (2009) Fernández Fontecha (2010) Martínez and Juncal Gutiérrez (2009) Moreno (2009) Ojeda (2009) Ruiz de Zarobe and Jiménez-Catalán (2009) San Isidro (2010) Villarreal and García (2009) Navés and Victori (2010) Lasagabaster and Ruiz de Zarobe (2010) Llinares and Dafouz (2010) Breidbach and Viebrock (2012) Agustín-Llach and Canga (2014) Lasagabaster (2015) Juan-Garau et al. (2015) Roquet and Pérez-Vidal (2015) Pérez-Vidal and Roquet (2015) Rumlich (2016)

(continued)

Table 1 (continued)

Research type	Author/s, year of publication
Research type 4 (RT4): Individual and Contextual variables	Dalton-Puffer (2007) Lasagabaster and Sierra (2009) Lasagabaster (2011) Llinares et al. (2012) Martín del Pozo (2013, 2015) Fernández and Canga (2014) Dafouz and Hibler (2013) Doiz et al. (2014) Arribas (2016) Pfenninger (2016) Lorenzo (2017) Navarro-Pablo and García-Jiménez (2018) Fernández-Sanjurjo et al. (2019) Madrid and Barrios (2018) Rascón and Bretones (2018) Alejo and Piquer (2016) Pavón (2018) Lancaster (2018) Lorenzo et al. (2019)

in Europe and their strengths and weaknesses, either through participants' opinions, classroom observation or both (Research type 2). Most of these are qualitative studies in which interviews, questionnaires or observation tools are used to analyse the effectiveness of CLIL teaching performance or stakeholders' beliefs and opinions on the matter. Their findings usually provide positive views on language learning in CLIL programmes. The third group is made up of studies that analyse students' English learning outcomes, making use of quantitative data from students' results in language tests (Research type 3). Finally, there is an increasing number of studies that focus on individual and contextual variables that affect language learning, such as motivation or extramural exposure (see Research type 4).

Research on CLIL characterisation (RT1) helps to understand the differences between CLIL and non-CLIL instruction. Studies concerning CLIL teaching performance and stakeholders' beliefs and opinions (RT2) offer information about the way these programmes are being implemented, and their effect on the participants. Studies that analyse the implementation of CLIL programmes comparing the language results of CLIL and non-CLIL groups (RT3) are especially relevant for this study as they provide pertinent empirical evidence in the field. Finally, the present study also considers the possible influence of different individual and contextual variables (setting, socioeconomic status, verbal intelligence, motivational variables, extramural exposure and language needed and produced in class) on students' linguistic performance. Agreeing with many researchers (RT4) on the importance of moderating variables, this study includes a discriminant analysis in order to determine the influence of these variables on linguistic outcomes.

In general, research provides positive results for CLIL strands, where CLIL learners outperform non-CLIL ones. However, there are also some conflicting and not fully comparable findings, for which some factors are accountable, namely:

- Context (country and region; monolingual or bilingual context) and languages involved: it is necessary to study bilingual and monolingual regions independently, since their experience with bilingual and immersion programmes is so different that it can affect students' results and hence research findings. In Spain, where the present study takes place, it is necessary to bear in mind the local context, as there are bilingual regions, such as the Basque Country or Catalonia, in which the language introduced through CLIL (mostly English) is a third language, whereas in monolingual regions, English tends to be the second language.
- Educational level: results vary depending on whether research takes place in Primary, Secondary or Higher Education.
- Language tests and competence categorisation: when analysing and comparing students' results, the same types of tests need to be used in order to measure the same linguistic aspects. This has not always been the case, and researchers have measured students' linguistic competence using different types of tests, thus placing the focus on some linguistic aspects over others (for example, skills or specific language components, such as grammar or vocabulary).
- Type of study (cross-sectional vs. longitudinal study): studies have tended to focus on students' results at a given point in time. More recently, longitudinal studies are starting to gain ground. In this way, the effectiveness of CLIL programmes can be measured throughout time, offering a wider perspective on CLIL (Dalton-Puffer, 2011).
- Language needed and produced in CLIL classrooms: when analysing students' linguistic outcomes, teachers' metalanguage and classroom discourse functions are essential to the teaching and learning of content curricular subjects. The BICS and CALP distinction made by Cummins (2008), the *Language Triptych* (Language of Learning, Language through Learning and Language for Learning) identified by Coyle et al. (2010) and t*he L4C model* (General Language, Academic Language, Subject/Domani Specific Language and Classroom Language) by Gierlinger (2013) provided a theoretical framework for the analysis of the different languages involved in CLIL classroom.

Bearing these factors in mind, this chapter presents the results of a study carried out in a monolingual region (Andalusia) at both Primary and Secondary level. The present analysis focuses on quantitative data obtained from a linguistic test administered to 351 Primary and Secondary Education students (more information on the data collection tools in the Methodology section below). More specifically, it focuses on learners' results regarding two specific linguistic aspects: grammar and vocabulary.

The last decade has seen a proliferation of research that disaggregates and compares CLIL learners' performance in specific skills and/or mastery of linguistic

components, rather than just drawing on their overall achievement levels. Most findings indicate that when specific skills or linguistic aspects are analysed, or other variables (individual, motivational or contextual) are introduced, results are not consistent (Ojeda, 2009; Roquet & Pérez-Vidal, 2015; Ruiz de Zarobe, 2008). This may be due to the fact that these studies measure aspects such as vocabulary or grammar, which are generally emphasised in regular EFL courses, as opposed to the emphasis placed on communication in CLIL. In spite of this, there is a general consensus in the literature regarding CLIL programmes' positive effects on learners' language development over those of regular EFL programmes.

The literature review that is offered in the following lines focuses on two different types of studies: longitudinal and cross-sectional ones. Then, some studies dealing specifically with vocabulary and grammar are considered.

Longitudinal studies on the long-term benefits of CLIL instruction are emerging progressively (cf. also Chapter 8 in this volume). In this line, Pérez-Vidal and Roquet (2015) provide empirical data from two studies where writing, reading and listening skills, together with lexico-grammatical abilities, are examined. Their findings indicate that 'larger relative gains are obtained by the FI + CLIL programme, however not in all domains and to the same degree' (Pérez-Vidal & Roquet, 2015, p. 80). As regards students' lexico-grammatical abilities, CLIL learners show higher relative gains, whereas results regarding only the writing skill indicate that 'the superiority of CLIL cannot be confirmed', since 'although improvement in the case of the FI + CLIL group is shown, results were only significant in the domain of accuracy' (Pérez-Vidal & Roquet, 2015, p. 1). The findings of their quantitative data reveal that CLIL learners' progress in syntactic and lexical complexity as well as fluency is better, although the differences between CLIL and non-CLIL strands are not statistically significant. The same results are found through the qualitative data they collect for grammar and vocabulary. Similarly, Juan-Garau et al. (2015) also research the impact of CLIL programmes by analysing CLIL learners' development of lexico-grammatical accuracy over a period of three years throughout Secondary Education. They conclude that CLIL and non-CLIL learners 'significantly improved their overall longitudinal lexical and grammatical ability' (2015, p. 189) and their results also suggest that lexico-grammatical development is faster in CLIL learners.

Turning now to cross-sectional studies, Llinares and Dafouz (2010) describe the UAM-CLIL project carried out at secondary school level in Madrid, where three types of data were collected (whole class discussion, written composition and oral interview) from a corpus of approximately 40,000 spoken and 25,000 written words. Dealing with both lexical and grammatical competence, and comparing the same students' written and spoken performance, their study indicates that CLIL 'students use appropriate lexis to express content-specific ideas' (2010, p. 106). Similarly, San Isidro (2010) analyses language competence improvement in Secondary schools considering three main variables: student type, gender and school type. It includes specific sections measuring grammar and vocabulary development where 'the examiners assessed the ability to use vocabulary, structure and paraphrase strategies to convey meaning' (2010, p. 67). Findings reveal that CLIL learners 'were able not

only to pass an objective skills-based test but also with much better results' (2010, p. 67).

Examples of both longitudinal and cross-sectional studies carried out in Primary and Secondary Education are provided by Navés and Victori (2010). They present two studies as part of the BAF project that aim at examining the effect of onset age in the acquisition of English as a foreign language. The studies compare students' marks in two Primary and three Secondary schools in Catalonia. The first study focuses on learners' general language proficiency and includes students from Years 5, 7, 8 and 9 in CLIL and non-CLIL strands. Cross-sectional findings reveal that CLIL learners outperform non-CLIL learners on all the tests administered: grammar, cloze, dictation and listening. Furthermore, for all the measures analysed, longitudinal results indicate younger learners in CLIL strands obtain similar results to those of the older EFL students (2010, p. 47). The second study analyses learners' writing skills from Years 5 to 12. Similarly, there were statistical significant differences for each of the areas tested: fluency, accuracy and syntactic and lexical complexity. Again, CLIL learners at lower levels performed better in writing than the older EFL students. Likewise, Ruiz de Zarobe (2008) measures the oral competence of Secondary school students. In order to do so, five categories are established: pronunciation, vocabulary, grammar, fluency and content. On the one hand, the cross-sectional analysis reveals that CLIL students perform significantly better in all the scales. On the other hand, the longitudinal study reveals that, after one year, the CLIL group also outperforms the non-CLIL one in all the categories. After two years, however, differences between both groups are not statistically significant in the vocabulary category.

Finally, we focus on research dealing specifically with vocabulary and grammar. Regarding vocabulary, some studies of this kind are described in Ruiz de Zarobe and Jiménez-Catalán (2009). The sample in all of them is comprised of learners in Year 6 of Primary Education. As regards lexical competence, Jiménez-Catalán and Ruiz de Zarobe aim at establishing connections between the type of instruction and its effect on FL vocabulary acquisition, based on the hypothesis that 'the type of language instruction relates positively to vocabulary knowledge' (2009, p. 82). The study contrasts the results of CLIL and non-CLIL groups in two receptive vocabulary tests, and concludes that, in both tests, CLIL learners outperform their non-CLIL counterparts. A further study is provided by Moreno (2009), who analyses the results of a free word association task as a means to explore learners' mental lexicon, assuming that this type of data can 'complement and corroborate findings that emerge from analyses of other lexical data' (2009, p. 93). One of the main purposes of this study is to describe the characteristics of the productive lexical profile of EFL Spanish learners in Primary Education, comparing both CLIL and non-CLIL instructional models. Findings show that CLIL learners produce more tokens and types than non-CLIL learners, which is indicative that they have a higher proficiency level. Findings also reveal that CLIL learners exhibit 'a slightly higher productive vocabulary size' and 'more lexical richness in the word association test [. . .] by recalling a higher number of infrequent words' (2009, p. 100). Thus, the study shows that statistically significant differences between CLIL and non-CLIL groups apply to both vocabulary size and vocabulary depth. However, the author also contends that CLIL learner results

'are not so overwhelming if we compare the great difference of formal instruction exposure between groups' (2009, p. 106).

In a similar line, the role of the L1 (Spanish) in CLIL and non-CLIL learners' FL vocabulary use is analysed by Agustín-Llach (2009). Taking into account variables such as students' proficiency levels, amount of FL exposure and instructional approach, the aim of the study is to reach conclusions regarding which group of students (CLIL or non-CLIL) has more transfer episodes from Spanish to English. In order to do so, three categories are established: borrowings, coinages and calques. Findings reveal that non-CLIL learners transfer from their L1 more frequently than their CLIL counterparts, showing a higher number of lexical errors for all categories distinguished. In another study, Agustín-Llach and Canga (2014) perform a cross-sectional and longitudinal analysis of Primary school learners' FL receptive vocabulary size and lexical growth. Whereas in the cross-sectional analysis, CLIL learners have a slightly larger vocabulary size, the longitudinal study reveals that differences between groups are not significant in the early years, and become progressively significant in the later years. Therefore, the study points to growing differences in favour of CLIL students as learners' educational and proficiency levels increase.

Also, Ojeda (2009) provides an interesting study on vocabulary and themes, drawing attention to the importance of considering the differences between the CLIL methodology and regular FL approaches, with special emphasis on CLIL and non-CLIL learners' view of the target language. In this regard, learners following an FL programme conceive the target language 'as a single object for language learning' (2009, p. 130) where the language is organised around linguistic components (grammar and vocabulary among others), whereas in CLIL language instruction is organised around non-linguistic topics and lessons, being the FL a vehicular language used for communication and for conveying meaning. The study analyses learners' written compositions in order to compare 'the vocabulary most frequently implemented by the two samples of participants' (2009, p. 132). The author draws upon the similarities and differences found in a total of 60 comparable texts following a lexical field theory where taxonomy serves to classify and organise the lexis found in the corpus. Findings are mixed concerning the similarities and discrepancies of both groups, depending on the lexical field analysed. In the interpretation of the results, the author suggests that the 'CLIL sample seems to have a slight tendency to use a wider range of types including both colloquial and even sophisticated words' (2009, p. 137) and refers to the 'non-CLIL sample's greater difficulties to express abstract ideas that entail a higher degree of complexity' (2009, p. 152). At the same time, Ojeda points to the non-CLIL group's 'higher lexical reiteration' (2009, p. 140) and CLIL learners' 'higher lexical variation' (2009, p. 153).

There are also some qualitative studies where stakeholders confirm the results reported in quantitative research concerning vocabulary improvement in CLIL programmes. The results in Pérez-Cañado (2014) are largely consistent with findings reported in Juan-Garau et al. (2015). Pérez-Cañado analyses data gathered from questionnaires for a European study in which in-service teachers across Europe provide their perceptions on teacher training needs for bilingual education. One of the thematic blocks considered in the study is related to participants' current level of

linguistic and intercultural competence. Findings reveal that 'all the items comprised within linguistic and intercultural competence are invariably considered to be appropriately mastered' and, more interestingly for our study, that 'this is especially the case for accurate pronunciation and knowledge of specialized academic vocabulary (within linguistic competence)' (2014, p. 11). Although studies carried out so far generally provide quite positive results for CLIL learners' gains, research also points to the impossibility of ascertaining that findings are due to the type of instruction, rather than, for example, to the increased number of hours that CLIL programmes imply (Juan-Garau et al., 2015; Ruiz de Zarobe, 2017).

However, not all the studies show such satisfactory results. For instance, Admiraal et al. (2006), who carry out an evaluation of bilingual education at Secondary school level in The Netherlands, administer a test that specifically measures receptive vocabulary, where findings show that there are no significant differences between bilingual and non-bilingual groups. Similarly, Fernández Fontecha (2010) provides results from a lexical availability task administered to Year 6 Primary school learners in which non-CLIL students outperform CLIL ones. The findings are interpreted as due to factors such as 'the type of test used, which requires that the learners produce types in a limited amount of time not in a communicative interaction, which is more typical of a CLIL environment [. . .] or the early stage of CLIL instruction at which learners had been tested' (2009, p. 87).

Turning now to research specifically focused on grammar development, Breidbach and Viebrock (2012) review recent CLIL research in Germany. Especially relevant for the present study is Berenbröker, as described by Breidbach and Viebrock (2012), where a comparative study between 195 CLIL and non-CLIL learners over a period of two years shows that, whereas CLIL has a very positive influence on FL competence in general, as regards grammar, differences are less accentuated. According to the study, this is due to the fact that 'regular foreign language teaching is often more concerned with an explicit focus on grammar, whereas CLIL is more concerned with implicit grammatical knowledge, which is acquired in the process of exchanging subject-specific information' (2000, p. 7). In turn, tense and agreement morphology in Secondary Education is analysed through a collection of oral narrations by Villarreal and García (2009). They compare affixal forms against suppletive forms. They find that the omission rate is very high across both groups of learners (CLIL and non-CLIL), which implies a parallel behaviour of the groups taken independently. However, when their overall performance is contrasted, the CLIL group outperforms the non-CLIL group in the production of affixal morphemes. Quite similarly, Martinez and Gutiérrez (2009) study the acquisition of syntax, also through the analysis of Secondary students' oral narrations. In order to do so, they select several morpho-syntactic features, null subjects, production of placeholders, negation and production of null objects, and conclude, 'CLIL learners significantly outperform non-CLIL learners only in the use of placeholders' (2009, p. 193).

As has been shown, findings regarding specific linguistic aspects are not clear-cut. Indeed, Juan-Garau et al. (2015) state that 'no conclusive results have so far been obtained regarding the development of lexico-grammatical competence in CLIL

contexts' (2015, p. 182). At the same time, Pérez-Vidal and Roquet (2015, p. 81) contend that:

> [G]eneral results [concerning the linguistic benefits of CLIL] seem to be by and large positive, although there are aspects which are either unaffected by CLIL or for which research is inexistent or inconclusive, namely syntax, productive vocabulary, written accuracy, discourse skills and pragmatic efficiency (see Llinares et al. 2012), and pronunciation, that is, degree of foreign accent. Such a positive impact has generally being attributed to higher quantity and quality of exposure. However, methodological issues are still unresolved in CLIL research and subject to debate.

The following conclusions may be extracted from this literature review:

- More research is needed in order to shed light on findings that are apparently contradictory or inconclusive. Any teaching approach needs time and fine-tuned research so that its theoretical bases can feed the results of research and produce better learning outcomes.
- CLIL is an approach that seeks contextualised learning based on meaning construction through the use of the target language. It is also a methodology that is eminently grounded on interaction and communication. Therefore, research instruments that measure learners' mastery of specific linguistic systems (typically trained in regular EFL courses) are likely to yield less positive results in CLIL groups. In view of this, specific instruments measuring learners' communicative competence and contextualised learning should be used.
- In order to design accurate research tools, it is also necessary to take into account the characteristic tasks of each instructional approach. In this sense, CLIL learners' vocabulary seems to be best measured by means of integrative vocabulary tests and word association tests than by discrete decontextualised receptive vocabulary tests.
- Different personal and contextual factors must be considered in any assessment of CLIL, as these have an effect on and may help explain learners' results.

3 The Study

Research Questions The aims of the present study are: (1) to examine the impact of CLIL on the English grammar and vocabulary of 351 Primary and Secondary school students and (2) to investigate the relationship between individual difference and contextual variables in order to determine which of them has a stronger influence on students' linguistic outcomes. The effects of the following intervening variables are analysed: setting, socioeconomic status, verbal intelligence, motivational variables and extramural exposition. Considering these aims, this study seeks to answer the following research questions:

- RQ1. Are there statistically significant differences between the achievement levels of CLIL and non-CLIL learners concerning grammar and vocabulary? If so, what is the effect size?

- RQ 2. Are there statistically significant differences between the achievement levels of CLIL and non-CLIL learners concerning individual differences and contextual variables? If so, which variable has a stronger influence on students' linguistic outcomes?

Scope In order to guarantee the homogeneity of the sample, three actions were undertaken: first, the researchers contacted the provincial coordinator of bilingual programmes in order to request a list of the state schools with English bilingual programmes. The schools selected for the sample have both CLIL and non-CLIL (or regular EFL) groups, which acted as experimental and control groups, respectively. Thus, participants in this study are streamed into two different instruction types: students enrolled in CLIL programmes and students who follow an EFL approach. Secondly, verbal intelligence and motivation tests were applied to each group. Finally, information regarding students' socioeconomic status, their English grades and their extramural exposure to English were also collected. The results of these actions allowed us to match students and ensure that these factors did not interfere with the results of the study. At the same time, this also allowed us to determine whether the differences in language attainment could be ascribed to the programme implementation rather than to any other factor related to the students' initial capacities, motivation or any other contextual variables. A total of eight schools participated in the study (four Primary schools and four Secondary schools), of which seven are state schools, and one is a charter school. From the overall total of 351 students, 193 are in Year 6 of Primary Education (in the age range of 11–12) and 158 are in the final year of Compulsory Secondary Education.

Instruments A total of four instruments were administered to the students: three tests (English, verbal intelligence and motivation tests) and a questionnaire. The language tests were designed following the *Common European Framework of Reference for Languages* (*CEFR*), the national Decrees and the regional Orders which establish the official curriculum for the educational stages assessed. The sections of the tests measuring learners' grammatical competence, both at Primary and Secondary level, combine traditional formal activities with exercises that require understanding meaning within a context. Similarly, the tests designed to measure students' lexical competence combine activities focused on form with exercises based on texts in which meaning has to be reconstructed from the context. Thus, in both instances, tests are suitable to assess lexico-grammatical competences in methodological approaches, CLIL and regular EFL courses. Verbal intelligence was measured by means of two different adapted versions (one for Primary students and one for Secondary students) of the EFAI (*Evaluación Factorial de las Aptitudes Intelectuales*) test (Santamaría, 2018). Both versions include analogies, antonyms and odd-one-out questions, of which students have to answer as many as possible in five minutes. In order to measure motivation, Pelechano's (1994) MA test was used. This test is composed of 35 items that isolate four motivational factors related to achievement and anxiety: (1) desire to work and self-esteem (comprising 10 items); (2) anxiety in the face of exams (composed of 9 items); (3) lack of interest in studying (made up of 9 elements); and (4) realistic personal self-demand (comprising

7 elements). An initial questionnaire provided personal data and information on students' socioeconomic status and extramural exposure to English.

Data analysis The data collected have been statistically analysed using SPSS 24.0. In order to answer RQ1, the t-test was used to identify significant differences between the two groups under study. Also, the effect size was calculated through Cohen's d coefficient. Finally, a discriminant analysis has been applied in order to address RQ2, as it can be considered a powerful technique for examining differences between the two groups with respect to several variables simultaneously. The grouping variable selected for the data analysis is the type of instructional programme followed by the students (CLIL vs. non-CLIL). The dependent variable is the students' results in both the grammar and vocabulary tests. The independent variables are setting, socioeconomic status, verbal intelligence, motivational variables and extramural exposure.

4 Results

Regarding the first research question, Table 2 presents the means scores, standard deviations and Cohen's d coefficient with the effect size for the grammar and vocabulary variables. Findings show that there are significant differences between CLIL and non-CLIL groups, both at Primary and Secondary level, in favour of CLIL groups. In Primary Education, the effect size is small (Cohen's d = 0.336) for grammar and medium (Cohen's d = 0.504) for vocabulary. Cohen's d sizes indicate that there are standard deviations of 0.336 and 0.504 between CLIL and non-CLIL groups, respectively. Both are considerably higher in Secondary Education, where both grammar and vocabulary differences show a large effect size (Cohen's d = 1.150 and 0.858), with standard deviations of 1.150 and 0.588 for CLIL and non-CLIL groups, respectively.

If we now take a closer look at the exercise type students were asked to answer, the t-test was used in order to determine whether learners' results vary depending on the task they had to carry out. In Primary Education (see Table 3), CLIL learners

Table 2 Means, standard deviations, Cohen's d coefficient and effect size

Educational Level	Independent variables	Non-CLIL		CLIL			
		Mean	SD	Mean	SD	Cohen's d	Effect size
Year 6 Primary Education	Grammar	9.084	5.600	10.900	5.208	0.336	0.166
	Vocabulary	5.930	3.309	7.567	3.180	0.504	0.245
Year 4 Secondary Education	Grammar	16.857	10.587	28.712	10.0104	1.150	0.499
	Vocabulary	6.825	3.420	9.661	3.187	0.858	0.394

SD standard deviation

Table 3 T-test exercise type. Primary Education

Exercises	Group	N	Mean	Sig. (bilateral)
A. Write questions for these answers	CLIL	35	2.89	.045[a]
	Non-CLIL	59	2.19	
B. Look at the picture on the right. Read and complete. Use: ON, IN, BY, WITH, UNDER	CLIL	59	2.44	.024[a]
	Non-CLIL	115	1.90	
C. Match the numbers with the letters	CLIL	61	3.00	.175
	Non-CLIL	121	2.61	
D. Complete using: MY, ITS, HIS, THEIR	CLIL	58	2.55	.006[a]
	Non-CLIL	120	1.89	
E. Complete using: DO, DOES, DID, AREN'T, ISN'T, IS	CLIL	56	3.27	.007[a]
	Non-CLIL	117	2.45	
F. Complete using: FEEDS, CHASES, DRIVES, GETS UP	CLIL	49	2.96	.046[a]
	Non-CLIL	116	2.50	
G. Write the parts of the body	CLIL	34	4.62	.099
	Non-CLIL	26	3.96	
H. Read and complete. Use: A GIRL, RIGHT, HERE, LOVELY, NAME	CLIL	53	3.96	.002[a]
	Non-CLIL	102	3.10	

([a]p = 0.05)

outperform their peers on all exercises. From the eight exercises testing lexical and grammatical competence, there are six in which statistically significant differences are found, and only two where differences are not statistically significant. Interestingly, neither of these involves the use of contextualised vocabulary within a meaningful text.

In Secondary Education, our results for grammar are consistent with those obtained at Primary level, as CLIL learners also show better results, with statistically significant differences on all the activities (see Table 4). However, regarding vocabulary, although CLIL learners also outperform non-CLIL learners, differences between both groups are not statistically significant.

Turning now to research question 2, Table 5 shows the results of the test of equality of CLIL and non-CLIL group means in Primary Education. It allows us to examine whether significant differences exist between the groups, in terms of predictor variables. Wilks' lambda reveals that the discriminant function is statistically significant only for three variables: setting (F = 10.897 and *p*-value = 0.001), vocabulary (F = 5.279 and *p*-value = 0.024) and verbal intelligence (F = 4.069 and *p*-value = 0.046). In the case of these variables, the null hypothesis is rejected (*p*-value < 0.05). The setting and verbal intelligence variables show higher means for non-CLIL groups; however, the results for grammar and vocabulary are better for CLIL groups. There are no statistically significant differences for the rest of the variables. Results from Box's test of Equality of Covariance Matrices (F = 1.398, *p*-value = 0.012 < 0.05)

Table 4 T-test exercise type. Secondary Education

Exercises	Group	N	Mean	Sig. (bilateral)
A. Rewrite these questions using the correct word order	CLIL	61	2.1311	.001[a]
	Non-CLIL	81	1.3951	
B. Change these sentences into the active. Example: Pictures are transmitted by television. Television transmits picture	CLIL	51	2.2353	.044[a]
	Non-CLIL	57	1.6491	
C. Complete the text with the verbs in the past tense	CLIL	54	8.1296	.003[a]
	Non-CLIL	66	6.7879	
D. Complete these sentences with: anything/something; tell/say (in the right form); bring/take (in the right form)	CLIL	51	6.4706	.000[a]
	Non-CLIL	56	4.2321	
E. Underline the right words in brackets	CLIL	64	6.3438	.001[a]
	Non-CLIL	89	5.3034	
F. Complete this text with the correct forms—the PAST (for example, went) or the PRESENT PERFECT (for example, has gone)—of the verbs in brackets	CLIL	46	5.2391	.026[a]
	Non-CLIL	61	4.3115	
G. Complete this text using one of the following words and expressions: ON YOUR OWN, OVERCOME, QUIT, CUT DOWN, IMPROVE, HEALTH, WILL POWER, HARM	CLIL	46	3.3261	.086
	Non-CLIL	50	2.6200	
H. Match symbol and text	CLIL	61	7.3279	.000[a]
	Non-CLIL	80	6.3250	

([a]p = 0.05)

Table 5 Tests of equality of group means. Primary Education

Variables		Wilks' Lambda	F	Df1	Df2	Sig.
Setting		.901	10.897	1	99	.001[a]
Socioeconomic status		.999	.141	1	99	.709
Verbal intelligence		.961	4.069	1	99	.046[a]
Motivation	Desire to work and self-esteem	.990	.980	1	99	.325
	Anxiety	.996	.411	1	99	.523
	Lack of interest	.978	2.254	1	99	.136
	Self-demand	.995	.457	1	99	.501
Extramural exposition		.988	1.237	1	99	.269
Language test	Grammar	.977	2.308	1	99	.132
	Vocabulary	.949	5.279	1	99	.024[a]
	Listening	.998	.246	1	99	.621
	Reading	.997	.299	1	99	.586
	Total	.983	1.693	1	99	.196

([a]p = 0.05)

Table 6 Classification results. Year 6 Primary Education

Group	Predicted group		Total
	Non-CLIL	CLIL	Non-CLIL
Non-CLIL	55	16	71
CLIL	9	21	30
Non-CLIL	77.5	22.5	100.0
CLIL	30.0	70.0	100.0

established that equal matrices of variances are rejected; therefore, the groups do not have the same variance matrix.

The discriminant analysis shows a canonical correlation of 0.507 with an eigenvalue of 0.346 and a statistical Wilks' lambda of 0.743, with 12 degrees and *p*-value = 0.006 (< 0.05), which leads us to reject the null hypothesis of equality of means and indicates the existence of a discriminant function that separates CLIL and non-CLIL groups significantly and accounts for 25% of the variance observed in their scores.

The standardised coefficients indicate that vocabulary is the best discriminating variable (−0.928), followed by the variable 'setting' (0.567).

The analyses also evaluate the accuracy of the classification. Table 6 shows that measures resulted in a fairly positive classification for students belonging to their corresponding groups. 75.2% of original grouped cases are correctly classified.

As regards Secondary Education, Table 7 shows the results of discriminant function analyses where Wilks' lambda reveals that the discriminant function is statistically significant to classify results obtained regarding the following six variables:

Table 7 Tests of equality of group means. Secondary Education

Variables		Wilk's Lambda	F	Df1	Df2	Sig.
Setting		.997	.403	1	120	.527
Socioeconomic status		.983	2.055	1	120	.154
Verbal intelligence		.957	5.358	1	120	.022[a]
Motivation	Desire to work and self-esteem	.978	2.674	1	120	.105
	Anxiety	.999	.075	1	120	.785
	Lack of interest	.957	5.350	1	120	.022[a]
	Self-demand	.988	1.441	1	120	.232
Extramural exposition		.992	.994	1	120	.321
Language test	Grammar	.749	40.261	1	120	.000[a]
	Vocabulary	.843	22.367	1	120	.000[a]
	Listening	.881	16.276	1	120	.000[a]
	Reading	.868	18.259	1	120	.000[a]
Total		.746	40.923	1	120	.000[a]

([a]p = 0.05)

Table 8 Classification results. Year 4 Secondary Education

Group	Predicted group		Total
	Non-CLIL	CLIL	Non-CLIL
Non-CLIL	48	15	63
CLIL	18	41	59
Non-grouped cases	1	0	1
Non-CLIL	76.2	23.8	100.0
CLIL	30.5	69.5	100.0
Non-grouped cases	100.0	.0	100.0

verbal intelligence (F = 5.358 and *p*-value = 0.022); lack of interest (F = 5.350 and *p*-value = 0.022); grammar (F = 40.261 and *p*-value = 0.000); vocabulary (F = 22.367 and *p*-value = 0.000); listening (F = 16.276 and *p*-value = 0.000); and reading (F = 18.259 and *p*-value = 0.000). Results from Box's test of Equality of Covariance Matrices (F = 0.987, *p*-value = 0.510) indicate that the null hypothesis is accepted; therefore, equal matrices of variances are assumed. There is a canonical correlation of 0.521 with an eigenvalue of 0.373 and a statistical Wilks' lambda (= 0.728) with 12 degrees of freedom of Chi-square and a critical significance level of 0.000 (< 0.05). This allows us to reject the null hypothesis and accept that at least one discriminant axis is significant. In this case, the standardised coefficients of the canonical discriminant functions indicate that the variable that has the greatest influence on the calculation of the function is grammar (0.787).

Finally, in the same way as with Primary Education, Table 8 shows a positive classification, as 73% of the original grouped cases are correctly classified.

5 Discussion

Positive results for Primary Education in favour of CLIL learners are in line with those presented in Navés and Victori (2010), where there are statistically significant differences between both groups on all the tests performed. In Secondary Education, our results coincide with Ruiz de Zarobe (2017), in which positive findings report greater lexical and syntactic complexity in CLIL learners. However, those results only coincide partially with Pérez-Vidal and Roquet (2015), in which CLIL learners outperform non-CLIL learners but without showing statically significant differences. With respect to the differences found between Primary and Secondary Education, our results are consistent with Agustín-Llach and Canga (2014), Garau et al. (2015) and Navés and Victori (2010). Their longitudinal studies show that, over time, CLIL learners improve their competence, giving rise to progressively larger differences between both groups. Although our study is not longitudinal, the fact that the effect size increases from Primary to Secondary Education can be related to the results presented in these longitudinal studies.

With respect to the differences found regarding the type of exercise included in the tests, in Primary Education our results are consistent with Jiménez-Catalán and Ruiz de Zarobe (2009), where CLIL learners also perform better than non-CLIL learners. However, in their study, there are statistically significant differences on all the tests performed, whereas in our study there are two exercises for which differences between both groups are not statistically significant. This may be due to the fact that these are exercises that do not require the use of lexico-grammatical elements within a meaningful context. The main implication of these results is that the type of instruction plays a role in the results. It is necessary to bear in mind that Jiménez-Catalán and Ruiz de Zarobe's study only focuses on receptive vocabulary and that their tests are different from ours, which may also account for the discrepancies in the results. There are also parallelisms between our study and Moreno's one (2009), in which there are statistically significant differences between CLIL and non-CLIL groups which apply not only to vocabulary size but also to vocabulary depth, pointing to the type of instruction as responsible for qualitative (and not only quantitative, as previous studies contend) differences in lexical competence. Also, our results may be related to the ones provided in Agustín-Llach (2009), where CLIL learners transfer from their L1 less frequently than their non-CLIL counterparts, showing a lower number of lexical errors for all the categories analysed. Agustín-Llach explains these differences alluding to the role of the target language and the way learners perceive it. Thus, the instructional programme is responsible for the differences insofar as it makes learners perceive the target language as a means of communication (CLIL instruction) or as merely a school subject (non-CLIL instruction): for CLIL students, 'the text becomes an exercise of communication rather than a language task' (124). In this sense, Ojeda (2009) also suggests the importance of the instructional programme in relation to how it makes students perceive their learning. Finally, our results confirm those in Admiraal, Westhoff and de Bot (2006), where the type of test administered is presented as accountable for CLIL's negative results: we agree with Admiraal, Westhoff and de Bot that the exercises used in tests should be adapted to both the instructional approach and the activities that are used in the classroom. Regarding Secondary Education, this study coincides with Ruiz de Zarobe (2008), where CLIL learners outperform non-CLIL learners in all scales measured. However, our results regarding vocabulary are not as satisfactory, due to the fact that differences between both groups are not statistically significant in any of the exercises. Again, regarding vocabulary, these results contradict those found in San Isidro (2010), where CLIL students show more strategies to convey meaning, and are more consistent with Ruiz de Zarobe (2008), where vocabulary is the only category in which Secondary students do not improve over time, and the only category for which differences between groups are not statistically significant. On the other hand, regarding grammar, our results do not coincide with Breidbach and Viebrock (2012), in which grammatical differences are less pronounced. Both groups of CLIL learners show a better grammatical competence with significant differences in both the global findings.

6 Conclusions

This study offers new empirically grounded insights into the current state of CLIL implementation and the effects of CLIL on students' language attainment. As regards RQ1, the results obtained complement previous research by offering CLIL outcomes regarding the impact of CLIL on the English grammar and vocabulary of 351 Primary and Secondary school students. In this respect, this study has shown that both at Primary and Secondary levels, there are statistically significant differences between CLIL and non-CLIL learners, in favour of the CLIL groups. Results on vocabulary and grammar show different effect sizes in Primary Education, being small for grammatical competence and medium for vocabulary. These differences in effect size increase in Secondary Education, which is indicative of students' improvement over time.

The results reported in this chapter also indicate that the difference between both groups of informants lies not only in language proficiency as reported by better overall results regarding lexical and grammatical competence, but also in the type of instruction as indicated by the comparison of results obtained in each exercise of the tests. Thus, this chapter has drawn attention to the central importance of considering the type of test administered in connection with the type of instruction implemented.

As far as RQ2 is concerned, this study has also investigated second-order interactions of individual difference variables and linguistic and contextual variables. The discriminant analyses evince different discriminant functions depending on the educational level under analysis. Therefore, it seems that, in general, it is important to contextualise findings, since individual and contextual variables do not have the same influence in Primary and Secondary Education.

On the one hand, in Primary Education, setting, verbal intelligence and vocabulary are the variables that display the greatest significance in the test of equality of group means. Vocabulary is the variable that best explains the statistically significant differences found between the groups. On the other hand, in Secondary Education, results show that, as it happens at Primary level, verbal intelligence carries a significant weight in explaining the differences between the groups. However, there are other variables that display significance in the test of equality of group means: lack of interest, vocabulary, grammar, listening and reading. At Secondary level, grammar is the variable that best explains the differences between the groups.

Taken together, these results suggest that, over the rest of the variables considered in this study, vocabulary and grammar are the variables that have the greatest influence on the calculation of the discriminant function. One of the most significant findings of this study is the fact that it confirms the effectiveness of the CLIL approach as far as students' language outcomes are concerned, providing better results even for the development of vocabulary and grammar, in spite of the importance that is traditionally given to them in FL instructional programmes.

Acknowledgements This work was supported by the Spanish Ministry of Economy and Competitiveness [research grant number FFI2012-32221] and the Government of Andalusia [research grant number HUM 2348] (Project: *The Effects of Content and Language Integrated Learning in Monolingual Communities: A Large-Scale Evaluation*).

References

Admiraal, W., Westhoff, G., & de Bot, K. (2006). Evaluation of bilingual Secondary Education in the Netherlands: Students' language proficiency in English. *Educational Research and Evaluation, 12*(1), 75–93.

Agustín-Llach, M. P. (2009). The role of Spanish L1 in the vocabulary use of CLIL and non-CLIL EFL learners. In Y. Ruiz de Zarobe & R. Jiménez-Catalán (Eds.), *Content and Language Integrated Learning: Evidence from research in Europe* (pp. 112–129). Bristol: Multilingual Matters.

Agustín-Llach, M. P., & Canga, A. (2014). Vocabulary growth in young CLIL and traditional EFL learners: Evidence from research and implications for education. *International Journal of Applied Linguistics, 26*(2), 211–227.

Alejo, R., & Piquer, A. (2016). Urban vs. rural CLIL: An analysis of input-related variables, motivation and language attainment. *Language, Culture and Curriculum, 29*(3), 245–262.

Arribas, M. (2016). Analysing a whole CLIL school: Students' attitudes, motivation, and receptive vocabulary outcomes. *Latin American Journal of Content and Language Integrated Learning, 9*(2), 267–292. https://doi.org/10.5294/la-clil.2016.9.2.2.

Breidbach, S., & Viebrock, B. (2012). CLIL in Germany—Results from recent research in a contested field of education. *International CLIL Research Journal, 1*(4), 5–16.

Cenoz, J., Genesee, F., & Gorter, D. (2014). Critical analysis of CLIL: Taking stock and looking forward. *Applied Linguistics, 35*(3), 243–262.

Coonan, C. (2007). Insider views of the CLIL class through teacher self-observation-introspection. *International Journal of Bilingual Education and Bilingualism, 10*(5), 625–646.

Coyle, D. (2008). Content and Language Integrated Learning: Towards a connected research agenda for CLIL pedagogies. *International Journal of Bilingual Education and Bilingualism, 10*(5), 543–562.

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

Cummins, J. (2008). BICS and CALP: Empirical and theoretical status of the distinction. In B. Street & N. H. Hornberger (Eds.), *Encyclopedia of language and education* (pp. 71–83). New York: Springer Science+Business Media LLC.

Dafouz, E., & Hibler, A. (2013). 'Zip your lips' or 'keep quiet': Main teachers' and language assistants' classroom discourse in CLIL settings. *The Modern Language Journal, 97*(3), 655–669.

Dalton-Puffer, C. (2007). *Discourse in Content and Language Integrated Learning (CLIL) classrooms*. Amsterdam/Philadelphia: John Benjamins Publishing Company.

Dalton-Puffer, C. (2009). Communicative competence and the CLIL lesson. In Y. Ruiz de Zarobe & R. Jiménez-Catalán (Eds.), *Content and Language Integrated Learning: Evidence from research in Europe* (pp. 197–214). Bristol: Multilingual Matters.

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

Dalton-Puffer, C. (2014). "You can stand under my umbrella": Immersion, CLIL and bilingual education. A response to Cenoz, Genesee, and Gorter. 2013. *Applied Linguistics, 35*(2), 213–218.

Doiz, A., Lasagabaster, D., & Sierra, J. M. (2014). CLIL and motivation: The effect of individual and contextual variables. *Language Learning Journal, 42,* 209–224.

Dueñas, M. (2004). The whats, whys, hows and whos of content-based instruction in second/foreign language education. *International Journal of English Studies, 4*(1), 73–96.

European Union. (2008). Council conclusions of 22 May 2008 on multilingualism. *EUR_Lex Official Journal of the European Union*, 2008/C 140/10. https://eur-lex.europa.eu/oj/direct-access.html. Accessed 22 November 2020.

Fernández, A., & Canga, A. (2014). A preliminary study on motivation and gender in CLIL and non-CLIL tupes of instruction. *International Journal of English Studies, 14*(1), 21–36.

Fernández-Sanjurjo, J., Fernández-Costales, A., & Arias Blanco, J. M. (2019). Analysing students' content-learning in science in CLIL vs. non-CLIL programmes: Empirical evidence from Spain. *International Journal of Bilingual Education and Bilingualism, 22*(6), 661–674. https://doi.org/10.1080/13670050.2017.1294142.

Fontecha, A. F. (2010). First steps of CLIL in Spanish monolingual community: The case of La Rioja. In D. Lasagabaster & Y. Ruiz de Zarobe (Eds.), *Spain: Implementation, results and teacher training* (pp. 79–94). Newcastle: Cambridge Scholars.

Gierlinger, E. (2013). *The languages for CLIL model*. http://clilingmesoftly.wordpress.com/clil-teachers-tl-competence/. Accessed 22 November 2020.

Jiménez-Catalán, R., & Ruiz de Zarobe, Y. (2009). The receptive vocabulary of EFL learners in two instructional contexts: CLIL versus non-CLIL instruction. In Y. Ruiz de Zarobe & R. Jiménez-Catalán (Eds.), *Content and Language Integrated Learning: Evidence from research in Europe* (pp. 81–92). Bristol: Multilingual Matters.

Juan-Garau, M., Prieto-Arranz, J., & Salazar, J. (2015). Lexico-grammatical development in secondary education CLIL learners. In M. Juan-Garau & J. Salazar (Eds.), *Content-based language learning in multilingual educational environments* (pp. 179–195). Educational Linguistics. Springer International Publishing Switzerland. https://doi.org/10.1007/978-3-319-11496-5_11.

Lancaster, N. (2018). Extramural exposure and language attainment: The examination of input-related variables in CLIL programmes. *Porta Linguarum, 29,* 91–114.

Lasagabaster, D. (2011). English achievement and student motivation in CLIL and EFL settings. *Innovation in Language Learning and Teaching, 5*(1), 3–18.

Lasagabaster, D. (2015). Different educational approaches to bi-or multilingualism and their effect on language attitudes. In M. Juan-Garau & J. Salazar (Eds.), *Content-based language learning in multilingual educational environments.* Educational Linguistics. Springer International Publishing Switzerland. https://doi.org/10.1007/978-3-319-11496-5_2.

Lasagabaster, D., & Sierra, J. M. (2009). Language attitudes in CLIL and traditional EFL classes. *International CLIL Research Journal, 1*(2), 4–17.

Lasagabaster, D., & Ruiz de Zarobe, Y. (Eds.). (2010). *CLIL in Spain: Implementation, results and teacher training*. Newcastle: Cambridge Scholars.

Llinares, A., & Dafouz, E. (2010). Content and language integrated programmes in the Madrid region: Overview and research findings. In D. Lasagabaster & Y. Ruiz de Zarobe (Eds.), *Spain: Implementation, results and teacher training* (pp. 95–114). Newcastle: Cambridge Scholars.

Llinares, A., Morton, T., & Whittaker, R. (2012). *The roles of language in CLIL.* Cambridge University Press.

Lorenzo, F. (2017). Historical literacy in bilingual settings: Cognitive academic language in CLIL history narratives. *Linguistic and Education, 37,* 32–41.

Lorenzo, F., Granados, A., & Ávila, I. (2019). The development of cognitive academic language proficiency in multilingual education: Evidence of a longitudinal study on the language of history. *Journal of English for Academic Purposes*, 41. https://doi.org/10.1016/j.jeap.2019.06.010.

Madrid, D., & Barrios, E. (2018). A comparison of students' educational achievemnet across programmes and school types with and without CLIL provision. *Porta Linguarum, 29,* 29–50.

Marsh, D. (Ed.). (2002). *CLIL/EMILE. The European dimension: Actions, trends, and foresight potential*. Jyväskylä: University of Jyväskylä.

Martín del Pozo, M. A. (2013). Formación lingüística del profesorado universitario para la docencia en inglés. *Revista de Docencia Universitaria. REDU, 11*(3), 197–218.

Martín del Pozo, M. A. (2015). Signposts for comprehensive knowledge of CLIL contexts. *Procedia—Social and Behavioral Sciences, 212,* 99–105.

Martínez, M., & Juncal Gutiérrez, M. (2009). The acquisition of English syntax by CLIL learners in the Basque Country. In Y. Ruiz de Zarobe & R. Jiménez-Catalán (Eds.), *Content and Language Integrated Learning: Evidence from research in Europe* (pp. 176–196). Bristol: Multilingual Matters.

Mehisto, P., Marsh, D., & Frigols, M. J. (2008). *Uncovering CLIL: Content and Language Integrated Learning in bilingual and multilingual education*. Oxford: Macmillan Education.
Moreno, S. (2009). Young learners' L2 word association responses in two different learning contexts. In Y. Ruiz de Zarobe & R. Jiménez-Catalán (Eds.), *Content and Language Integrated Learning: Evidence from research in Europe* (pp. 93–111). Bristol: Multilingual Matters.
Navarro-Pablo, M., & García-Jiménez, E. (2018). Are CLIL students more motivated? An analysis of affective factor and their relation to language attainment. *Porta Linguarum, 29,* 71–90.
Navés, T., & Victori, M. (2010). CLIL in Catalonia: An overview of research studies. In D. Lasagabaster & Y. Ruiz de Zarobe (Eds.), *Spain: Implementation, results and teacher training* (pp. 31–54). Newcastle: Cambridge Scholars.
Nikula, T. (2007). Speaking English in Finnish content-based classroom. *World Englishes, 26*(2), 206–223.
Nikula, T., Daufouz, E., Moore, P., & Smit, U. (2016). *Conceptualising integration in CLIL and Multilingual Education*. Bristol: Multilingual Matters.
Ojeda, A. (2009). Themes and vocabulary in CLIL and non-CLIL instruction. In Y. Ruiz de Zarobe & R. Jiménez-Catalán (Eds.), *Content and Language Integrated Learning: Evidence from research in Europe* (pp. 130–156). Bristol: Multilingual Matters.
Pavón, V. (2018). Learning outcomes in CLIL programmes: A comparison of results between urban and rural enviroments. *Porta Linguarum, 29,* 9–28.
Pelechano, V. (1994). Prueba MA. *Análisis y Modificación de la Conducta, 20*, 71–72.
Pérez-Cañado, M. (2012). CLIL research in Europe: Past, present and future. *International Journal of Bilingual Education and Bilingualism, 15*(3), 315–341. https://doi.org/10.1080/13670050.2011.630064.
Pérez-Cañado, M. (2014). Teacher training needs for bilingual education: In-service teacher perceptions. *International Journal of Bilingual Education and Bilingualism, 19*(3), 266–295. https://doi.org/10.1080/13670050.2014.980778.
Pérez-Vidal, C., & Roquet, H. (2015). The linguistic impact of a CLIL Science programme: An analysis measuring relative gains. *System, 54,* 80–90.
Pfenninger, S. (2016). All good things come in threes: Early English learning, CLIL and motivation in Switzerland. *Cahiers de l'ILSL, 48,* 119–147.
Rascón, D., & Bretones, C. M. (2018). Socioeconomic status and its impact on language and content attainment in CLIL contexts. *Porta Linguarum, 29,* 115–135.
Roquet, H., & Pérez-Vidal, C. (2015). Do productive skills improve in Content and Language Integrated Learning contexts? The case of writing. *Applied Linguistics, 38*(4), 489–511. https://doi.org/10.1093/applin/amv050.
Ruiz de Zarobe, Y. (2008). CLIL and foreign language learning: A longitudinal study in the Basque Country. *International CLIL Research Journal, 1*(1), 61–72.
Ruiz de Zarobe, Y. (2017). Language awareness and CLIL. In J. Cenoz, D. Gorter, & S. May (Eds.), *Language awareness and multilingualism* (pp. 1–11). Springer, Cham. https://doi.org/10.1007/978-3-319-02325-0_13-2.
Ruiz de Zarobe, Y., & Jiménez-Catalán, R. (2009). *Content and Language Integrated Learning: Evidence from research in Europe*. Bristol: Multilingual Matters.
Rumlich, D. (2016). *Evaluating Bilingual Education in Germany: CLIL students' general proficiency, EFL self-concept and interest*. Berlin: Peter Lang.
San Isidro, X. (2010). An insight into Galician CLIL: Provision and results. In D. Lasagabaster & Y. Ruiz de Zarobe (Eds.), *Spain: Implementation, results and teacher training* (pp. 55–78). Newcastle: Cambridge Scholars.
Santamaría, P. (2018). *EFAI: Evaluación Factorial de las Aptitudes Intelectuales*. Tea.
Turner, M. (2013). CLIL in Australia: The importance of context. *International Journal of Bilingual Education and Bilingualism, 16*(4), 395–410. https://doi.org/10.1080/13670050.2012.691086.

Villarreal, I., & García, M. del P. (2009). Tense and agreement morphology in the interlanguage of Basque/Spanish bilinguals: CLIL versus non-CLIL. In Y. Ruiz de Zarobe & R. Jiménez-Catalán (Eds.), *Content and Language Integrated Learning: Evidence from research in Europe* (pp. 157–175). Bristol: Multilingual Matters.

The Effects of CLIL on FL Learning: A Longitudinal Study

María del Mar Gálvez Gómez

Abstract This article reports on a longitudinal study conducted with Content and Language Integrated Learning (CLIL) and English as a Foreign Language (EFL) students in seven public and private schools situated in the Spanish monolingual community of Andalusia, more concretely in the province of Jaén. This research aims at evaluating the impact of CLIL on the development of 223 students' English linguistic competence by assessing their attainment in grammar, vocabulary, reading, listening and speaking. To this end, the bilingual and the non-bilingual strands were matched on a pre-test in terms of English level, verbal intelligence and motivation and they were then administered post- and delayed post-tests to trace their progress from Primary Education to Compulsory Secondary Education to Baccalaureate. Within-cohort and across-cohort comparisons are carried out in order to determine the evolution of both CLIL and non-CLIL students in the different educational levels in terms of the linguistic components and skills. Discriminant analyses are finally carried out taking into account all the intervening variables considered in the study, with the objective of determining if CLIL is actually responsible for the differences detected or if, on the contrary, other variables account for a greater proportion of the variance.

1 Introduction

After almost two decades since the appearance of Content and Language Integrated Learning in Europe, this 'major trend' (Fernández Sanjurjo et al., 2019, p. 661) has been implemented in all educational settings throughout the continent (Hüttner & Smit, 2014, p. 160), with the objective of pushing plurilingualism forward and meeting the so-called 'mother tongue + 2 objective' (Pérez Cañado, 2018, p. 52), by means of which all citizens should become proficient in their mother tongue and in two foreign languages at least (European Comission, 1995). Hence, with the support of many European Union Institutions, CLIL has been embraced enthusiastically by

M. d. M. Gálvez Gómez (✉)
University of Jaén, Jaén, Spain
e-mail: mmgg0007@red.ujaen.es

M. L. Pérez Cañado (eds.), *Content and Language Integrated Learning in Monolingual Settings*, Multilingual Education 38,
https://doi.org/10.1007/978-3-030-68329-0_8

all the stakeholders (Lasagabaster & Doiz, 2016, p. 1) 'as a lever for change and success in language learning' (Pérez Cañado & Ráez Padilla, 2015, p. 1), becoming 'a well-established part of education systems across Europe' (Surmont et al., 2016, p. 320).

As a consequence of this widespread implementation of CLIL, the research into the effects of this methodology on students and their skills has also increased. As Jäppinen (2005, p. 149) underscores, CLIL has become 'an extremely prolific phenomenon', making foreign language learning more naturalistic (Nieto Moreno de Diezmas, 2016, p. 81) and demonstrating that nowadays 'multilingualism is the norm while monolingualism is the exception' (Ouazizi, 2016, p. 113). However, despite the great amount of publications on the effects of CLIL, some of the already existing studies present a series of methodological shortcomings which might endanger the validity of the results (Bruton, 2011, 2013, 2015; Paran, 2013; Pérez Cañado, 2011, 2012). These lacunae, classified by Pérez Cañado and Ráez Padilla (2015) in terms of *variables, research design and statistical methodology*, make us recognise that 'we simply do not have enough evidence' (Paran, 2013, p. 331) and that there is still 'a need for unbiased, unskewed and methodologically sound research to continue driving the CLIL agenda forward' (Pérez Cañado & Lancaster, 2017, p. 2). In this respect, according to many scholars (Bruton, 2011; Lasagabaster & Ruiz de Zarobe, 2010; Pérez Cañado, 2017a, 2018; Pérez Cañado & Ráez Padilla, 2015; Ruiz de Zarobe, 2011), more importance should be given to longitudinal investigations which can examine the effects of CLIL across the different educational levels.

This is exactly the starting point of the present investigation, which will report on the results of a longitudinal study on the effects of CLIL on foreign language (FL) outcomes across educational levels (Primary, Compulsory Secondary and non-compulsory Secondary Education). As Vez (2009, p. 18) claims, 'There is not yet empirical evidence from EU countries on which to base definitive claims about the educational (or other) advantages of multilingual education'. Moreover, 'longitudinal studies with pre-, post-, and follow-up assessments are still rare' (Piesche et al., 2016, p. 109). Therefore, this investigation seeks to offer a rigorous monitoring of CLIL implementation, which is 'key for a better understanding of the processes and outcomes of these courses' (Coyle et al., 2010 cited in Pascual Bajo, 2018, p. 222).

After framing the topic against the backdrop of prior investigations, the research design of the study will be described, reporting on the results obtained within and across cohorts in terms of English as a foreign language competence. The evolution of the bilingual and non-bilingual groups, which were previously matched on a pre-test phase in terms of English level, verbal intelligence and motivation, from Primary Education to Compulsory Secondary Education (CSE) to Baccalaureate, is depicted through the administration of post- and delayed post-tests. Within-cohort comparisons in relation to the intervening variables considered are also presented, together with the discriminant analyses carried out to find out if the independent variable (CLIL) is actually responsible for the differences detected.

2 A Critical Reading of Prior Research

CLIL practice has spread rapidly in the past ten years, currently spanning the continent from north to south, and from east to west (Pérez Cañado, 2012, p. 319). However, this interest in CLIL has not always been accompanied by the same amount of publications on the issue, especially when regarding research projects with a longitudinal focus. The existing longitudinal investigations tend to focus on the four fields around which Wolff (2005) considers CLIL investigations should be articulated: the effects of CLIL on FL, L1 and subject content competence and motivational aspects. Some of these investigations have focused on Primary or Secondary Education and those evaluating FL learning have considered both receptive and productive skills, although they generally have not done it simultaneously.

Among these longitudinal studies analysing FL competence, the investigation by Admiraal et al. (2006) in The Netherlands is worthy of mention. After measuring the vocabulary knowledge, the reading comprehension level and the oral proficiency of 1,305 Secondary students who had received four years of CLIL education through English in five Dutch schools, the results revealed higher scores for the oral and reading parts of the investigation, whereas no differences appeared when dealing with vocabulary. No negative effects were found either for subject matter learning or the L1. Nevertheless, the study lacks statistical analyses that confirm the outcomes can be attributed to CLIL (Admiraal et al., 2006, p. 91).

One year later, this time in Switzerland, Serra (2007) conducted a longitudinal study to evaluate German-speaking Primary Education students' FL competence, focusing on their oral and written comprehension in Italian or Romansch as a second language. Similar results were obtained for the experimental and the control groups, suggesting then that both cohorts performed equally well on these skills.

Turning now to the Spanish scenario, special attention should be paid to the investigation by Ruiz de Zarobe (2008), who compared the oral and written competence of 89 students in their 3rd and 4th year of CSE and at the post-compulsory level. The results were extremely positive since CLIL students obtained better outcomes than the non-CLIL counterparts. Further evidence on the differences between skills and abilities in favour of CLIL was provided, coinciding with previous research (Jiménez Catalán et al., 2006; Ruiz de Zarobe, 2007; Ruiz de Zarobe & Jiménez Catalán, 2009 cited in Lasagabaster & Ruiz de Zarobe, 2010, p. 17).

In 2015, Rallo Fabra and Jacob carried out another longitudinal research in Secondary Education to evaluate the effects of CLIL on pupils' oral skills at the onset of the programme and two years after its implementation. They worked with 43 students from state-run Secondary schools in the Balearic Islands, who were divided into the experimental CLIL group and the control non-CLIL cohort. Special attention was paid to students' fluency and the number of vowel errors in English. Outcomes revealed pupils' pronunciation of English vowels was unaffected by CLIL and there was no significant improvement over the two years considered. Moreover, CLIL students' pronunciation was not better than their EFL peers' and no significant differences were detected in fluency either (Rallo Fabra & Jacob, 2015).

Similarly, in the study by Pladevall-Ballester and Vallbona (2016), the superiority of CLIL is not proved. The project, which was developed over the course of two academic years, focused on analysing students' receptive skills in a minimal CLIL input context at Primary school level. 287 pupils from four different state-funded private schools participated in the investigation, whose results showed an important progress regarding the achievement and development of reading and listening skills in both CLIL and non-CLIL cohorts between the first and the last test. However, the EFL group outperformed the CLIL cohort in relation to listening skills. As for reading, no significant differences were detected.

More recently, Pascual Bajo (2018) developed an investigation aimed at evaluating CLIL from a qualitative and a quantitative point of view within the context of two educational institutions: a public school with CLIL and non-CLIL streams, and a semi-private school with no CLIL provision in the province of Valencia. 63 pupils constituted the sample of the longitudinal quantitative study. Outcomes confirmed the superior English language competence of CLIL pupils. The CLIL cohort outperformed the non-CLIL stream on all the skills and aspects considered, with particularly significant differences in the use of English, vocabulary and reading (Pascual Bajo, 2018, p. 382). Six months after the end of the CLIL programme, in the delayed post-test phase, the outcomes improved for all skills studied, except for that of writing.

To finish, moving on to the Andalusian community, where our study has also been carried out, two longitudinal investigations must be foregrounded: Pérez Cañado and Lancaster's (2017) and Pérez Cañado's (2018). The former is a longitudinal, quantitative, quasi-experimental case study on oral comprehension and production in Andalusia. Their research aimed at determining whether CLIL students acquired greater oral comprehension and production skills in comparison to non-CLIL pupils. Moreover, it tried to find out if the possible differential effects of CLIL continued after the CSE CLIL programme finished. 24 secondary students participated in the investigation and sat two FL competence tests to assess their oral comprehension and production skills. An initial pre-test was used to guarantee the homogeneity of the cohorts in terms of English language proficiency. Outcomes evinced CLIL students had a higher level in English oral comprehension and production when compared to those obtained by the EFL group. It was also found that the effects of CLIL pervaded six months after the programme was discontinued, but only in the case of oral production.

The latter (Pérez Cañado, 2018) focused on the effects of CLIL on L2 learning. This investigation is especially interesting in the context of our study because of its similarities with our study in terms of research questions, instruments and variables. A total of 1,033 students enrolled in a CLIL programme and 991 EFL pupils took part in the project. These learners, who were completing 6th grade of Primary Education or finishing the 4th grade of CSE, came from different public, semi-public and private schools situated in three Spanish monolingual communities: Andalusia, The Canary Islands and Extremadura. Participants were firstly matched in terms of FL proficiency, motivation and verbal intelligence to guarantee the homogeneity of the cohorts. The outcomes revealed the CLIL group at both Primary and Secondary levels had a higher linguistic competence (on grammar, vocabulary, listening, speaking and

reading). The linguistic competence differential was especially marked at Secondary Education, where the CLIL group clearly outperformed its non-CLIL counterpart on all aspects considered. As for the durability of the effects of CLIL, it was shown that outcomes pervaded six months after the programmes were discontinued. However, no statistically significant differences were detected between the EFL semi-private stream and the CLIL groups (public and private) in 1st year of non-compulsory Secondary Education. The delayed post-test phase results coincide, to some extent, with Pladevall-Ballester and Vallbona's findings (2016), according to which CLIL had more positive effects on the productive skills than on the receptive skills. To finish, the discriminant analysis carried out to study the competence differential between the treatment and comparison groups allowed Pérez Cañado (2018) to confirm CLIL is truly responsible for the differences found.

After this brief summary of the most important longitudinal studies developed hitherto in Europe and more concretely in Spain (cf. also Chapter 7 in this volume), let us now turn to our investigation, which strives to provide updated empirical evidence on CLIL practice by overcoming the main shortcomings presented by prior investigations.

3 The Study

The present investigation is framed within a broader research project which has developed a large-scale evaluation of CLIL programmes in one of the Spanish monolingual communities with the least tradition in bilingual education: Andalusia. With a mixed-research design, the study examines the effects of CLIL from quantitative and qualitative perspectives. The impact of CLIL on FL learning, content learning (Natural Sciences) and L1 learning of Primary (6th grade) and Secondary (4th grade) Education Students is analysed in the quantitative part of the project, in which participants in the experimental CLIL group are assessed in comparison to a control non-CLIL cohort, in order to find out whether they develop superior language and content skills to those promoted by a traditional EFL programme. Moreover, the study aims to determine if the possible effects exerted by CLIL pervade six months after the programme is discontinued, when the same CSE students are in the first grade of Baccalaureate, or if they gradually disappear. This quantitative part is then completed from a qualitative point of view by means of a SWOT analysis on the satisfaction generated by CLIL programmes among all the agents involved. Stakeholders' perspectives (teachers, students and parents) are collected via questionnaires, and personal and focus-group interviews are carried out with teachers and students. The present study is inserted within the quantitative side of the project and focuses specifically on the effects of CLIL on English as an FL learning through the following research questions.

3.1 Research Questions

RQ1. Do CLIL programmes implemented with Primary and Secondary school students (experimental group) develop superior linguistic competence (in grammar, vocabulary, receptive skills and oral production) to that promoted by EFL programmes with students from the same level (control group)? More simply, is there a linguistic competence differential between CLIL and EFL groups at Primary and Secondary school level in the province of Jaén?

RQ2. What is the modulating (differential) effect exerted on the Primary and Secondary students' English language competence by the following intervening variables: type of school (public and private), setting (rural–urban), gender, socioeconomic status (SES), motivation, verbal intelligence and English level?

RQ3. Do the possible differential effects exerted by CLIL programmes on English language competence pervade six months after the CLIL programme is discontinued or do they gradually disappear?

RQ4. If there is a competence differential between the treatment and comparison groups, is it truly ascribable to language learning based on academic content processing?

3.2 Research Design

The quantitative side of the broader study is an example of applied, primary, quasi-experimental research, with a pre-test/post-test control group design, to which a delayed post-test has also been added. Thus, as Rossell and Baker (1996), together with Cummins (1979 cited in Lancaster, 2015, p. 137) specified, four benchmarks are necessary for studies to be methodologically sound, and this study meets them:

1. Studies must compare students in a bilingual programme to a control group of similar students.
2. The design must ensure that initial differences between treatment and control groups are controlled statistically.
3. Results must be based on standardised test scores.
4. Differences between the scores of treatment and control groups must be determined by means of appropriate statistical tests.

3.3 Sample

The final sample that took part in the quantitative part of the investigation was made up of 223 students from public and private centres in the Andalusian province of Jaén, situated in the north-eastern part of the autonomous community, in the south of Spain. Most of the participants are studying the fourth grade of Compulsory Secondary

Education (60.1%), being the rest Primary Education students (39.9%). In the same vein, the majority of the cohort belong to a public school (76.2%) where CLIL and EFL branches co-exist, while 23.8% are enrolled in a bilingual private school (23.8%). 76.2% of these schools are located in urban settings, while the remaining 23.8% are placed in rural areas. Regarding gender, practically equal percentages are found, with a slightly higher number of female participants (50.7%). Finally, a significantly higher percentage of CLIL pupils have participated in the study (81.3%), with the rest of pupils being enrolled in the traditional EFL programme (18.7%).

The homogeneity of the sample, both in the experimental and the control group, has been guaranteed from the beginning of the research. In fact, the first year was entirely devoted to matching students within schools in terms of verbal intelligence, motivation and English level. Pupils were administered initial motivation and verbal intelligence tests that would allow us to select the really comparable groups. Moreover, the English grades of these students were collected to compare the results obtained by CLIL and non-CLIL pupils. Those schools whose outcomes evinced the greatest homogeneity comprise the final sample of the investigation.

3.4 Variables

Three different types of variables have been incorporated in the study: dependent, independent and moderating.

- The *dependent variable* is the students' English language (FL) competence (grammar, vocabulary, receptive skills and oral productive skill).
- The *independent variable* corresponds to the CLIL programme implemented in the different types of schools (public and private).
- Finally, the *moderating variables* are the following:
  - Verbal intelligence
  - Motivation
  - Socioeconomic status (SES)
  - Gender
  - Type of School
  - Setting

3.5 Instruments

For the collection of data, four different instruments were used, depending on the stage of the investigation. An initial questionnaire was firstly administered to students. It comprised personal information and data on their parents' age and educational level, which was taken as a proxy for SES. Moreover, verbal intelligence and motivation tests were employed in this initial phase, together with the English

language tests. All of these were already existing and validated instruments which belong to language teaching and psychology research areas.

The verbal intelligence test, which was part of the Evaluación Factorial de las Aptitudes Intelectuales (EFAI), designed by Santamaría et al. (2016), and the motivation test (Pelechano Barberá, 1994) were applied in each of the schools over the course of an hour almost at the end of the academic year 2014–2015, after exactly ten years of CLIL implementation in the community. Two different versions of the verbal intelligence tests were applied, adapted to the sixth grade of Primary Education and fourth grade of Compulsory Secondary Education. The former version comprises 26 items, while the latter is reduced to 23. In both cases, pupils had to choose from four multiple-choice options involving analogies, antonyms and odd-one-out and they had five minutes to complete as many items as they could. In turn, Pelechano's MA test (1994) comprises a total of 35 items aimed at measuring students' motivation, and it isolates four motivational factors of achievement and anxiety: (i) vain desire to work and self-esteem (10 items); (ii) anxiety when facing exams (9 items); (iii) lack of interest in studying (9 items); and (iv) realistic personal self-demand (7 items).

Finally, the English competence tests applied were originally devised for the project (Madrid et al., 2018) and incorporated three different batteries of six tests each (grammar, vocabulary, reading, writing,[1] listening and speaking) which corresponded to the levels at which our study has been developed, namely, 6th grade of PE, 4th grade of CSE and 1st grade of Baccalaureate. A rubric was also designed and validated for the assessment of oral production, following five main criteria: grammatical accuracy, lexical range, fluency and interaction, pronunciation and task fulfilment (Pérez Cañado & Lancaster, 2017).

3.6 Data Analysis

The data obtained from all the tests has been analysed statistically with the aid of the SPSS program, in its 23.0 version. To guarantee the homogeneity and comparability of the sample, participants have been matched for verbal intelligence, motivation and English level through the ANOVA and the T-test. Moreover, in order to determine the existence of any statistically significant differences within and across groups in terms of the different identification variables considered, the ANOVA, the T-test, the Mann–Whitney U test and Tukey's HSD test have been employed. To calculate effect sizes, Cohen's *d* and eta squared have been used. Lastly, to address RQ4, successive discriminant analyses have been carried out to establish which variable(s) are truly responsible for the differences detected.

[1]The results corresponding to the writing skill are not included in this article since they are still under analysis.

4 Results and Discussion

Taking into consideration the research questions set out at the beginning of the investigation, a detailed examination of the FL level attained by both CLIL and non-CLIL students will be offered in this section, with a special focus on productive and receptive skills, the effects of the different intervening variables considered on the students' FL proficiency, and lastly, the durability or medium-term effects of the CLIL and EFL programmes on FL competence.

4.1 Across-Cohort Comparison

After the initial overall comparison was carried out, statistically significant differences were detected in favour of the experimental group ($p = < 0.001$). High confidence levels were then found on most of the aspects analysed, with listening being the only skill for which no differences between the CLIL and the non-CLIL group were observed, coinciding with Serra (2007). However, CLIL students clearly outperformed their non-CLIL counterparts since the means obtained were 7.99 and 6.56, respectively (cf. Table 1).

Focusing on speaking, statistically significant differences in favour of the CLIL group were detected ($p = .0240$). Moreover, more differences emerged when attending to the five subaspects mentioned above. Thus, CLIL students outperformed their non-CLIL counterparts in their knowledge and use of grammar ($p = .0330$) and in their pronunciation in the FL ($p = .0050$). Nevertheless, no statistically significant differences appeared when analysing students' lexical range, their fluency in English and their adaptability to the task provided ($p = .0840$, .0580 and .0590, respectively). However, the means obtained by the CLIL group in each one of these subaspects were

Table 1 Foreign language competence: Across-cohort comparison

Educational Level	Skills	Group	Mean	Standard Deviation	Cohen's d	*p*-value
General	Use of English	Non-CLIL	13.34	8.323	−0.995	<0.001
		CLIL	23.23	10.263		
	Vocabulary	Non-CLIL	7.98	4.150	−0.807	<0.001
		CLIL	10.75	3.250		
	Listening	Non-CLIL	6.56	4.707	−0.301	0.084
		CLIL	7.99	4.761		
	Reading	Non-CLIL	3.15	2.963	−0.730	<0.001
		CLIL	5.83	3.810		
	Total	Non-CLIL	31.02	15.244	−1.231	<0.001
		CLIL	47.79	13.223		

Table 2 Foreign language competence. General speaking analysis

	Group	Mean	Standard Deviation	Rosenthal's R	*p*-value
Speaking_Total	Non-CLIL	5.806	2.3772	.7577	.0240
	CLIL	7.296	1.7279		
Grammatical	Non-CLIL	1.028	.5278	.6799	.0330
	CLIL	1.370	.4722		
Lexical_Range	Non-CLIL	1.194	.5461	.4465	.0840
	CLIL	1.426	.3008		
Fluency_Interaction	Non-CLIL	1.111	.6077	.5377	.0580
	CLIL	1.444	.4237		
Pronunciation	Non-CLIL	1.222	.3919	1.1575	.0050
	CLIL	1.556	.3203		
Task_Fulfilment	Non-CLIL	1.250	.4926	.5295	.0590
	CLIL	1.500	.3669		

always higher than the non-CLIL group. These results corroborate Pérez Cañado and Lancaster's outcomes (2017), according to which CLIL students outperformed their EFL counterparts in oral production (cf. Table 2).

When analysing the data from Primary and Secondary students separately, the superior linguistic competence of CLIL pupils was again confirmed. When comparing 4th year CSE EFL students with CLIL learners at the same educational level, it is proved that the latter outperformed the former on all the skills and aspects sampled. On the contrary, when dealing with 6th year Primary Education students, no differences appeared in terms of speaking, in contrast to the tendency found in previous research at Primary Education level (Madrid & Barrios, 2018; Nieto Moreno de Diezmas, 2016; Ruiz de Zarobe, 2008), according to which the CLIL group was significantly superior in terms of oral production. However, Primary school students belonging to CLIL are found to be superior on all skills analysed ($p = 0.931$) (cf. Table 3).

4.2 Differential Effect of Intervening Variables on FL Competence

To fully understand the effects of CLIL and EFL programmes on English proficiency, an analysis of the data in terms of the different intervening variables considered will now be included, addressing our RQ2.

Taking into account the gender of the participants, in the initial overall analysis, no statistically significant differences were detected between female and male participants in any of the skills evaluated, coinciding with the results obtained by Heras and Lasagabaster (2015) and Pascual Bajo (2018). However, in a deeper analysis carried

Table 3 FL competence comparison per group (T-test)

Educational level	Skill	Group	Mean	Standard Deviation	Cohen's d	*p*-value
Primary Education	Use of English	Non-CLIL	10.94	7.197	−0.777	0.006
		CLIL	15.99	6.339		
	Vocabulary	Non-CLIL	11.13	2.964	−0.575	0.040
		CLIL	12.75	2.803		
	Listening	Non-CLIL	11.88	2.247	−0.665	0.018
		CLIL	13.29	2.098		
	Reading	Non-CLIL	5.25	3.357	−1.072	<0.001
		CLIL	9.14	3.679		
	Speaking	Non-CLIL	6.083	2.354	−0.077*	0.931
		CLIL	5.800	2.049		
	Total	Non-CLIL	39.19	14.520	−0.901	0.002
		CLIL	51.16	13.015		
Compulsory Secondary Education	Use of English	Non-CLIL	14.88	8.76	−1.438	<0.001
		CLIL	28.27	9.43		
	Vocabulary	Non-CLIL	5.96	3.52	−1.156	<0.001
		CLIL	9.35	2.78		
	Listening	Non-CLIL	3.16	1.70	−0.761	0.004
		CLIL	4.30	1.46		
	Speaking	CLIL	5.667	2.480	0.434*	0.011
		Non-CLIL	7.636	1.497		
	Reading	Non-CLIL	1.80	1.66	−1.118	<0.001
		CLIL	3.52	1.51		
	Total	Non-CLIL	25.80	13.52	−1.508	<0.001
		CLIL	45.45	12.91		

*The analyses for speaking are done by means of Rosenthal's R

out within each separate cohort, the T-test evinced women in the CLIL group have a higher level of vocabulary in the FL and generally outperform their male peers. Regarding the non-CLIL group, no statistical confirmation of clear differences could be reported between girls and boys (cf. Table 4).

Following with the variable of setting, our first evaluation detected statistically significant differences in favour of those students belonging to an urban centre, both generally and in their level of use of English, although pupils in the rural schools outperformed their counterparts on the oral receptive skill. Similar results appeared when analysing each cohort separately, as differences were also found in favour of the urban school pupils within the CLIL group. However, the rural school students outperformed their counterparts on all skills tested except for use of English within the non-CLIL group (cf. Table 5).

Table 4 Comparison per gender (T-test)

Group	Skills	Gender	Mean	Standard Deviation	Cohen's d	*p*-value
General	Use of English	Male	20.68	10.75	−0.111	0.408
		Female	21.86	10.45		
	Vocabulary	Male	9.68	3.89	−0.257	0.056
		Female	10.61	3.32		
	Listening	Male	7.63	4.70	−0.002	0.988
		Female	7.64	4.87		
	Reading	Male	4.88	3.68	−0.197	0.144
		Female	5.63	3.90		
	Total	Male	42.87	15.26	−0.189	0.160
		Female	45.73	15.02		
Non-CLIL	Use of English	Male	15.25	9.94	0.454	0.161
		Female	11.52	6.13		
	Vocabulary	Male	8.10	4.71	0.058	0.854
		Female	7.86	3.65		
	Listening	Male	7.10	4.54	0.222	0.481
		Female	6.05	4.91		
	Reading	Male	3.40	3.45	0.166	0.599
		Female	2.90	2.47		
	Speaking	Male	6.250	2.699	.1683*	0.515
		Female	5.450	2.166		
	Total	Male	33.85	17.59	0.363	0.252
		Female	28.33	12.46		
CLIL	Use of English	Male	22.15	10.64	−0.207	0.170
		Female	24.26	9.83		
	Vocabulary	Male	10.21	3.49	−0.329	0.030
		Female	11.26	2.92		
	Listening	Male	7.92	4.74	−0.028	0.850
		Female	8.05	4.81		
	Reading	Male	5.34	3.65	−0.248	0.100
		Female	6.29	3.92		
	Speaking	Male	7.036	1.759	.1742*	0.375
		Female	7.577	1.718		
	Total	Male	45.62	13.65	−0.325	0.032
		Female	49.87	12.53		

*The analyses for speaking are done by means of Rosenthal's R

Table 5 Cohort comparison regarding setting

Group	Skills	Setting	Mean	Standard Deviation	Cohen's d	*p*-value
General	Use of English	Rural	15.32	8.57	−0.776	<0.001
		Urban	23.14	10.49		
	Vocabulary	Rural	10.13	3.73	−0.007	0.963
		Urban	10.16	3.62		
	Listening	Rural	9.23	4.06	0.444	0.002
		Urban	7.14	4.89		
	Reading	Rural	5.26	3.61	0.001	0.99
		Urban	5.26	3.88		
	Total	Rural	39.94	14.18	−0.383	0.016
		Urban	45.69	15.25		
Non-CLIL	Use of English	Rural	12.50	8.71	−0.179	0.573
		Urban	14.00	8.14		
	Vocabulary	Rural	10.67	3.16	1.402	<0.001
		Urban	5.87	3.61		
	Listening	Rural	11.11	3.08	3.401	<0.001
		Urban	3.00	1.65		
	Reading	Rural	5.17	3.17	1.516	<0.001
		Urban	1.57	1.50		
	Total	Rural	39.44	14.32	1.119	0.001
		Urban	24.43	12.68		
CLIL	Use of English	Rural	16.77	8.25	−0.822	<0.001
		Urban	24.81	10.11		
	Vocabulary	Rural	9.86	4.01	−0.343	0.133
		Urban	10.97	3.01		
	Listening	Rural	8.26	4.19	0.070	0.711
		Urban	7.92	4.90		
	Reading	Rural	5.31	3.86	−0.167	0.377
		Urban	5.95	3.80		
	Total	Rural	40.20	14.32	−0.743	<0.001
		Urban	49.65	12.30		

Socioeconomic status (SES) was also incorporated in our investigation taking into account the educational level of parents. Hence, three different groups were established according to their educational attainment: high (Tertiary Education), medium (vocational training or Secondary) and low (school qualifications or no studies). As a result, after applying the ANOVA test, statistically significant differences were detected on all skills analysed in favour of those students having a high socioeconomic status, except for use of English (cf. Table 6). These results coincide with

Table 6 Comparison among cohorts regarding SES

Group	Skills	SES	Mean	Standard Deviation	Eta Squared	*p*-value
General	Use of English	Low	21.13	10.25	0.001	0.872
		Medium	21.81	11.13		
		High	21.00	10.18		
	Vocabulary	Low	8.28	3.26	0.150	0.000
		Medium	9.54	3.48		
		High	11.71	3.22		
	Listening	Low	5.19	3.18	0.236	0.000
		Medium	5.97	3.77		
		High	10.32	4.90		
	Reading	Low	3.23	1.97	0.197	0.000
		Medium	4.23	2.99		
		High	7.18	4.24		
	Total	Low	37.83	13.04	0.120	0.000
		Medium	41.56	14.51		
		High	50.20	14.19		

Anghel et al. (2016) and Fernández Sanjurjo et al. (2019) outcomes, according to which CLIL programmes tend to have a negative effect on students from a low socioeconomic background. However, recent investigations have proved that CLIL has cancelled out differences among social classes (Lorenzo, 2019; Pavón Vázquez, 2018; Pérez Cañado, 2017b; Rascón Moreno & Bretones Callejas, 2018). No statistically significant differences appeared among non-CLIL learners in any of the skills evaluated, except for their knowledge of vocabulary.

When analysing the cohorts separately, the same tendency was repeated among CLIL students, where those pupils with a high socioeconomic level are revealed to also have the highest level in all skills tested but use of English. In the case of non-CLIL students, statistically significant differences in favour of those having a high socioeconomic level exist only when dealing with their knowledge of vocabulary (cf. Table 7).

Valuable results were also obtained when taking into consideration type of school. In general terms, statistically significant differences were found for all skills considered in favour of those students belonging to a private bilingual school, with the exception of use of English, for which pupils from public bilingual schools were revealed to obtain better outcomes (cf. Table 8).

Different results are obtained within cohorts. In the case of Primary Education students, the same pattern is repeated, showing that private bilingual school students obtain higher results in vocabulary, use of English, listening and reading. However, when dealing with Secondary Compulsory Education pupils, results vary, as no private bilingual schools were considered in our study. Thus, statistically significant differences were found on all skills analysed in favour of public bilingual school

Table 7 Comparison within cohorts in terms of SES (ANOVA)

Group	Skills	SES	Mean	Standard Deviation	Eta Squared	*p*-value
Non-CLIL	Use of English	Low	14.82	7.26	0.021	0.684
		Medium	12.70	7.15		
		High	12.00	9.43		
	Vocabulary	Low	6.00	3.32	0.177	0.030
		Medium	8.25	4.15		
		High	10.88	3.27		
	Listening	Low	5.45	4.32	0.098	0.157
		Medium	6.35	4.66		
		High	9.50	4.78		
	Reading	Low	2.18	1.54	0.046	0.427
		Medium	3.50	3.59		
		High	3.75	2.76		
	Speaking	Low	3.333	1.6073	0.235	0.102
		Medium	6.182	2.3798		
		High	6.625	1.9738		
	Total	Low	28.45	8.69	0.036	0.518
		Medium	30.80	15.99		
		High	36.13	16.52		
CLIL	Use of English	Low	23.06	10.33	0.021	0.161
		Medium	25.28	10.48		
		High	21.93	9.98		
	Vocabulary	Low	8.97	2.95	0.143	0.000
		Medium	10.05	3.17		
		High	11.96	2.96		
	Listening	Low	5.11	2.82	0.275	0.000
		Medium	5.95	3.47		
		High	10.58	4.82		
	Reading	Low	3.56	1.99	0.214	0.000
		Medium	4.58	2.72		
		High	7.62	4.18		
	Speaking	Low	6.692	1.5075	0.125	0.141
		Medium	7.688	1.6677		
		High	8.083	2.0595		
	Total	Low	40.69	12.87	0.116	0.000
		Medium	45.86	11.65		
		High	52.09	12.89		

Table 8 Comparison across cohorts in terms of type of school (ANOVA)

Group	Skills	Type of school	Mean	Standard Deviation	Cohen's d	*p*-value
General	Use of English	Public school non-CLIL	13.34	8.32	0.252	0.000
		Public school CLIL	25.88	10.62		
		Private school CLIL	16.98	5.72		
	Vocabulary	Public school non-CLIL	7.98	4.15	0.271	0.000
		Public school CLIL	9.65	3.06		
		Private school CLIL	13.34	1.98		
	Listening	Public school non-CLIL	6.56	4.71	0.556	0.000
		Public school CLIL	5.46	3.08		
		Private school CLIL	13.96	1.66		
	Reading	Public school non-CLIL	3.15	2.96	0.455	0.000
		Public school CLIL	4.14	2.50		
		Private school CLIL	9.81	3.39		
	Total	Public school non-CLIL	31.02	15.24	0.249	0.000
		Public school CLIL	45.12	13.41		
		Private school CLIL	54.09	10.41		

students, in line with previous investigations which attest to the superiority of CLIL over EFL programmes (Lasagabaster & Ruiz de Zarobe, 2010; Lasagabaster, 2009; Pascual Bajo, 2018) (cf. Table 9).

Table 9 Comparison in terms of type of school (ANOVA)

Group	Skills	Type of school	Mean	Standard Deviation	Cohen's d	*p*-value
Primary Education	Use of English	Public school non-CLIL	10.94	7.20	0.131	0.002
		Public school CLIL	13.35	7.26		
		Private school CLIL	16.98	5.72		
	Vocabulary	Public school non-CLIL	11.13	2.96	0.138	0.002
		Public school CLIL	11.20	3.94		
		Private school CLIL	13.34	1.98		
	Listening	Public school non-CLIL	11.88	2.25	0.273	0.000
		Public school CLIL	11.50	2.12		
		Private school CLIL	13.96	1.66		
	Reading	Public school non-CLIL	5.25	3.36	0.213	0.000
		Public school CLIL	7.35	3.90		
		Private school CLIL	9.81	3.39		
	Total	Public school non-CLIL	39.19	14.52	0.206	0.000
		Public school CLIL	43.40	16.06		
		Private school CLIL	54.09	10.41		
Compulsory Secondary Education	Use of English	Public school non-CLIL	14.88	8.762	0.246	0.000
		Public school CLIL	28.27	9.434		
		Private school CLIL				
	Vocabulary	Public school non-CLIL	5.96	3.518	0.174	0.000
		Public school CLIL	9.35	2.784		

(continued)

Table 9 (continued)

Group	Skills	Type of school	Mean	Standard Deviation	Cohen's d	*p*-value
		Private school CLIL				
	Listening	Public school non-CLIL	3.16	1.700	0.084	0.001
		Public school CLIL	4.30	1.455		
		Private school CLIL				
	Reading	Public school non-CLIL	1.80	1.658	0.165	0.000
		Public school CLIL	3.52	1.513		
		Private school CLIL				
	Total	Public school non-CLIL	25.80	13.515	0.264	0.000
		Public school CLIL	45.45	12.915		
		Private school CLIL				

4.3 Durability of Effects of the CLIL Programme on FL Competence

The across- and within-group comparisons presented above are also complemented with general and group-focused analyses in order to confirm if the effects of CLIL remained once the programme was discontinued or if, on the contrary, they gradually disappeared. Hence, vis-à-vis RQ3, a comparison between the results obtained in the post-test phase and the outcomes of the delayed post-tests sat by the same students six months later was carried out.

Starting with the analysis of the CLIL group, our findings indicate that students obtained slightly better results on most skills tested, although no statistically significant differences were detected except for listening ($p = < 0.001$). Slightly worse results were found for use of English in the delayed post-test phase, although still not significant (cf. Table 10).

Turning now to the analysis of the non-CLIL students, no statistically significant differences were detected between the two phases except for the reading skill ($p = 0.007$). Moreover, slightly worse results were detected in the delayed post-tests for use of English and listening, something which clearly differs from the results obtained in the experimental group. Non-CLIL pupils were also found to have a

Table 10 Post- to delayed post-test comparison of CLIL cohort's skills

Group	Skills	Post/Delayed	Mean	Standard Deviation	Cohen's d	*p*-value
CLIL	Use of English	Post	33,51	9,456	0.031	0.687
		Delayed	33,22	8,792		
	Vocabulary	Post	10,43	3,373	−0.216	0.097
		Delayed	11,08	2,637		
	Listening	Post	4,71	1,683	−0.393	<0.001
		Delayed	5,35	1,535		
	Reading	Post	4,20	1,472	−0.208	0.213
		Delayed	4,51	1,474		
	Total	Post	52,86	13,519	−0.100	0.106
		Delayed	54,16	12,686		

greater knowledge of vocabulary, although no statistically significant differences were ascertained for this skill (cf. Table 11).

After having expounded on the results obtained in the post-test phase by both CLIL and non-CLIL students, let us now analyse if statistically significant differences emerge across the cohorts. Clear-cut tendencies were discerned, pointing to the supremacy of the bilingual programme, as statistically significant differences emerged on all skills considered in favour of CLIL pupils.

As can be observed in the table below, CLIL students outperformed their non-CLIL peers on use of English ($p = < 0.001$) and on their knowledge of vocabulary ($p = < 0.001$) in the FL. In the same vein, the means obtained for both listening and reading skills were significantly higher for CLIL students (5.35 and 4.51, respectively), coinciding with previous studies by Pladevall-Ballester and Vallbona (2016) and Pascual Bajo (2018). However, these results must be interpreted with caution, since the large effect sizes shown by Cohen's *d* have to be taken into consideration in our assessment (cf. Table 12).

Table 11 Post- to delayed post-test comparison of non-CLIL cohort's skills

Group	Skills	Post/Delayed	Mean	Standard Deviation	Cohen's d	*p*-value
Non-CLIL	Use of English	Post	15.58	5.551	0.031	0.899
		Delayed	15.42	5.248		
	Vocabulary	Post	5.33	2.348	−0.682	0.056
		Delayed	6.83	2.038		
	Listening	Post	3.00	1.348	0.306	0.422
		Delayed	2.58	1.379		
	Reading	Post	1.33	1.303	−1.201	0.007
		Delayed	2.83	1.193		
	Total	Post	25.25	7.124	−0.334	0.281

Table 12 Comparison per group (T-test)

Group	Skills	Group	Mean	Standard Deviation	Cohen's d	*p*-value
General	Use of English	Non-CLIL	15.42	5.25	−2.159	<0.001
		CLIL	33.22	8.79		
	Vocabulary	Non-CLIL	6.83	2.04	−1.675	<0.001
		CLIL	11.08	2.64		
	Listening	Non-CLIL	2.58	1.38	−1.834	<0.001
		CLIL	5.35	1.54		
	Reading	Non-CLIL	2.83	1.19	−1.176	<0.001
		CLIL	4.51	1.47		
	Total	Non-CLIL	27.67	7.36	−2.231	<0.001
		CLIL	54.16	12.69		

4.4 Language Competence Differential: Discriminant Analyses

Finally, the effect of the different intervening variables considered is quantified. Successive discriminant analyses have helped us determine which variables are the most significant in explaining the differences detected between the CLIL and non-CLIL strands.

As for the differences found between the experimental and the control group in English, it can be clearly seen that the independent variable (Group—CLIL) together with the moderating variables of socioeconomic status (SES) and verbal intelligence are the ones which display the greatest significance ($p = 0.000$). Accordingly, these variables were later used in a discriminant function, which proved that all of them were significant, as the *p*-value obtained was 0,000. Thus, we can confirm that the CLIL programme has the greatest weight in explaining the language competence differential between the experimental and control groups, mirroring Pascual Bajo's (2018) and Pérez Cañado's (2018) outcomes (cf. Tables 13 and 14).

5 Conclusion

This study has addressed one of the most important current areas of interest in Second Language Acquisition (SLA) research: the analysis of how CLIL is playing out in a Spanish monolingual region which lacks a firm tradition for foreign language learning. More concretely, the investigation has been developed in the province of Jaén, an area where little research on the topic has been published so far. In order to overcome the main lacunae presented by previous investigations into the topic in terms of homogeneity, variables or statistical analyses, the project has worked

Table 13 Test of equality group means

	Wilks' Lambda	F	df1	df2	Sig.
Exposure to English	.991	2.013	1	214	.157
SES	.854	36.672	1	214	.000
Verbal Intelligence	.837	41.827	1	214	.000
Goodwill	.999	.138	1	214	.711
Anxiety	.992	1.660	1	214	.199
Lack of interest	.954	10.400	1	214	.001
Selfcommitment	.986	3.050	1	214	.082
Type	.864	33.604	1	214	.000
Group	.940	13.670	1	214	.000
Gender	.998	.425	1	214	.515

Table 14 Canonical discriminant functions

Function	Eigenvalue	% of Variance	Cumulative %	Canonical Correlation
1	.362[a]	100.0	100.0	.515
Test of Functions	**Wilks' Lambda**	**Chi-square**	**Df**	**Sig.**
1	.734	65.586	3	.000

with students from three different educational levels (Primary Education, Compulsory Secondary Education and Baccalaureate), divided into two different cohorts according to the educational programme they are following (experimental CLIL group and control non-CLIL group) and taking into account different moderating variables.

Regarding RQ1, our outcomes have allowed us to confirm the superior linguistic competence in English of those learners following the CLIL programme. As detailed in Sect. 1 in Chapter "Are CLIL Settings More Conducive to the Acquisition of Digital Competences? A Comparative Study in Primary Education", the CLIL cohort outperformed the non-CLIL stream on all skills analysed, being listening the only skill for which no statistically significant differences were detected. The oral production of CLIL and non-CLIL students was also analysed, taking into account their use of English (grammar and vocabulary), their pronunciation and fluency and their adequacy to the task. The outcomes obtained attest to the superiority of CLIL students, especially in their knowledge and use of grammar and in their pronunciation in the FL (Madrid & Barrios, 2018; Nieto Moreno de Diezmas, 2016; Pascual Bajo, 2018; Ruiz de Zarobe, 2008). However, no statistically significant differences appeared when analysing learners' lexical range, fluency or their adaptability to the task presented.

A total confirmation of these results was provided when evaluating students separately according to their educational stage. Thus, when comparing 4th year of CSE students following a traditional EFL programme with their CLIL peers, it was proved

that the latter outperformed the former on all the skills and aspects considered. On the contrary, slightly more negative results were obtained regarding the oral production of 6th year Primary Education students, since no differences appeared between cohorts.

In line with RQ2, the data obtained by means of the English tests was also analysed taking into consideration the different intervening variables considered. Valuable conclusions can be drawn in this respect, as numerous differences arose across and within cohorts. From a global perspective, no differences were detected in terms of gender, while the variable of setting offered statistically significant differences in favour of students studying in an urban context. However, students from rural centres obtained better results when analysing their receptive skills. Regarding socioeconomic status, statistically significant differences arose in favour of students with a high socioeconomic status on all skills considered, except for use of English, corroborating Anghel, Cabrales and Carro's (2016) and Fernández Sanjurjo et al. (2019) outcomes, according to which bilingual programmes affected negatively those students whose socioeconomic status was lower. No statistically significant differences appeared among non-CLIL learners in any of the skills evaluated, except for their knowledge of vocabulary. Our outcomes thus run counter to the tendency detected by some of the most recent research (Lorenzo, 2019; Pavón Vázquez, 2018; Pérez Cañado, 2017b; Rascón Moreno & Bretones Callejas, 2018), where CLIL has cancelled out differences among social classes.

In the overall comparisons, the type of school variable yielded differences on all skills considered in favour of those students belonging to a private bilingual school, except for use of English, in which learners from public bilingual schools obtained better outcomes. The same tendency was repeated when analysing the results from Primary schools, since students from private centres obtained better results in vocabulary, use of English, listening and reading. In Secondary Education, the situation changed, as no private bilingual schools were considered in our research. Consequently, statistically significant differences were found on all skills tested in favour of the CLIL group from public centres, corroborating Pérez Cañado's (2018) outcomes.

Vis-à-vis RQ3, the results obtained in the post-test phase were compared with the outcomes of the delayed post-test phase. No statistically significant differences were found for the CLIL group for any of the skills tested, except for listening. That is, although CLIL students obtained slightly better results in the delayed post-tests on all the skills considered, the differences were only significant for their oral receptive skill. In the case of non-CLIL students, no statistically significant differences appeared between the two phases, with the exception of reading. Differences in means were also detected for vocabulary, in favour of the delayed post-tests, although they did not reach statistical significance. However, slightly worse results were obtained in the delayed-post phase for use of English and listening, something which clearly differs from CLIL group's results.

In the across-cohort comparisons, the results obtained attest to the supremacy of the bilingual programme, since significant differences arose for all skills considered in favour of CLIL pupils. Coinciding with the results of previous studies (Lorenzo, 2019; Pascual Bajo, 2018; Pladevall-Ballester & Vallbona, 2016), CLIL students

outperform their EFL counterparts on the receptive skills as well as in use of English and knowledge of vocabulary.

Finally, regarding the last RQ, the successive discriminant analyses performed have confirmed that CLIL programmes, together with motivation and SES, are the variables which best explain the differences detected.

As a conclusion, we can affirm our data point to the general improvement of skills in the FL. Hence, our results support the continuity of CLIL programmes in post-Secondary stages, something which would help to consolidate their positive effects among students. In line with Pérez Cañado (2018), our results indicate that, although many of the effects of bilingual education remain, these can gradually disappear if the programmes are discontinued. That is why it is highly recommendable to promote their continued implementation.

Acknowledgements This work was supported by the Spanish Ministry of Economy and Competitiveness [research grant number FFI2012-32221] and the Government of Andalusia [research grant number HUM 2348] (Project: *The Effects of Content and Language Integrated Learning in Monolingual Communities: A Large-Scale Evaluation*).

References

Admiraal, W., Westhoff, G., & de Bot, K. (2006). Evaluation of bilingual secondary education in the Netherlands: Students' language proficiency. *English Educational Research and Evaluation, 12*(1), 75–93.

Anghel, B., Cabrales, A., & Carro, J. M. (2016). Evaluating a bilingual education programme in Spain: The impact beyond foreign language learning. *Economic Inquiry, 54*(2), 1202–1223.

Bruton, A. (2011). Are the differences between CLIL and non-CLIL groups in Andalusia due to CLIL? A reply to Lorenzo, Casal and Moore (2010). *Applied Linguistics, 32*(2), 236–241. https://doi.org/10.1093/applin/amr007.

Bruton, A. (2013). CLIL: Some of the reasons why… and why not. *System, 41*(3), 587–897. https://doi.org/10.1016/j.system.2013.07.001.

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.

European Commission. (1995). *White paper on education and training: Teaching and learning: Towards the learning society*. http://europa.eu/documents/comm/white_papers/pdf/com95_590_en.pdf. Accessed 20 October 2017.

Fernández-Sanjurjo, J., Fernández-Costales, A., & Arias Blanco, J. M. (2019). Analysing students' content-learning in science in CLIL vs. non-CLIL programmes: Empirical evidence from Spain. *International Journal of Bilingual Education and Bilingualism, 22*(6), 661–674. https://doi.org/10.1080/13670050.2017.1294142.

Heras, A., & Lasagabaster, D. (2015). The impact of CLIL on affective factors and vocabulary learning. *Language Teaching Research, 19*(1), 70–88. https://doi.org/10.1177/1362168814541736.

Hüttner, J., & Smit, U. (2014). CLIL (Content and Language Integrated Learning): The bigger picture: A response to: A. Bruton. 2013. CLIL: Some of the reasons why…and why not. *System, 44,* 160–167.

Jäppinen, A. K. (2005). Thinking and content learning of Mathematics and Science as cognitional development in Content and Language Integrated Learning (CLIL): Teaching through a foreign language in Finland. *Language and Education, 19*(2), 147–168.

Jiménez Catalán, R. M., Ruiz de Zarobe, Y., & Cenoz, J. (2006). Vocabulary profiles of English foreign language learners in English as a subject and as a vehicular language. *VIEWZ (Vienna English Working Papers), 15*(3), 23–27.

Lancaster, N. K. (2015). *The effects of Content and Language Integrated Learning on the oral skills of Compulsory Secondary Education students: A longitudinal study*. Jaén: Universidad de Jaén.

Lasagabaster, D. (2009). The implementation of CLIL and attitudes towards trilingualism. *ITL, International Journal of Applied Linguistics, 159*, 23–45.

Lasagabaster, D., & Doiz, A. (2016). CLIL students' perceptions of their language learning process: Delving into self-perceived improvement and instructional preferences. *Language Awareness*. https://doi.org/10.1080/09658416.2015.1122019.

Lasagabaster, D., & Ruiz de Zarobe, Y. (2010). Ways forward in CLIL: Provision issues and future planning. In D. Lasagabaster & Y. Ruiz de Zarobe (Eds.), *CLIL in Spain: Implementation, results and teacher training* (pp. 278–295). Newcastle upon Tyne: Cambridge Scholars Publishing.

Lorenzo, F. (2019). *Educación Bilingüe en Andalucía: Informe de Gestión, Competencias y Organización*. Junta de Andalucía: Consejería de Educación.

Madrid, D., & Barrios, E. (2018). A comparison of students' educational achievement across programmes and school types with and without CLIL provision. *Porta Linguarum, 29,* 29–50.

Madrid, D., Bueno, A., & Ráez, J. (2018). Investigating the effects of CLIL on language attainment: Instrument design and validation. In M. L. Pérez Cañado (Eds.), *Content and Language lntegrated Learning in monolingual settings: New insights from the Spanish context*. Amsterdam: Springer.

Nieto Moreno de Diezmas, E. (2016). The impact of CLIL on the acquisition of language competences and skills in L2 in primary education. *International Journal of English Studies, 16*(2), 81–101.

Ouazizi, K. (2016). The effects of CLIL education on the subject matter (mathematics) and the target language (English). *Latin American Journal of Content and Language Integrated Learning, 9*(1), 110–137. https://doi.org/10.5294/laclil.2016.9.1.5.

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/pplirev-2013-0014.

Pascual Bajo, C. (2018). *The impact of Content and Language Integrated Learning (CLIL) on Valencian Secondary Education students' English language acquisition and academic achievement: A case study*. Valencia: Universidad de Valencia.

Pavón Vázquez, V. (2018). Learning outcomes in CLIL programmes: A comparison of results between urban and rural environments. *Porta Linguarum, 29,* 9–28.

Pelechano Barberá, V. (1994). Prueba MA. *Análisis y Modificación de la Conducta, 20,* 71–72.

Pérez Cañado, M. L. (2011). The effects of CLIL within the APPP: Lessons learned and ways forward. In R. Crespo & M. A. García de Sola (Eds.), *Studies in honour of Ángeles Linde López* (pp. 389–406). Granada: Universidad de Granada.

Pérez Cañado, M. L. (2012). CLIL research in Europe: Past, present and future. *International Journal of Bilingual Education and Bilingualism, 15*(3), 315–341. https://doi.org/10.1080/13670050.2011.630064.

Pérez Cañado, M. L. (2017a). Stopping the 'pendulum effect' in CLIL research: Finding the balance between Pollyanna and Scrooge. *Applied Linguistics Review, 8*(1): 79–99. doi.org/https://doi.org/10.1515/applirev-2016-2001.

Pérez Cañado, M. L. (2017b). The evolution of bilingual education in monolingual settings: An Andalusian case study. In M. Jedynack & P. Romanowski (Eds.), *The many faces of bilingualism-living with two languages* (pp. 207–241). Amsterdam: Springer.

Pérez Cañado, M. L. (2018). CLIL and educational level: A longitudinal study on the impact of CLIL on language outcomes. *Porta Linguarum, 29,* 51–70.

Pérez Cañado, M. L., & Lancaster, N. K. (2017). The effects of CLIL on oral comprehension and production: A longitudinal case study. *Language, Culture and Curriculum, 30*(3), 300–316. doi.org/https://doi.org/10.1080/07908318.2017.1338717.

Pérez Cañado, M. L., & Ráez Padilla, J. (2015). Introduction and overview. In D. Marsh, M. L. Pérez Cañado, & J. Ráez Padilla (Eds.), *CLIL in action: Voices from the classroom* (pp. 1–12). Newcastle upon Tyne: Cambridge Scholars Publishing.

Piesche, N., Jonkmann, K., Fiege, C., & Kebler, J. (2016). CLIL for all? A randomized 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., & Vallbona, A. (2016). CLIL in minimal input contexts: A longitudinal study of primary school learners' receptive skills. *System, 58,* 37–48.

Rallo Fabra, L., & Jacob, K. (2015). Does CLIL enhance oral skills? Fluency and pronunciation errors by Spanish-Catalan learners of English. In M. Juan-Garau & J. Salazar Noguera (Eds.), *Content-based language learning in multilingual educational environments*, 63–177. Berlin: Springer.

Rascón Moreno, D., & Bretones Callejas, C. M. (2018). Socioeconomic status and its impact on language and content attainment in CLIL contexts. *Porta Linguarum, 29,* 115–136.

Rossell, C., & Baker, K. (1996). The educational effectiveness of bilingual education. *Research in the Teaching of English, 30*(1), 7–74.

Ruiz de Zarobe, Y. (2007). CLIL in a bilingual community: Similarities and differences with the learning of English as a foreign language. *Vienna English Working Papers, 16*(3), 47–53.

Ruiz de Zarobe, Y. (2008). CLIL and foreign language learning: A longitudinal study in the Basque Country. *International CLIL Research Journal, 1,* 60–73.

Ruiz de Zarobe, Y. (2011). Which language competencies benefit from CLIL? An insight into applied linguistic research. In Y. Ruiz de Zarobe, J. Sierra, & F. Gallardo del Puerto (Eds.), *Content and foreign language integrated learning* (pp. 129–153). Berne: Peter Lang.

Santamaría, P., Arribas, D., Pereña, J., & Seisdedos, N. (2016). *EFAI: Evaluación Factorial de las Aptitudes Intelectuales.* Madrid: TEA Ediciones.

Serra, C. (2007). Assessing CLIL at primary school: A longitudinal study. *International Journal of Bilingual Education and Bilingualism, 10*(5), 582–602.

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

Vez, J. M. (2009). Multilingual education in Europe: Policy developments. *Porta Linguarum, 12,* 7–24.

Wolff, D. (2005). Approaching CLIL. In D. Marsh (Eds.), *The CLIL quality matrix: Central workshop report*, coord (pp. 10–26). http://www.ecml.at/mtp2/CLILmatrix/pdf/wsrepD3E2005_6.pdf. Accessed 20 October 2017.

The Effects of CLIL on Subject Matter Learning: The Case of Science in Primary and Secondary Education

Juan de Dios Martínez Agudo

Abstract This chapter focuses on an under-researched topic in relation to Content and Language Integrated Learning (CLIL): the effects of bilingual programmes on content learning. It was carried out with a total of ten schools (in Primary and Secondary Education) and 318 students in public bilingual and charter non-bilingual schools in the autonomous community of Extremadura. Students' performance in the subjects of Science in Primary Education and Natural Science in Compulsory Secondary Education is compared across schools in order to provide data on the important issue of whether CLIL programmes are watering down subject matter learning or promoting it as successfully as in monolingual streams. Factor and discriminant analyses are performed with the type of school and educational level variables in order to determine whether CLIL is truly responsible for the differences ascertained or whether other variables account for a greater proportion of the variance. The statistical analysis confirms the positive effects of CLIL on content subject learning, showing that bilingually educated students outperform monolingually educated ones at both educational stages, although this difference is clearly significant when finishing their Secondary Education studies. The results also provide some clear-cut differences in terms of type of school and educational level, as public bilingual schools only outperform charter non-bilingual schools only at the end of Compulsory Secondary Education, which suggests the long-term effects of CLIL on subject matter learning.

1 Introduction

Based on the pioneering immersion programmes in Canada and bilingual programmes in the USA which have led to the progressive introduction of bilingual education worldwide, there has been an explosion of interest and debate in the

J. d. D. Martínez Agudo (✉)
Faculty of Education, University of Extremadura, Badajoz, Spain
e-mail: jdmtinez@unex.es

M. L. Pérez Cañado (eds.), *Content and Language Integrated Learning in Monolingual Settings*, Multilingual Education 38,
https://doi.org/10.1007/978-3-030-68329-0_9

last two decades about the potential of the European approach to bilingual education—CLIL (Content and Language Integrated Learning)—, not only in Europe but also around the world. CLIL has been conceived as an alternative to the Communicative Language Teaching (CLT) approach (Coyle et al., 2010), an extension of it (Dalton-Puffer, 2007; Lasagabaster & Sierra, 2010) or just a new paradigm of education (Ouazizi, 2016). While Coyle et al. (2010) do not view CLIL as 'simply another step in language teaching or a new development in content-subject methodology' (p. ix) but rather as 'a major rethink of how we teach what we teach' (pp. ix–x), Dalton-Puffer (2011, p. 195) is cautious and warns that, contrary to most expectations, CLIL is not a panacea, as evidenced by the fact that some downsides have been recently reported by CLIL research as well as the fact that there is a wide gap between what is provided in CLIL teaching and what comes out in terms of CLIL learning. Due to the idiosyncrasies of the European context and the unceasing search for improved language teaching methods aimed at increasing L2 competence at a time when European integration is idealised, what becomes clear is that CLIL has emerged as a 'timely solution to European plurilingual education' (Pérez Cañado, 2012, p. 315) in an increasingly globalised world where '(bi)multilingualism is the norm whereas monolingualism is the exception' (Ouazizi, 2016, p. 113).

Regarding the different conceptualisations of CLIL viewed as an umbrella construct which lacks conceptual precision as well as the fact that its scope is still not clear-cut, Cenoz et al. (2014) recognise that the varied interpretations of the CLIL approach suggest that CLIL is understood in different ways, which makes it difficult to pin down its uniqueness. For example, the language and content balance as well as the intensity of the exposure to the foreign language, among other CLIL core characteristics, are understood in different ways, as Cenoz et al. (2014) point out. Based on the succinct definition by Coyle et al. (2010, p. 1), 'Content and Language Integrated Learning (CLIL) is a dual-focused educational approach in which an additional language is used for the learning and teaching of both content and language', what becomes clear is that the core feature of CLIL is, no doubt, its dual-focused nature, as this 'two for one' approach strives to promote the integration of both L2 learning and content learning. Such a duality is certainly seen as the main strength of this educational approach, as both foreign language and subject matter content should be learnt and taught in an integrated way. Needless to say, such integration, in turn, involves the major challenge facing CLIL teachers and learners; as Ruiz de Zarobe (2013, p. 235) claims, 'the challenge remains of how to enable learners to make best use of both areas in the classroom'. In the same vein, Cenoz et al. (2014, p. 244) also make it clear that

> The dual role of language and content has been understood in different ways. According to Ting (2010: 3), 'CLIL advocates a 50:50/Content: Language CLIL-equilibrium'. However, research conducted in actual CLIL classrooms shows that it is difficult to achieve a strict balance of language and content (Dalton-Puffer, 2007; Mehisto, 2008; Pérez-Vidal & Juan-Garau, 2010).

Besides its dual nature, another essential feature of CLIL pedagogy is precisely its diversity (Cenoz, 2015; Coyle, 2010). This diversity of models or formats is a visible

trend in the European context. Coyle (2007, p. 49) considers that there is no model 'which suits all CLIL contexts'. Certainly, CLIL comprises many different variants as it has been implemented in a variety of forms since the 1990s (Cenoz et al., 2014). In relation to this diversity, Coyle (2008, p. 101) also makes it clear that 'there is a lack of cohesion around CLIL pedagogies. There is neither one CLIL approach nor one theory of CLIL'. In the same vein, Mehisto et al. (2008, p. 12) claim that CLIL implementation takes different forms as 'CLIL is an umbrella term covering a dozen or more educational approaches (e.g. immersion, bilingual education, multilingual education, language showers and enriched language programmes)'. Similarly, Ruiz de Zarobe (2013, p. 233) states that almost all EU states implement some form of CLIL with varying degrees of success, responding in different ways (Eurydice, 2006): 'Under the acronym CLIL we recognize a wide range of models, which show divergences as regards the age of implementation of the model or the intensity of the exposure to the foreign language (…), to name but a few differences'. According to Cenoz (2015, p. 21), 'There is great diversity in the implementation of CBI/CLIL programmes and these programmes are dynamic and change because they have to keep up with new challenges in society'. To be more specific, such diversity can be seen in the differences observed in teaching methodology, as some programmes are more content-oriented than others (Cenoz, 2015). In this respect, Kerstin (2013) also concludes that discrepancies in results obtained across CLIL contexts in Europe may be due to nation-specific contextual factors such as policy framework, teacher education, age of implementation and extramural exposure to English. Lastly, Cenoz et al. (2014, p. 258) also suggest that 'Rather than insisting on the uniqueness of CLIL, efforts might be better spent establishing a taxonomy of different common forms of CLIL/CBI so as to circumscribe the diverse contexts in which CLIL is found'. Needless to say, this diversity of CLIL programme formats also involves great challenges when carrying out research on CLIL, as Cenoz et al. (2014) rightly point out.

Much of what we currently know about CLIL approach comes from Applied Linguistics research (Marsh & Frigols, 2013). Specifically, Second Language Acquisition (SLA) research studies provide, as Lasagabaster and López (2015, p. 43) remind us, 'some arguments in favour of CLIL programmes on the grounds that they create conditions for naturalistic language learning, increase the time of exposure to the foreign language and provide an aim for language use in the classroom (Dalton-Puffer & Smit, 2007)'. While several theoretical arguments propose that CLIL promotes content learning, other theories, in contrast, suggest negative effects. Based on information processing theories, Piesche et al. (2016, p. 109) argue that bilingual education in general and CLIL contexts in particular are assumed to lead to 'a greater cognitive control and selective attention which prevents the working memory from being overloaded and thus leads to more effective cognitive processes'. Additionally, it is also made clear that bilingual students are expected 'to process information more deeply because they have to invest more mental effort' (Piesche et al., 2016, p. 109). On the contrary, the perspective of cognitive load theory (Sweller et al., 2011) sustains that students' working memory is overloaded by simultaneously processing new content and the foreign language.

The potential of CLIL in terms of linguistic and cognitive benefits has been fully discussed and documented within the international research literature over the last two decades (Casal & Moore, 2009; Cenoz, 2015; Coyle, 2002; Coyle et al., 2010; Dalton-Puffer, 2008; Halbach, 2008; Lasagabaster & Ruiz de Zarobe, 2010; Muñoz, 2002; Madrid & Hughes, 2011; Pérez Cañado, 2016, 2017, 2018; cf. also Chapters 'CLIL and L1 Competence Development', 'The Impact of CLIL on FL Grammar and Vocabulary', and 'The Effects of CLIL on FL Learning: A Longitudinal Study' in this volume). However, recent studies seem to move beyond the impact of CLIL in terms of linguistic benefits and, consequently, address its effects on the development of content subject knowledge, which has been a neglected research area so far. In relation to this, Nikula (2016) makes it clear that there seems to be a shift in emphasis in CLIL research from studies focusing exclusively on the potential of CLIL in terms of L2 learning outcomes to studies that point towards the need to adopt a truly integrated view on language and content, thus exploring the effects of CLIL on the development of content subject learning. Very little is known for certain about the effects of CLIL on the development of subject matter knowledge (Dalton-Puffer, 2011). Such effects still remain unclear, as Nikula and Mard-Miettinen (2014, p. 14) rightly state: 'the overall image to date remains rather inconclusive, suggesting that this is an area where more research is needed'. By comparing different models of bilingual education, Cenoz et al. (2014, p. 10) also recognise that 'there is a greater focus on language than on academic achievement in Canadian immersion research, the same can be said of research on CLIL where research on content is extremely limited'. This argument is also supported by Pérez Cañado (2012, p. 315) who claims that 'there is still a well-documented paucity of research in this area'. Such lack of research into content outcomes may be due to the fact that 'CLIL research is conducted by language educators rather than subject specialists, and therefore focuses almost exclusively on language, with content knowledge rarely examined or measured' (Paran, 2013, p. 323). In the same vein, Cenoz et al. (2014, p. 257) argue that 'Specifically, much, if not most, research on CLIL has been conducted by ESL/EFL scholars who have compared CLIL and non-CLIL groups of learners and reported higher achievement in English for CLIL learners'. Similarly, Fernández-Sanjurjo et al. (2017, p. 1) argue that 'So far, CLIL research has focused primarily on language attainment in the L2 and the L1, but students' achievements as regards content subjects have been largely ignored'. Accordingly, as Dallinger et al. (2016, p. 25) conclude, 'the effects of CLIL on content learning remain an open question'. What becomes clear is future CLIL research agenda needs to address this under-researched strand in depth (Cenoz et al., 2014; Lasagabaster & Ruiz de Zarobe, 2010; Paran, 2013; Pérez Cañado, 2016, 2017) because 'we simply do not have enough evidence' (Paran, 2013, p. 331). The few existing research studies focusing on CLIL-effects on academic content learning to date are in fact contradictory and present mixed results, as Piesche et al. (2016, p. 109) remind us. In the European CLIL context, while most of the studies conducted so far report positive outcomes for academic content learning, other studies have recently found no differences between bilingual and non-bilingual students in terms of content knowledge and some studies have even revealed negative effects on content subject competence, as will be described below in further detail. Accordingly,

as will be examined below, CLIL research offers contradictory results which vary across European contexts.

Overall, CLIL research has provided empirical evidence on the benefits of CLIL education on content learning, concluding that bilingual learners assimilate the content of the academic subjects at the same pace or even better than their non-bilingual counterparts (Jäppinen, 2005; Murray, 2010; Madrid & Hughes, 2011; Mattheoudakis et al., 2014; Ouazizi, 2016; Pérez Cañado, 2012, 2016, 2017, 2018; Serra, 2007; Surmont et al., 2016; Ullmann, 1999; Wode, 1999; Xanthou, 2011). The study by Ullmann (1999) in the United Kingdom reveals that CLIL Secondary Education students assimilating subject-contents (in History and Geography) show enhanced subject matter learning. In Germany, Wode (1999) also reports that CLIL Secondary Education students perform better in Geography and History than their monolingual peers. The longitudinal study by Jäppinen (2005) conducted in Finland concludes that content subject learning, Maths and Science particularly, might be promoted by CLIL as a result of the stimulation of cognition/thinking processes which seem to have positive repercussions on subject matter learning. Similarly, the longitudinal study by Serra (2007) in Switzerland reveals that CLIL Primary Education students obtain higher scores in Mathematics than their non-CLIL counterparts. The cross-sectional study performed in Spain by Madrid (2011) reports that CLIL students learning History and Geography perform better than their non-CLIL peers. Identical results are also reported by Xanthou (2011), whose study with CLIL Primary Education students in Cyprus shows that assimilating academic contents (Science, specifically) through English is beneficial. In the same vein, the study by Madrid and Hughes (2011) in Spain also provides positive results for CLIL in terms of academic content learning (Science in Primary Education and Social Science in Secondary Education) by factoring in type of school as an intervening variable, as in the present study. The study by Mattheoudakis et al. (2014) in Greece, where CLIL was introduced as a pilot project in 2010, also reveals both language and content gains for CLIL students learning Geography in the context of Primary Education. Ouazizi (2016) also reports positive findings with CLIL Secondary Education students learning Mathematics in Belgium, concluding that CLIL exerts a positive influence on content knowledge due to the cognitive benefits of CLIL which seems to stimulate cognitive flexibility (Coyle et al., 2010) and/or cognitive development. Lastly, Surmont et al.'s (2016) study carried out in Belgium also reports that CLIL appears to have a positive impact on the mathematical knowledge of Secondary Education students, even after a very short period of time (three months). As shown by all these studies conducted in various European countries, subject matter knowledge is positively affected by the CLIL approach. Contrary to what might be expected, Van de Craen et al. (2007) hold that subject matter knowledge is not of less quality in CLIL than in traditional education.

A neutral position is also visible in the present discussion as different studies have reported no differences between monolingually and bilingually educated students concerning their content subject knowledge. For example, Bergroth (2006) argues that CLIL students learning Mathematics in Finland do not obtain lower results than their non-CLIL counterparts when finishing their Secondary Education studies,

and indeed perform just as well as their non-CLIL peers. In the same vein, the longitudinal study conducted in the Netherlands by Admiraal et al. (2006) also reports that no negative impact was found in CLIL Secondary Education students' content knowledge in History and Geography. Similar results are also reported by Stehler (2006) in Switzerland, who concludes that CLIL has neither a positive nor a negative influence on academic content knowledge. The quantitative and qualitative research study carried out by Alonso et al. (2008) in the Basque autonomous community in Spain relating to the effectiveness of plurilingual education through CLIL approach in Secondary schools concludes that the assimilation of academic contents taught in English is similar, if not superior, to those relating to non-CLIL students. All in all, these studies reveal that academic content knowledge is not threatened by CLIL in view of a lack of differences observed between both student cohorts (CLIL/non-CLIL).

At the opposite end of the debate are those recent studies which report the negative effects of CLIL on content subject knowledge (Anghel et al., 2016; Dallinger et al., 2016; Fernández-Sanjurjo et al., 2017; Piesche et al., 2016; Sotoca, 2014). For example, the study by Sotoca (2014) conducted in bilingual and non-bilingual public Primary Education schools in Madrid (Spain) reports statistically significant differences in favour of non-bilingual schools in Science, which may be due, according to the author, to a greater level of exigency for academic subjects in bilingual schools. The study carried out in Spain by Anghel et al. (2016) factored in parents' educational level and also reveals significantly negative effects in Natural Science knowledge for those CLIL Primary Education students of less educated parents. Another research study also conducted in Spain is that of Fernández-Sanjurjo et al. (2017), who conclude that monolingual students learning Science achieve better results than their bilingual peers. Similarly, the study performed in Germany by Piesche et al. (2016) shows that monolingually educated students outperform bilingually educated ones in learning Science, although it is also made clear that the negative effects of CLIL on students' content learning are small. In the same vein, Dallinger et al. (2016) in Germany also report a negative CLIL-effect on content learning, concluding that CLIL students progress more slowly and need to receive more input to achieve the same results in terms of content learning. In short, all these studies report a detrimental effect of CLIL education on academic achievement. Additionally, several research studies also point towards students' difficulties in expressing subject knowledge through the foreign language (Jäppinen, 2005; Piesche et al., 2016). Perhaps one convincing reason for this CLIL negative effect might be, as Marsh et al. (2000) identified, the high linguistic demands of the content areas.

Once the initial euphoria of this innovative educational approach has passed since its emergence on the European scene in 1994, a more critical attitude has recently emerged in response to the need to address some 'problematic issues of CLIL' (Paran, 2013, p. 334), calling into question certain controversial aspects or challenges. In relation to the present CLIL research scenario, Pérez Cañado (2018, p. 20) argues that

> the so-called 'pendulum effect' (Pérez Cañado, 2016, p. 1) can be seen at work within the CLIL research scenario, as we have moved from an initial period of unbridled enthusiasm and 'celebratory rhetoric' (Paran, 2013, p. 334) on the effects of CLIL to a more critical moment (…) a much more pessimistic outlook on CLIL implementation.

Despite the widely recognised benefits attributed to CLIL approach, certain critical voices have recently warned about the possible drawbacks of this approach (Bruton, 2011, 2013, 2015), thus making it clear that the initial enthusiasm for CLIL should not neglect the real challenges of this new educational approach (Fernández-Sanjurjo et al., 2017; Paran, 2013; Pérez Cañado, 2016, 2017). By exclusively insisting on the uniqueness and potential of CLIL (often without substantial empirical evidence), what is true is that CLIL shortcomings have not been addressed in detail. Hence, Cenoz et al. (2014, p. 256) recognise that 'There is a need for more balanced reflection on both the strengths and shortcomings or gaps in our understanding of CLIL and its effectiveness in diverse contexts'. Given the ambiguity of CLIL, researchers like Cenoz (2013) and Cenoz et al. (2014) demand more critical research beyond exclusively analysing CLIL language gains. While it is true that research results have confirmed the benefits of CLIL in terms of L2 competence, the effects of CLIL on content knowledge, in contrast, still remain an open research question, an unexplored research terrain, as little is known for certain about its real effects on the development of the subject content knowledge (Cenoz et al., 2014; Lasagabaster & Ruiz de Zarobe, 2010). In relation to this under-researched topic, Pérez Cañado (2018, p. 20) concludes that 'the research carried out thus far presents potentially serious flaws which could compromise the validity of its outcomes'. Since CLIL research has recently pointed out the neglect of influential intervening variables which need to be examined in detail, Pérez Cañado (2012, p. 330) argues that there is a 'need of solid empirical research which builds in rigorous assessment of the variables under scrutiny (…) to determine whether the gains observed are truly ascribable to CLIL practice'. Further investigation is also needed on the way language and content are integrated into CLIL classrooms. In view of such empirical gaps in our understanding of CLIL effectiveness, Cenoz et al. (2014, pp. 256–257) point out that 'Without empirical evidence concerning these issues, we simply do not know (…) there is a need to examine more carefully if content is acquired to the same extent when taught through the medium of the L2 in comparison with students' native language'. Additionally, Cenoz et al. (2014, p. 257) also clarify that 'Although these results provide general support for CLIL (although see Bruton, 2011 for an opposing view), they do not establish a clear causal link between integrated language and content teaching and learner outcomes'. Before leaving this discussion, it is undeniable that the development of CLIL pedagogy in the European context presents both strengths and weaknesses, hence the need for a more critical classroom-based research on CLIL, as Cenoz et al. (2014, pp. 258–259) suggest,

> We believe that it is time for CLIL scholars to move from celebration to a critical empirical examination of CLIL in its diverse forms to better identify its strengths and weaknesses in different learning contexts (…) In other words, research is needed that goes beyond examining simply whether teaching content in an L2 or a foreign language promotes L2 competence to examining how teaching content in an L2 works and how it can be improved.

> Classroom-based research on how best to integrate language and content is necessary if we are to enhance teacher effectiveness in CLIL settings (…) However, there are many aspects of the integration of language and content instruction that require careful theoretical, empirical, and pedagogical attention.

2 Research Questions

Given the scarcity of research studies addressing the effects of CLIL approach on content subject learning in monolingual contexts in Spain (Anghel et al., 2016; Fernández-Sanjurjo et al., 2017; Madrid & Hughes, 2011; Pérez Cañado, 2018; Sotoca, 2014), this chapter aims to shed some light on this still under-researched topic, assessing whether CLIL programmes water down content subject knowledge or rather promote it as successfully as in monolingual streams. Bearing in mind the literature reviewed so far on the effects of CLIL on content subject knowledge, this chapter aims to address the following research questions:

RQ1: Does CLIL education positively or negatively affect subject content knowledge?
RQ2: Does CLIL education lead to equal or better subject matter knowledge than traditional education?
RQ3: When do positive CLIL-effects become visible, in the short or long term?
RQ4: What is the differential effect exerted on the Primary and Secondary CLIL students' Science learning outcomes by the following two intervening variables: type of school (public and charter) and educational stage (Primary and Secondary Education)?

3 Method

This study forms part of a broader research project focusing on a three-year longitudinal large-scale evaluation of CLIL programmes carried out in those Spanish monolingual communities with the least tradition in bilingual education (Andalusia, Extremadura and the Canary Islands). In view of the scarce research literature available so far on the effects of CLIL on subject matter learning which presents contradictory empirical evidence, the main emphasis of this quantitative study is on the impact of CLIL education on students' Science subject knowledge at the end of Primary (6th grade) and Secondary (4th grade) Education.

Efforts have been made to ensure the homogeneity of the experimental (CLIL) and control (non-CLIL) groups in terms of motivation, verbal intelligence and English level. Pre-, post- and delayed post-tests were administered to Primary and Secondary Education students. In view of the very limited number of research studies focused on controlling the differential effect of particular intervening variables, factor and discriminant analyses were consequently conducted to ascertain the relationship or interaction between CLIL education and the intervening variables under control in

this study (type of school—public and charter—and educational stage—Primary and Secondary Education—) which may account for the differences detected between both student cohorts. To be more specific, dependent (content subject learning results), independent (CLIL programmes) and intervening (type of school and educational level) variables have been taken into consideration in the present study so as to determine whether CLIL is truly responsible for the potential differences observed or whether the aforementioned intervening variables can account for a greater proportion of the variance. Lastly, Cohen's *d* was employed to measure effect sizes.

3.1 Context and Participants

The context of the present study is the monolingual autonomous community of Extremadura, which is situated in the south-west of Spain, on the border with Portugal, and which has very little tradition in bilingual education (from 2004 onwards). At the present time there are 274 CLIL schools in Extremadura at Primary and Secondary Education stages.

The sample under control comprises 318 students from 10 schools (public and charter). The control group (non-CLIL) consists of 162 learners, while the remaining 156 learners form the experimental group (CLIL). Accordingly, the achievement results of both student cohorts on the subjects of Science in Primary Education and Natural Science in Secondary Education are compared across schools, examining the impact of CLIL on the intervening contextual variables (type of school and educational level). It is noteworthy that no private school participated in the present study, so the comparison with this type of school has not been possible in Extremadura. Table 1 provides an outline of the participating sample.

3.2 Instrument

The data were gathered through an initial questionnaire aimed at collecting personal data on the participants such as age and educational stage. Science subject knowledge was measured by CLIL students' final grades provided by the participating schools out of a total score of 10, which is the highest grade in the Spanish educational system.

4 Results and Discussion

RQ1 investigates whether subject matter knowledge is positively or negatively affected by CLIL education. As can be observed in Table 2, the results of our

Table 1 The research sample

		N	Mean (%)
Student cohorts			
	CLIL	156	49.1
	Non-CLIL	162	50.9
Gender			
	Male	174	54.7
	Female	144	45.3
Educational stages			
	Primary Education (6th grade)	162	50.9
	Compulsory Secondary Education (4th grade)	156	49.1
Type of school			
	Public	252	79.2
	Charter	66	20.8
Setting			
	Urban	113	35.5
	Rural	205	64.5

Table 2 Mean difference scores of the experimental (CLIL) and control (non-CLIL) groups on the subject matter achievement results at both educational stages

Educational level	Group	Mean	Standard deviation	Cohen's *d*	*p* value
Primary Education	Non-CLIL	6.98	1.694	−0.110	.528
	CLIL	7.16	1.587		
Secondary Education	Non-CLIL	6.21	1.817	−0.497	.002
	CLIL	7.03	1.443		

analysis confirm the positive effects of CLIL programmes on the development of content subject knowledge by comparing the resulting data of both student cohorts (CLIL/non-CLIL) (Martínez, 2020). This result is backed up by numerous research studies which indicate the positive effects of CLIL education on content subject learning (Jäppinen, 2005; Madrid & Hughes, 2011; Mattheoudakis et al., 2014; Murray, 2010; Ouazizi, 2016; Pérez Cañado, 2012, 2018; Serra, 2007; Surmont et al., 2016; Ullmann, 1999; Wode, 1999; Xanthou, 2011).

Once this positive effect has been reported, RQ2 analyses whether CLIL education leads to equal or better subject matter knowledge than traditional educational approaches, particularly whether bilingually educated students learning Science in Primary Education and Natural Science in Secondary Education perform equally well or outperform their monolingually educated peers. Unlike Fernández-Sanjurjo et al. (2017) and Piesche et al.'s (2016) studies, which show that monolingually educated students perform slightly better than bilingually educated ones when learning subject

matter knowledge, the results of the present study confirm the opposite view, as CLIL students' learning gains are higher than their non-CLIL counterparts' at both educational stages, but especially at the end of Compulsory Secondary Education. According to the results of the present study, bilingual learners assimilate the subject matter content at more or less the same pace in Primary Education, but clearly outperform their non-bilingual peers at the end of Compulsory Secondary Education. This is in line with several research studies which show comparable or even better results between both student cohorts (CLIL/non-CLIL) regarding content subject knowledge. For example, Mattheoudakis et al. (2014) confirm that content knowledge is clearly not negatively affected by CLIL education, reporting that CLIL Greek learners score higher than their non-CLIL counterparts in Geography tests. In the same vein, Ouazizi (2016) also concludes that CLIL education leads to better subject matter knowledge than traditional learning approaches, as Belgian CLIL students obtain better scores than monolinguals in Mathematics knowledge. Such a difference in global performance between both student cohorts may be due, among other aspects, both to the prior careful selection and to the high motivation and interest on the part of the families and students involved in such bilingual programmes, as suggested by Alonso et al. (2008). As can be seen in Table 2, while no statistically significant differences emerge between the experimental (CLIL) and control (non-CLIL) groups at the end of Primary Education as Cohen's *d* is quite low, the differences between both cohorts are, in contrast, statistically significant when finishing their Secondary Education studies, with a higher Cohen's *d*.

Given that RQ3 addresses whether the impact of CLIL education on content learning becomes visible in the short or long term, the results of this study confirm that the differences in academic achievement results between the experimental group (CLIL) and the control group (non-CLIL) are higher or become more visible in the long term, particularly when finishing their Secondary Education studies, in line with other studies in the Spanish context (Alonso et al., 2008; Madrid & Hughes, 2011; Pérez Cañado, 2018). While CLIL students obtain similar scores or slightly outperform their non-CLIL peers concerning Science knowledge at the end of Primary Education, bilingually educated students clearly outstrip their monolingually educated counterparts when finishing their Secondary Education studies. This difference in achievement results seems to become more visible as time goes by. Perhaps this may be due to the influence of accumulated experience in bilingual education. In relation to the impact of such experience, the study by Piesche et al. (2016) reminds us of the possible negative effects of CLIL education on content learning for students without CLIL experience. In short, this study reveals that the positive effects of CLIL education require a longer period of time, after which they will become more visible. However, this finding is not congruent with those obtained by Van de Craen et al. (2007), who reported that subject matter learning through CLIL education seems to be boosted more significantly in Primary Education than in Secondary Education. This result is not congruent either with the longitudinal study by Surmont et al. (2016), who conclude, in contrast to their expectations, that CLIL education's positive effects become visible even after a very short period of time (three months).

Our last research question inquires into the differential effect which the intervening contextual variables under control in this study (type of school and educational stage) exert on the Primary and Secondary CLIL students' Science learning outcomes. Consequently, public bilingual and monolingual schools, as well as charter monolingual ones, were compared in this study.

Considering only public bilingual schools, bilingual students achieve better results in Science than their non-bilingual counterparts at the end of both educational stages, which corroborates the benefits of CLIL in terms of content subject learning. Such results corroborate Madrid and Hughes's (2011) and Pérez Cañado's (2018) findings, relating to the fact that bilingual strands outstrip monolingual ones at the end of Primary and Secondary Education in public schools. While no substantial differences were observed between both student cohorts at the end of Primary Education, statistically significant differences were, in contrast, found in favour of CLIL learners when finishing their Secondary Education studies, in this case, with a higher Cohen's *d*, as can be seen in Table 3.

By comparing both public bilingual schools and charter non-bilingual schools, the resulting data surprisingly reveal that charter monolingual schools obtain slightly better results in Science than public bilingual schools only at the end of Primary Education. In this respect, unsubstantial differences were detected in view of the low Cohen's *d*. However, statistically significant differences between public bilingual branches and charter non-bilingual ones were found in favour of the former at the end of Compulsory Secondary Education, with a higher Cohen's *d*. In other words, the results of this study suggest that public bilingual schools outstrip charter non-bilingual schools only at the end of Secondary Education, which confirms once again that the positive effects of CLIL on content learning are mainly witnessed or verified in the long term. However, this finding completely differs from Madrid and Hughes's (2011) study, in which charter monolingual schools obtained significantly better results at the end of Compulsory Secondary Education, thus outperforming public bilingual ones even in the long term.

Lastly, it is noted that charter non-bilingual schools outperform public non-bilingual ones at the end of Primary Education, although the differences cannot be considered statistically significant, with a relatively low Cohen's *d*. However, similar results are obtained by both cohorts in both types of schools at the end of Secondary Education, with an extremely low Cohen's *d*. Consequently, such a finding is not congruent with that obtained by Madrid and Hughes (2011), who reported that the public monolingual strands lagged behind the rest of the groups at both educational stages.

Based on the discriminant analyses performed, statistically significant differences in fact emerge between the experimental (CLIL) and control (non-CLIL) groups in terms of the different intervening variables in this study (type of school and educational level). As can be seen in the tests of equality of group means, the discriminating potential of such variables becomes visible. To be more specific, Wilks' Lambda test reports that there are differences between the mean scores of both student cohorts on the content subject results, particularly at the end of Compulsory Secondary Education. In short, such differences between the experimental (CLIL) and control

Table 3 Subject content results according to educational level and type of school

Educational level	Type of school/Group	Mean	Standard deviation	Cohen's *d*	*p* value
Primary Education	Public school/Non-CLIL	6.72	1.508	−0.281	.178
	Public school/CLIL	7.16	1.587		
	Charter school/Non-CLIL	7.35	1.898	0.114	.638
	Public school/CLIL	7.16	1.587		
	Public school/Non-CLIL	6.72	1.508	−0.375	.177
	Charter school/Non-CLIL	7.35	1.898		
Educational level	Type of school/Group	Mean	Standard deviation	Cohen's *d*	*p* value
Secondary Education	Public school/Non-CLIL	6.22	1.821	−0.497	.005
	Public school/CLIL	7.03	1.443		
	Charter school/Non-CLIL	6.17	1.855	−0.562	
	Public school/CLIL	7.03	1.443		.035
	Public school/Non-CLIL	6.22	1.821	0.027	
	Charter school/Non-CLIL	6.17	1.855		.915

(non-CLIL) groups cannot be exclusively ascribed to the impact of CLIL education, as the type of school and educational level contextual variables also have a significant influence in explaining the differences found between both cohorts, as can be seen in Tables 4 and 5.

All in all, statistical analysis allows us to conclude that CLIL education does not negatively affect subject content knowledge but rather the opposite. Additionally, the

Table 4 Test of equality of group means

Primary Education	Wilk's Lambda	*F*	df1	df2	Sig.
Type of school	.568	121.481	1	160	.000
Secondary Education	Wilk's Lambda	*F*	df1	df2	Sig.
Type of school	.882	20.546	1	154	.000
Subject content results	.941	9.596	1	154	.002

Table 5 Summary of canonical discriminant functions

Primary education				
Function	Eigenvalue	% of variance	Cumulative %	Canonical correlation
1	.759	100.0	100.0	.657
Test of function	Wilks' Lambda	Chi-square	df	Sig.
1	.568	90.100	1	.000
Secondary education				
Function	Eigenvalue	% of variance	Cumulative %	Canonical correlation
1	.194	100.0	100.0	.403
Test of function	Wilks' Lambda	Chi-square	df	Sig.
1	.838	27.129	2	.000

results also reveal that the effects of CLIL education (as an independent variable) are substantial on the content subject learning results (as the dependent variable), especially at the end of Compulsory Secondary Education; that is, positive CLIL-effects are particularly felt in the long term, which corroborates Pérez Cañado's (2018) findings. The discriminating potential of type of school and educational level (as intervening contextual variables) may also account for the differences detected between the experimental (CLIL) and control (non-CLIL) groups.

5 Conclusion

After reviewing what the research literature has revealed so far, the present study aims to shed some light on this unexplored research topic, providing updated empirical evidence on the positive effects of CLIL education on the development of subject matter knowledge in Primary and Secondary Education when compared to traditional educational approaches in a monolingual Spanish region with very little bilingual education tradition.

In response to RQ1, the results of the present study confirm that subject matter knowledge is not diminished or detrimentally affected by the impact of CLIL education, but quite the opposite. Turning now to RQ2 and in view of the data obtained, it can be concluded that CLIL education strands lead to better subject matter knowledge than traditional mainstream school programmes. While no substantial differences are found at the end of Primary Education, statistically significant differences are, in contrast, detected when finishing their Secondary Education studies. In relation to RQ3, the results suggest that positive CLIL-effects are clearly observable in the experimental group (CLIL students) at the end of Compulsory Secondary Education, which indicates that positive CLIL-effects become more noticeable in the long term. As regards the last RQ, the results allow us to state that the two intervening contextual variables (type of school and educational level) have a discriminating

potential, as bilingually educated students obtain better results than monolingually educated students at both educational stages in public schools only. However, when comparing both public bilingual and charter non-bilingual schools, the results surprisingly reveal that there are clear-cut differences in the learning achievement results of both student cohorts depending on type of school and educational stage: while charter non-bilingual strands outperform public bilingual schools at the end of Primary Education (unsubstantial differences), the bilingually educated students' learning gains, in contrast, are higher than the monolingually educated ones when finishing their Secondary Education studies (statistically significant differences). Perhaps a possible convincing reason for the higher scores of bilingually educated students in Natural Science at the end of Compulsory Secondary Education lies precisely in the value of the accumulated experience in bilingual education. To a lesser extent, the discriminating potential of the intervening variables is also observed when comparing both non-bilingual charter schools and non-bilingual public ones, since while the former outperform the latter at the end of Primary Education, similar achievement results are surprisingly found at the end of Compulsory Secondary Education in both types of schools. It must be added that no statistically significant differences are in fact detected at both educational stages in both types of schools. All in all, the results of the present study confirm the differential effect of the two intervening contextual variables of this study, which may account for the differences ascertained between both student cohorts.

In short, the findings of this study confirm the educational value and effectiveness of CLIL education in comparison with traditional educational approaches as the experimental group (CLIL) obtains better results in subject matter knowledge than the control group (non-CLIL), especially in the long term. In this chapter, two factors have been targeted as influential for explaining variation in content subject learning results: type of school and educational stage.

Since the current state of CLIL research is somewhat sparse, methodologically limited and contradictory (Cenoz et al., 2014; Dallinger et al., 2016), further longitudinal research studies are certainly needed in this direction to investigate the real impact of CLIL education on the development of subject matter knowledge. Particularly, Mattheoudakis et al. (2014) advocate the need for further investigations into the strategies CLIL learners use in order to comprehend the concepts presented in the foreign language. Future research studies need to address the impact of CLIL on subject matter knowledge, over shorter and longer periods of time, but also in different learning contexts and with different age groups, as suggested by Surmont et al. (2016). Lastly, the emotional impact of CLIL education on subject matter learning, which remains an unexplored research area to date, calls for further investigation in the future so as to be able to understand how and under what affective and contextual conditions content subject learning actually develops.

Acknowledgements This work was supported by the Spanish Ministry of Economy and Competitiveness [research grant number FFI2012-32221] and the Government of Andalusia [research grant number P12-HUM-2348] (Project: *The Effects of Content and Language Integrated Learning in Monolingual Communities: A Large-Scale Evaluation*).

References

Admiraal, W., Westhoff, G., & de Bot, K. (2006). Evaluation of bilingual Secondary Education in the Netherlands: Students' language proficiency in English. *Educational Research and Evaluation, 12*(1), 75–93.

Alonso, E., Grisaleña, J., & Campo, A. (2008). Plurilingual education in Secondary schools: Analysis of results. *International CLIL Research Journal, 1*(1), 36–49.

Anghel, B., Cabrales, A., & Carro, J. M. (2016). Evaluating a bilingual education program in Spain: The impact beyond foreign language learning. *Economic Inquiry, 54*(2), 1202–1223.

Bergroth, M. (2006). Immersion students in the matriculation examination three years after immersion. In S. Björklund, K. Mard-Miettinen, M. Bergström, & M. Södergard (Eds.), *Exploring dual-focussed education: Integrating language and content for individual and societal needs* (pp. 123–134). Vaasa: University of Vaasa, Centre for Immersion and Multilingualism.

Bruton, A. (2011). Is CLIL so beneficial, or just selective? *Re-evaluating Some of the Research. System, 39*(4), 523–532. https://doi.org/10.1016/j.system.2011.08.002.

Bruton, A. (2013). CLIL: Some of the reasons why… and why not. *System, 41,* 587–597. https://doi.org/10.1016/j.system.2013.07.001.

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.

Casal, S., & Moore, P. (2009). The Andalusian bilingual sections scheme: Evaluation and consultancy. *International CLIL Research Journal, 1*(2), 36–46.

Cenoz, J. (2013). Discussion: Towards an educational perspective in CLIL language policy and pedagogical practice. *International Journal of Bilingual Education and Bilingualism, 16,* 389–394. https://doi.org/10.1080/13670050.2013.777392.

Cenoz, J. (2015). Content-based instruction and Content and Language Integrated Learning: The same or different? *Language, Culture & Curriculum, 28*(1), 8–24. https://doi.org/10.1080/07908318.2014.1000922.

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.

Coyle, D. (2002). Relevance of CLIL to the European Commission's language learning objectives. In D. Marsh (Ed.), *CLIL / EMILE—The European dimension: Actions, trends, and foresight potential* (pp. 27–28). DG Education & Culture, European Commision, University of Jyväskylä.

Coyle, D. (2007). The CLIL quality change. In D. Marsh & D. Wolff (Eds.), *Diverse contexts-converging goals: CLIL in Europe* (pp. 47–58). Bern: Peter Lang.

Coyle, D. (2008). CLIL—A pedagogical approach from the European perspective. In M. Van Dusen-Scholl & N. H. Hornberger (Eds.), *Encyclopedia of language and education* (pp. 97–111). New York: Springer.

Coyle, D. (2010). Preface. In D. Lasagabaster & Y. Ruiz de Zarobe (Eds.), *CLIL in Spain: Implementation, results and teacher training.* Newcastle upon Tyne: Cambridge Scholars Publishing.

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

Dallinger, S., Jonkmann, K., Hollm, J., & Fieg, C. (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. (2007). *Discourse in Content and Language Integrated (CLIL) classrooms.* Amsterdam: John Benjamins.

Dalton-Puffer, C. (2008). Outcomes and processes in Content and Language Integrated Learning (CLIL): Current research from Europe. In W. Delanoy & L. Volkmann (Eds.), *Future Perspectives for English Language Teaching* (pp. 139–157). Heidelberg: Carl Winter.

Dalton-Puffer, C. (2011). Content-and-Language Integrated Learning: From practice to principles? *Annual Review of Applied Linguistics, 31,* 182–204.

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

Fernández-Sanjurjo, J., Fernández-Costales, A., & Arias, 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.

Halbach, A. (2008). Bilingual methodology in Primary schools. *Revista de Educación, 346,* 455–466.

Jäppinen, A. K. (2005). Thinking and content learning of Mathematics and Science as cognitional development in Content and Language Integrated Learning (CLIL): Teaching through a foreign language in Finland. *Language and Education, 19*(2), 147–168.

Kerstin, L. (2013). CLIL in Spain—Why does it not work? A metaperspective on CLIL across contexts in Europe. *International Journal of Bilingual Education and Bilingualism, 16*(3), 301–320.

Lasagabaster, D., & López, R. (2015). The impact of type of approach (CLIL versus EFL) and methodology (book-based versus project work) on motivation. *Porta Linguarum, 23,* 41–57.

Lasagabaster, D., & Ruiz de Zarobe, Y. (Eds.). (2010). *CLIL in Spain. Implementation, results and teacher training.* Newcastle upon Tyne: Cambridge Scholars Publishing.

Lasagabaster, D., & Sierra, J. M. (2010). Immersion and CLIL in English: More differences than similarities. *ELT Journal, 64,* 376–395.

Madrid, D. (2011). Monolingual and bilingual students' competence in social sciences. In D. Madrid & S. Hughes (Eds.), *Studies in bilingual education* (pp. 195–222). Frankfurt-am-Main: Lang.

Madrid, D., & Hughes, S. (2011). Introduction to bilingual education. In D. Madrid & S. Hughes (Eds.), *Studies in bilingual education* (pp. 17–50). Frankfurt-am-Main: Peter Lang.

Marsh, D., & Frigols, M. J. (2013). Content and Language Integrated Learning. In C. Chapelle (Ed.), *Blackwell encyclopedia of applied linguistics* (pp. 911–920). Oxford: Blackwell.

Marsh, H., Hau, K.-T., & Kong, C.-K. (2000). Late immersion and language of instruction in Hong Kong High Schools: Achievement growth in language and nonlanguage subjects. *Harvard Educational Review, 70,* 302–347.

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.* Advance online publication. https://doi.org/10.1080/13670050.2020.1715916.

Mattheoudakis, M., Alexiou, T., & Laskaridou, C. (2014). To CLIL or not to CLIL? The case of the 3rd experimental Primary school in Evosmos. In N. Lavidas, T. Alexiou, & A. M. Sougari (Eds.), *Major trends in theoretical and applied linguistics: Selected papers from the 20th international symposium from theoretical and applied linguistics* (pp. 215–233). London: DeGruyter Versitas Publications.

Mehisto, P., Marsh, D., & Frigols, M. J. (2008). *Uncovering CLIL, Content and Language Integrated Learning in bilingual and multilingual education.* Oxford: Macmillan.

Muñoz, C. (2002). Relevance & potential of CLIL. In D. Marsh (Ed.), *The European dimension: Actions, trends and foresight potential* (pp. 33–36). Strasbourg, European Commission.

Murray, D. R. (2010). Irish-medium language immersion programs' effect on mathematics education. *Journal of Mathematics Education at Teachers College, 1,* 28–32.

Nikula, T. (2016). CLIL: A European approach to bilingual education. In N. Van Deusen-Scholl & S. May (Eds.), *Second and foreign language education* (pp. 1–14). New York: Springer.

Nikula, T., & K. Mard-Miettinen. (2014). Language learning in immersion and CLIL classrooms. In J.-O. Östman & J. Verschueren (Eds.), *Handbook of pragmatics* (pp. 1–26). John Benjamins Publishing Group.

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. L. (2012). CLIL research in Europe: Past, present, and future. *International Journal of Bilingual Education and Bilingualism, 15,* 315–341. https://doi.org/10.1080/13670050.2011.630064.

Pérez Cañado, M. L. (2016). From the CLIL craze to the CLIL conundrum: Addressing the current CLIL controversy. *Bellaterra Journal of Teaching Language and Literature, 9*(1), 9–31. https://doi.org/10.5565/rev/jtl3.667.

Pérez Cañado, M. L. (2017). Stopping the "pendulum effect" in CLIL research: Finding the balance between Pollyanna and Scrooge. *Applied Linguistics Review, 8*(1), 79–99. https://doi.org/10.1515/applirev-2016-2001.

Pérez Cañado, M. L. (2018). The effects of CLIL on L1 and content learning: Updated empirical evidence from monolingual contexts. *Learning and Instruction, 57,* 18–33. https://doi.org/10.1016/j.learninstruc.2017.12.002.

Piesche, N., Jonkmann, K., Fiege, C., & KeBler, J.-U. (2016). CLIL for all? A randomised controlled field experiment with sixthgrade students on the effects of content and language integrated science learning. *Learning and Instruction, 44,* 108–116.

Ruiz de Zarobe, Y. (2013). CLIL implementation: From policy-makers to individual initiatives. *International Journal of Bilingual Education and Bilingualism, 16*(3), 231–243.

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

Sotoca, E. (2014). La repercusión del bilingüismo en el rendimiento académico en alumnos de colegios públicos de la Comunidad de Madrid. *Revista Complutense de Educación, 25*(2), 481–500.

Stehler, U. (2006). The acquisition of knowledge in bilingual learning: An empirical study on the role of language in content learning. *Vienna English Working Papers, 15*(3): 41–46.

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

Sweller, J., Ayres, P., & Kalyuga, S. (2011). *Cognitive load theory*. Berlin: Springer.

Ullmann, M. (1999). History and geography through French: CLIL in a UK Secondary school. In J. Masih (Ed.), *Learning through a foreign language: Models, methods and outcomes* (pp. 96–105). London: Centre for Information on Language Teaching and Research.

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

Wode, H. (1999). Language learning in European immersion classes. In J. Masih (Ed.), *Learning through a foreign language: Models, methods and outcomes* (pp. 16–25). London: Centre for Information on Language Teaching and Research.

Xanthou, M. (2011). The impact of CLIL on L2 vocabulary development and content knowledge. *English Teaching: Practique and Critique, 10*(4), 116–126.

Advancing CLIL as Bilingual Pedagogy and Performance in Spain: A Commentary from 'el Otro Lado del Charco'

Cristian R. Aquino Sterling

—A Rodolfo, *un tesoro*

Abstract Being mindful of his positionality as a scholar working in the United States of America, *del otro lado del charco*, the author suggests future challenges and ways forward for CLIL characteriszation, implementation, and research in light of the state-of-the-art work presented in this timely edited volume. In asserting how this work advances CLIL as bilingual pedagogy and performance, Aquino-Sterling outlines how Pérez-Cañado and colleagues, by virtue of their sound and comprehensive study, render CLIL as an evidence-based approach to bilingual education; one that, when implemented in contextually and pedagogically informed ways, has the potential to yield positive impactful results. Furthermore, this commentary suggests ways in which CLIL could continue to serve as a powerful pedagogical force for advancing diversity, equity, and inclusion within Spain's culturally and linguistically diverse classroom contexts. Drawing on the critical multilingual tradition in Spain, the author frames CLIL within its continued potential to serve as a humanising and socially engaged bilingual pedagogy for all.

At the beginning of the fall 2018 semester, I arrived in the great—*alucinante*—City of Madrid to carry out a pilot research study aimed at understanding approaches for preparing bilingual and foreign language teachers.[1] The study was part of a greater comparative research proposal seeking to identify areas of theoretical, methodological, and pedagogical convergence and/or divergence relative to how the work of bilingual and foreign language teacher education is conceptualised and conducted

[1] I am grateful to Dr. Luisa Martín Rojo—Professor at the Universidad Autónoma de Madrid (UAM) and Principal Investigator of the Multilingualism, Social Identities, Intercultural Relations, and Communication Research Group (https://www.mircouam.com/en/)—for her expert guidance and kind invitation to serve as a MIRCo visiting researcher for the Fall 2018 semester.

C. R. Aquino Sterling (✉)
San Diego State University, San Diego, CA, USA
e-mail: caquino@sdsu.edu

M. L. Pérez Cañado (eds.), *Content and Language Integrated Learning in Monolingual Settings*, Multilingual Education 38,
https://doi.org/10.1007/978-3-030-68329-0_10

in the United States (specifically in the State of California), Spain (specifically in the Community of Madrid), and Brazil (specifically in Rio de Janeiro).[2] Although I am well aware these countries—and even particular regions within them—have had unique histories in the conceptualisation, implementation, research, and development of "bilingual education," I felt—and still feel *strongly*—that obtaining comparative and evidence-based understandings on the state of the art of the field across national contexts could serve as a way to continue to establish sustained scholarly collaborations for advancing bi-multi-lingual education at regional, national, and cross-national level (Aquino-Sterling et al., 2021). Given these experiences, as I prepared to write this brief commentary, I felt compelled to review my '2018 Pre-sabbatical Research Notes'. In those pages, I came across a set of research questions that continue to spark my intellectual curiosity. As a way to contextualise the observations I have the honour to provide in these pages, and since I am still 'living the questions' (Rilke, 1934), I find it fitting to list them below:

1. What is the state of the art in the field of bilingual and foreign language teacher education in relation to research, theory, and practice in la Comunidad de Madrid and Rio de Janeiro?[3]
2. What innovations characterise the field in relation to research, theory, and pedagogical-instructional practices?
3. What opportunities and challenges pervade the field?
4. What is the vision (or are the visions) for advancing the field?
5. How do these understandings (questions 1–4) compare-contrast with U.S. realities?
6. Could a comparative analysis help us develop a healthy critical eye in regards to 'making the familiar strange' in our respective regional and national contexts?
7. What can we possibly learn from each other's experiences?

Drawing on some of the critical and essential questions characterising research in the field of bilingual/dual language education in the United States—as well as on some of the works emerging out of Spain's robust critical multilingual education studies tradition (see, for example, Martín Rojo, 2003, 2010)—I decided to add four more questions to my '2018 Pre-sabbatical Research Notes'. Although not related to bilingual and foreign language teacher education proper, I consider these questions to be fundamental for advancing our understandings of bilingual and foreign language education from diversity, equity, inclusion, and social justice standpoints. I remember

[2] I am grateful to Dr. Ralph Bannell (Departamento de Educação), Dr. Alexandre Montaury Baptista Coutinho (Departamento de Letras), Dr. Inés Miller (Departamento de Letras), Dr. Adriana Nobrega (Departamento de Letras), and Dr. Liliana Cabral Bastos (Departamento de Letras) at Pontificia Universidade Católica do Rio de Janeiro for their expert guidance and kind invitation to serve as a visiting scholar for the Spring 2019 semester.

[3] Although in Spain (as well as in Brazil) these fields are closely related, in the United States we understand bilingual (teacher) education and foreign language or world language (teacher) education to be separate academic fields of research and practice. These do converge in foundational theoretical and methodological-instructional aspects, but in the academic imaginary these are two actual and separate fields.

thinking of these questions as *non-negotiable* topics for inquiry once I had addressed questions 1–7 above:

1. How effective is (public) bilingual/foreign language education from the perspectives of (a) bilingual and/or foreign language teacher educators; (b) K-12 preservice and in-service teachers; (c) K-12 students; (d) K-12 parents/guardians, and (e) the general public?
2. How is this effectiveness perceived in relation to student diversity or 'superdiversity' (i.e. immigrant/migrant status, socio-economic status (SES), race, ethnicity, cultural background[s], home language[s]; language varieties)?
3. How do these students fair (quantitatively and qualitatively) in terms of linguistic, academic, and personal and social achievement?
4. How do these understandings (questions 1–3) compare and/or contrast with U.S. realities?

Although my aim in this commentary is not to discuss the results of the *actual*[4] pilot study I was able to conduct in la Comunidad de Madrid, it is refreshing to read and comment on this timely and innovative edited volume that, in seeking to illumine critical aspects of the current Spanish-European Content and Language Integrated Learning (CLIL) controversy, to a degree addresses questions relevant to the ones posed above in conceptually and methodologically sound, rigorous, and substantive ways.

While I recognise that I am an 'outsider' to the European-Spanish bilingual and foreign language education tradition, what I've learned through my relatively short, yet eye-opening research experiences in Madrid, allows me to read *Content and Language Integrated Learning in Monolingual Settings: New Insights from the Spanish Context* with a complementary set of informed and critical lenses. My end goal, however, is to suggest future challenges and ways forward for CLIL characterisation, implementation, and research in light of the current CLIL debate and the quantitative and qualitative evidence signalling CLIL's positive impact and effectiveness in a variety of school contexts serving diverse student populations, as well as specific areas for improvement in the scholarly quest and commitment to continue to advance CLIL as an effective bilingual pedagogy and performance.

In so doing, however, I am deeply aware of my particular academic *positionality* and the fact that these insights are offered from *el otro lado del charco,* and particularly from a place where bilingual dual language education—or the use of *two* languages in the instruction and assessment of learners (García, 2009)—is conducted within relatively distinct political, cultural, linguistic, historical, and educational landscapes that have given rise to varying programmatic approaches for educating diverse bilingual learners. It is important to note that although the CLIL European-Spanish orientation to bilingual education—preceded by North American immersion and bilingual education programmes, as well as European international schools (Pérez-Cañado, 2012)—is perhaps closest to U.S. Sheltered Content Area Instruction—*grade-level content area instruction that is provided in English*

[4]The final pilot study did not address all the preliminary questions formulated and here included.

but in ways that makes it comprehensible to ELLs while promoting their English language development—and Structured English Immersion—*self-contained grade-level classrooms with teachers trained to provide language and content instruction for English Language Learners*—(see Wright, 2019), it is clear that U.S. bilingual programme labels and categorisations are insufficient for grasping or characterising the complex nature of the CLIL European-Spanish bilingual tradition.

Coyle, Hood, and Marsh (2010, p. 1) define CLIL as: '[…] a dual-focused educational approach in which an additional language is used for the learning and teaching of both content and language' (cited in Cenoz, 2015, p. 11) and, therefore, in trying to understand the CLIL bilingual tradition in relation to U.S. realities, rather than comparing CLIL to U.S. *programmatic models*, it might be most useful to focus on the *pedagogical or instructional orientations* that serve as foci for implementing the so-called 'dual-focused educational approach' in language and content education in Spain, namely: Content and Language Integrated Learning (CLIL) and Content-Based Instruction (CBI). In addtition, it is important to note that although Cenoz (2015) has found that '[…] there are no differences between CBI and CLIL regarding their essential properties' (p. 21), within a U.S. context, CBI is usually employed within foreign language (*not bilingual/dual language*) programmes (i.e. Foreign Language in Elementary School or 'FLES') where students are *not* necessarily engaged in learning a foreign or world language through the content areas or through disciplinary knowledge and standards-based educational approaches (Tedick & Wesley, 2015). Within U.S. elementary and secondary foreign language programmes where CBI serves as the organising orientation to content and language teaching and learning, teachers are not required to scaffold language and learning (Gibbons, 2015) in order to make sure bilingual learners are provided access to the school grade-level curriculum. Even though CLIL and CBI teacher *knowledge* and *competencies* have much in common, the *performance* and *implementation* of these language and content pedagogies could and at times do serve different educational purposes. These comparisons could also be relevant to the CLIL-English as a Lingua Franca (EFL) commonalities Pérez Cañado indicates in this volume. CLIL and EFL could most certainly serve complementary educational purposes; yet it is the differences in instructional purposes that create the *real* distinction between these teaching modalities. CLIL's *raison d'être* goes beyond the teaching and learning of English *through* the disciplines. CLIL's sphere expands to greater personal, academic, and social purposes for the thousands of students who are learning through this medium. For example, it is through CLIL as a language and content pedagogy (and not through an EFL pedagogy) that students are able to develop disciplinary literacies or shared ways of reading, writing, thinking, and reasoning within academic fields (Moje, 2007; Shanahan & Shanahan, 2008):

> Each discipline has unique ways of asking questions and solving problems. Similarly, each discipline has unique expectations for the types of claims that are made and the way those claims are supported. These differences play out in the ways that texts are written and in the demands those texts place on the readers. For these reasons, we can say that each discipline has its own discourse community, a shared way of using language and constructing knowledge. (Rainer & Moje, 2012, p. 73)

1 Advancing CLIL as Bilingual Pedagogy and Performance

With the previous in mind, and alluding to the distinction between *linguistic competence* (knowledge of language) and *linguistic performance* (the way we put that knowledge to use in real situations) (see Chomsky 1965), the work documented in this edited volume reflects *cierta postura* in advancing CLIL both as a sound *bilingual pedagogy* and as *performance*. This distinction could assist scholars, educational stakeholders, and the greater public in reframing the current CLIL debate-controversy in ways that characterise CLIL as an evidence-based and conceptually and pedagogically sound approach to bilingual education in Spain (*a knowledge and a competence*) that when displayed (or *performed*) in optimal ways, yields positive and impactful results. However, when certain conditions are not met, the way CLIL as a bilingual pedagogy is performed in real situations can yield undesirable effects that if/when properly addressed will, most definitely, serve to advance the Spanish CLIL bilingual tradition to new heights. Such is the evidence we find in this opportune edited volume.

In its aim 'to provide an updated picture of where [Spain] stands in the CLIL arena' (Pérez Cañado, this volume), *Content and Language Integrated Learning in Monolingual Settings: New Insights from the Spanish Context* helps us to understand the pedagogical, educational, and social value of CLIL as seen across the 53 schools and 12 Spanish regions studied. On the other hand, courageously and ethically, the book does not shy away from identifying areas where the CLIL performance needs to be augmented. Yet, what are some ways in which, despite the CLIL innovations, advancements, and hurdles evidenced in this work, CLIL *characterisation*, *implementation*, and *research* could continue to move forward? I offer, *desde el otro lado del charco*, some very brief suggestions below.

2 Future Challenges and Ways Forward

2.1 CLIL Characterisation

In the quest to design a CLIL prototype (a fixed or demarcated form of bilingual pedagogy and performance), as suggested in the Andalusian case study included in this volume (Rascón Moreno & Casas Pedrosa), I am reminded of the so-called 'SIOP Model' (Echevarría et al., 2012), a comprehensive, coherent, and structured model for implementing Sheltered English Instruction published in the United States (https://www.cal.org/siop/lesson-plans/index.html). Given the level of variability in CLIL performance across Spanish contexts, the value of identifying and documenting effective practices that are representative or reflective of the goals sought, cannot be underestimated. This could certainly assist in making sure CLIL teachers are properly prepared, and that the methodology is faithful and effectively implemented. One must recognise, however, that this technicist, instrumentalist, and potentionally

reductionistic approach at (re)configuring and automatising a bilingual pedagogy could also serve in the deskilling of teachers and the *macdonalisation* of bilingual education (Ritzer, 2019). One must avoid the pitfalls of a behaviourist approach to bilingual teaching and learning (Crawford & Adelman Reyes, 2015) while also making sure the classroom (and the lives of children) are not negatively impacted by an 'anything goes' instructional approach. It is ultimately this delicate balance that will prove fruitful in advancing CLIL as bilingual pedagogy and performance in Spain and the rest of Europe.

Moreover, characterising CLIL as a bilingual pedagogy aimed at just developing 'standard' English competencies will not be sufficient in twenty-first-century language education. How can CLIL as a bilingual pedagogy *plus* engage the teaching and learning of language and content from a *multiliteracies* (Kalantzis et al., 2010; The New London Group, 1996) approach? How can CLIL be reconceptualised to espouse the theorising of language acquisition as multimodal, multisensory, multilateral, and, therefore, multidimensional (Canagarajah, 2007), as referenced by Parra-Velasco (2020)? In the same way, could CLIL serve an explicit agenda for linguistic democracy and social change, especially during these challenging times where the Covid-19 pandemic and social and racial challenges continue to pervade our world? I believe framing CLIL as a humanising and socially engaged bilingual pedagogy has the potential to advance CLIL beyond the content area and linguistic needs of the critical world citizens the world needs (Aquino-Sterling, 2020). If bilingual education is to serve a greater purpose in our world societies, it must begin to reenvision and broaden its curriculum and reductionist focus on language. Within a U.S. reality, for example, bilingual education must begin to pay systemnatic, sustained, and substantive attention to issues of racial justice and anti-Black racism, global warming, social violence and extremism, and other pressing issues (Aquino-Sterling, C., Gort, M., & Flores, B.B., Forthcoming). We can no longer afford to reducing the field to issues of "translanguaging," a topic that has dominated the conversation for the past decades.

2.2 CLIL Implementation

Reimagining the future of CLIL implementation requires the design and use of valid assessment and evaluation instruments that measure both student language and content attainment (Madrid Fernández, Bueno González & Ráez Padilla, this volume). However, how can the role of formative assessments be expanded to inform both micro (classroom) and macro (local, regional, national) processes of CLIL implementation? Even if/when a CLIL prototype is developed, the performance of CLIL, or the way it is implemented within particular contexts given the knowledge and competencies of CLIL teachers and other factors, cannot and will not always be uniform across contexts (just as the local implementation of macro-level educational policies also varies across contexts). How can the development of valid assessment instruments be employed in conjunction with formative assessments to obtain a more

succinct and realistic picture of student linguistic and content achievement and performance? Could the design and implementation of valid formative language and content assessments be explicitly addressed in the CLIL teacher education curriculum?

2.3 CLIL Research

This volume attests to the fact that the present and future of CLIL research in Spain look bright. In my understanding, this is the very first study characterised by such comprehensibility and rigour in CLIL effectiveness research in Spain. I would like to encourage the field, *desde este lado del charco*, to continue to address questions on the effectiveness of CLIL—such as the ones that guided this mixed-method study—while also placing emphasis on identifying how CLIL meets standards of diversity, equity, inclusion, and social justice for low SES and immigrant student populations who attend CLIL programmes in Spain. These complementary questions have the potential to establish CLIL as a distinctly democratic and culturally and linguistically relevant and sustaining bilingual pedagogy (Paris & Alim, 2017) aimed at continuing to advance human development and freedoms for all.

References

Aquino-Sterling, C. (2020). *Humanizing bilingual education: The value of lifelong learning pedagogies.* Virtual Presentation at the São Paulo Open Centre (June 27), São Paulo, Brasil.

Aquino-Sterling, C., Faltis, C., Feng, A., Hamel, R.E., Pereira Siqueira, I.C., Truscott de Mejía, A.M., Pérez Cañado, M.L., Machado Casas, M. (April 2021). Bilingual Education Around the World: Cross-National Dialogues on the State of the Art. International Round Table, Diálogos Virtuales, College of Education, San Diego State University: https://education.sdsu.edu/about/dei-ia/diversity/dialogos.

Aquino-Sterling, C., Gort, M., Flores, B.B. (forthcoming). Innovative curricular and pedagogical designs in bilingual teacher education: Bridging the distance with P-12 contexts. Charlotte, NC: Information Age Publishers.

Canagarajah, S. (2007). Lingua Franca English, multilingual communities, and language acquisition. *The Modern Language Journal, 91*(1), 923–939.

Cenoz, J. (2015). Content-based instruction and content and language integrated learning: The same or different? *Language, Culture and Curriculum, 28*(1), 8–24.

Chomsky, N. (1965). *Aspects on the theory of syntax*. Cambridge, MA: The MIT Press.

Crawford, J., & Adelman Reyes, S. (2015). *The trouble with SIOP®: How a behaviorist framework, flawed research, and clever marketing have come to define—and diminish—Sheltered Instruction.* Portland, OR: Institute for Language & Education Policy.

Echevarría, J., Vogt, M., & Short, D. (2012). *Making content comprehensible for English learners: The SIOP Model* (4th ed.). New York, NY: Pearson.

García, O. (2009). *Bilingual education in the 21st century: A global perspective.* Malden and Oxford: Wiley-Blackwell.

Gibbons, P. (2015). *Scaffolding language scaffolding learning* (2nd ed.). Portsmouth, NH: Heinemann.

Kalantzis, M., Cope, B., & Cooland, A. (2010). A multiliteracies perspective on the new literacies. In E. A. Baker (Ed.), *The New literacies: Multiple perspectives on research and practice* (pp. 61–87). New York, NY: The Guilford Press.

Martín Rojo, L. (Ed.). (2003). *¿Asimilar o integrar? dilemas ante el multilingüismo en las aulas.* Madrid: Ministerio de Educación Cultura y Deporte.

Martín Rojo, L. (2010). *Constructing inequality in multilingual classrooms.* Berlin, Germany: De Gruyter Mouton.

Moje, E. B. (2007). Developing socially just subject-matter instruction: A review of the literature on disciplinary literacy teaching. *Review of Research in Education, 31*(1), 1–44.

Paris, D., & Alim, H. (2017). *Culturally sustaining pedagogies: Teaching and learning for justice in a changing world*. New York, NY: Teachers College Press.

Parra-Velasco, M. L. (2020). *La enseñanza del español en México: Caleidoscopio de retos y posibilidades para los profesores del siglo XXI*. Virtual Lecture at Universidad Autónoma de Baja California.

Pérez-Cañado, M. L. (2012). CLIL research in Europe: Past, present, and future. *International Journal of Bilingual Education and Bilingualism, 15,* 315–341.

Rainer, E., & Moje, E. B. (2012). Building insider knowledge: Teaching students to read, write, and think within ELA and across the disciplines. *English Education, 45*(1), 71–90.

Rilke, R. M. (1934). *Letters to a young poet*. London: Norton & Company.

Ritzer, G. (2019). *The Macdonalization of society*. Thousand Oaks, CA: SAGE .

Shanahan, T., & Shanahan, C. (2008). Teaching disciplinary literacy to adolescents: Rethinking content area literacy. *Harvard Education Review, 78,* 40–59.

Tedick, D., & Wesley, P. (2015). A review of research on content-based foreign/second language education in US K-12 contexts. *Language, Culture and Curriculum, 28,* 25–40.

The New London Group. (1996). A pedagogy of multiliteracies: Designing social futures. *Harvard Educational Review, 66*(1), 60–92.

Wright, W. (2019). *Foundations for teaching English language learners* (3rd ed.). Philadelphia: Caslon.