

<https://doi.org/10.17323/jle.2024.18150>

Analysis of CLIL-related research in school settings: A systematic review

Belen Poveda-Garcia-Noblejas , Svetlana Antropova 

Villanueva University, Spain

ABSTRACT

Background: Content Language Integrated Learning (CLIL) is an emerging approach in the global educational landscape, and as such, there is a lack of a systematic review of this field.

Purpose: To explore CLIL-related scientific publications in school settings around the world.

Method: A systematic review was performed following the PRISMA guidelines in WoS and Scopus databases. A total of 142 articles published in the period 2018-2022 were analysed according to three types of variables: extrinsic to the scientific process, methodological, and content based. The results of the methodological and content-based variables were contrasted with the portfolio of CLIL evaluation measures and analysed through the lens of the 4Cs framework.

Results: The findings revealed that CLIL studies were performed in a wide range of countries across continents. It was the secondary school which drew most scientific interest. Apropos of the methodological variable, there was a balance between qualitative and quantitative studies, and a questionnaire as a tool was favoured by the researchers. The major scientific interest lay in the communication principle, while cognition was understudied.

Conclusion: There was a growing scientific interest in CLIL. Although the major interest laid in linguistic gains, other fields of research transpired. The conclusions provide further agenda for CLIL research.

KEYWORDS

Education, schools, second language instruction, Content Language Integrated Learning (CLIL), systematic review

INTRODUCTION

For the increasingly globalized world in the 21st century, the active pursuit of innovative methodologies, which can prepare future generations to integrate in the global community of speakers from diverse linguistic and cultural backgrounds, is on all educational agendas. The dual-focused educational approach, both content- and L2-driven, which is being implemented in practically all educational stages in the European Union is CLIL. However, the question whether CLIL is unique for the European context, or whether it has transpired European borders remains open.

Due to the recent implementation of CLIL, and although the body of research in this field is growing, it is still piecemeal and not “coherent as a package” (Coyle et al., 2010, p. 135). As a result, there has

been little systematic review research so far. The current systematic review studies have approached the analysis of CLIL from diverse angles: CLIL’s implementation in the European context (Cimermanová, 2021; Goris et al., 2019; Palacios-Hidalgo et al., 2021), foreign language learning in Physical Education (Gil-López et al., 2021), curriculum evaluation of CLIL on a global scale (Li et al., 2020), and the analysis of content and language outcomes in CLIL, CBI, EMI (Graham et al., 2018). The latter concluded that while CLIL was constantly present on the research agendas of certain countries, especially in the EU, other countries remained understudied, thus further systematic reviews on the implementation of CLIL worldwide are required.

With the purpose of filling in the gap in the systematic review analysis, this study aims to provide an analysis of CLIL-re-

Citation: Poveda-Garcia-Noblejas B., & Antropova S. (2024). Analysis of CLIL-related research in school settings: A systematic review. *Journal of Language and Education*, 10(1), 146-159. <https://doi.org/10.17323/jle.2024.18150>

Correspondence:
Svetlana Antropova,
santropova@villanueva.edu

Received: October 20, 2023

Accepted: March 15, 2024

Published: March 30, 2024



lated scientific research in school settings, both within and outside the EU borders, according to the following specific objectives:

- (1) to analyse CLIL-related research in school settings according to countries, year of publication, and educational stages;
- (2) to delve into methodological variables: type of research, data collection tool, population group, area of research;
- (3) to inquire which of the 4Cs receives more/ less scientific interest;
- (4) to analyse recent areas of scientific research on CLIL;
- (5) to contrast the results with the portfolio of evaluation measures.

LITERATURE REVIEW

The Essence of CLIL

In response to the demands of the modern age, specifically to enhance foreign language teaching to promote bilingualism/multilingualism in a European context¹, a group of experts in different fields of education launched CLIL. This novel approach appealed to educators as a wide range of pilot studies on the adaptation of CLIL sprouted all over Europe, and CLIL was recommended to teach subjects in L2 in pre-primary, primary, general secondary, secondary vocational and further education (Marsh, 2002).

CLIL is defined as “a dual-focused educational approach in which an additional language is used for learning and teaching of both content and language” (Coyle et al., 2010, p. 1) inasmuch as the integration of content learning and language learning becomes the essence of CLIL (Mehisto et al., 2008). However, there have been critical voices which have pointed to the ambiguity of this definition and called for its clarification (Cenoz et al., 2014; Linares & Morton, 2017). Due to the novelty of CLIL, there is no blueprint. Nevertheless, rather than being a drawback, this lack of a specific model may be considered an advantage, since it allows educators and practitioners to contribute to the shaping of the method with good practices. As such CLIL is historically and pedagogically unique and can easily fit into different national curricula due to its flexibility (Merino & Lasagabaster, 2018). To this end, under the umbrella of CLIL, different pedagogies and models emerged responding to the socio-cultural context of the countries (Van Mensel et al., 2020). Taking stock of these new models of CLIL implementations, the Eurydice Report² stated that although some of the pedagogies had

promising results, there was no consensus reached on the theoretical principles of CLIL, which inevitably led to either more language-focus teaching, where content was being used as a mere vehicle for language development, or to content-focus instruction, where less attention was paid to interaction in L2.

To supply a rigorous theoretical basis for this methodology, Coyle (2007) provided a conceptualization of CLIL, placing the focus on an innovative 4Cs framework: Content, Communication, Cognition, and Culture. Drawing on a wide range of theories from different fields of knowledge, Coyle (2007, pp. 550–552) tackled the 4Cs as follows:

Content is viewed as a construction of knowledge of the subject based more on high-order thinking skills rather than pure memorization of the subject-matter. Therefore, thinking processes must be reflected upon for further linguistic demands as there is an integration of linguistic content with content knowledge. Communication encompasses contextualized use of the target language, which is defined not only by the linguistic needs of the subject-matter but also by social interaction. The latter becomes the way of acquisition of knowledge and skills fundamental for learning. The target language is seen as a vibrant construct which allows the learner to access other fields of knowledge and to be able to interact with others.

Cognition is an integrative component of this framework as the progression from low-order thinking skills (hereinafter LOTs) to high-order thinking skills (hereinafter HOTs) is viewed as a requirement to progress both in content and target language. Since the development of creative and critical thinking is one of the educational demands of the 21st century, problem-solving and decision-making were recommended to form part of CLIL classes (Cimermanová, 2021).

Culture is deeply embedded in the language and determines the way we interpret the world, therefore cultural understanding and awareness of the conventions form part of this method. Hence, Byram’s concept of intercultural competence, “the ability to communicate and operate effectively with people from another culture” (Byram, 1997, p. 5), became one of the cornerstones of CLIL. Furthermore, Byram (2012) included citizenship education in this competence, which was incorporated as one of the goals of CLIL (Coyle et al., 2010), and Mehisto et al. (2008) added Community to the principle of Culture.

When providing the examples of CLIL classes in primary and secondary education, Mehisto et al. (2008) contemplated the inclusion the 4Cs as the guiding principles which can contribute to successful outcomes.

¹ Eurydice Report. (2006). Content and Language Integrated Learning (CLIL) at school in Europe (Y. European Commission, Directorate-General for Education Sport and Culture, Ed.). Publications Office. <https://op.europa.eu/en/publication-detail/-/publication/756ebdaa-f694-44e4-8409-21eef02c9b9b>

² *ibid*

Portfolio of Evaluation Measures

To gain more insight into CLIL implementation and outcomes, Coyle et al. (2010) proposed a series of evaluation criteria, contemplated in the portfolio of evaluation measures. The four areas of research are the following:

Performance evidence englobes the empirical research in which its major aim is to assess the learners' outcomes in CLIL subjects, as well as to compare the results with established expectations from the national curriculum. Both quantitative (statistical data) and qualitative studies (e.g., portfolios) are included in this evaluation measure. For the analysis of progression in subject-matter, a contextual comparison of outcomes in L1 and L2 was recommended.

Affective evidence is the field of research which aims at gathering and evaluating learners' and teachers' testimonies as far as motivation, L2 anxiety, self-esteem, etc. The instruments for qualitative research include open-ended questionnaires, focus groups, and individual interviews. For a more in-depth analysis of CLIL outcomes, a joint evaluation of performance and affective evidence through "a fuller cross-referenced portfolio using the range of students across the ability range" was proposed (Coyle et al., 2010, p. 137)

Both process evidence and materials and task evidence deal with the actual in-class procedures. Whilst process evidence evaluates the learners' verbal performance (individual, pair, groups), task evidence aims at the analysis of the tasks and the materials used. These evaluation elements can be most complex for assessment from a "logistical standpoint", as certain precision is required with data collection procedures and tools of analysis (Coyle et al., 2010, p. 137).

Table 1

Inclusion and Exclusion Criteria

Inclusion Criteria	Exclusion Criteria
R1. Type of publication: journal article	Books, book chapters, theses, conference proceedings
R2. Field of knowledge: Education and Social Science	Other fields of knowledge
R3. Language: English, Spanish	Other languages
R4. Educational stage: pre-primary (3-6), primary (7-11), secondary (12-18)	Other educational stages: undergraduate, postgraduate, etc. Duplicated
R5. Empirical studies	Theoretical studies, systematic reviews
R6. CLIL	Other methodologies

Note. R=Reason in Flow Diagram for automation tools

METHOD

Protocol and Registration

This systematic review followed the guidelines of PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses, 2020) statement, to ensure that the results have adequate quality criteria and to reduce biases during the selection process.

The stages carried out to perform this systematic review were the following: (1) formulation of the research problem, which in this case was to analyse the trajectory and current situation of the most relevant research on CLIL. To this end, (2) a selection of articles on CLIL in indexed publications with double peer review in Scopus and WoS databases were performed. (3) Inclusion and exclusion criteria, which are further explained in this article, were applied. And, finally, (4) the analysis was carried out considering extrinsic and intrinsic variables.

Procedures

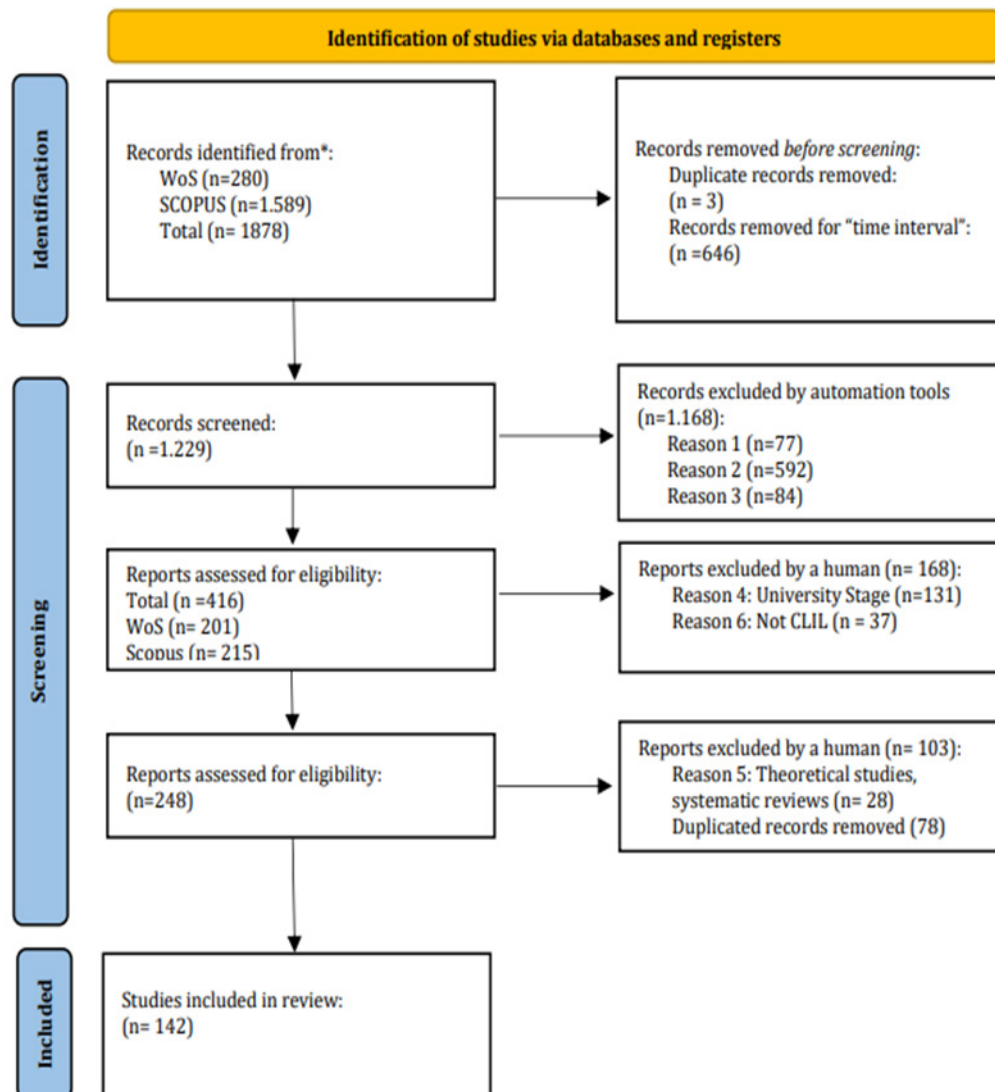
The search was carried out in the databases of Scopus and Web of Science (WoS), during the months of March – April 2023. To ensure that all the articles fall into the domain of the present research the following terms were introduced in the English language, following ERIC indicators: TI= (CLIL AND/OR Content Language Integrated Learning).

Search Strategies

To narrow the search of the articles published, and, therefore, to certify the relevance and the current state of the art on the topic of the present study, the period from 2018-2022 was selected. Applying this first criterion, a total of 1878

Figure 1

Flow Diagram for Systematic Reviews Adapted from PRISMA 2020



(WoS - 280 articles; Scopus - 1.589 articles) were found. From then on, given the size of the sample, inclusion and exclusion criteria were applied (Table 1).

Selecting Studies Procedure

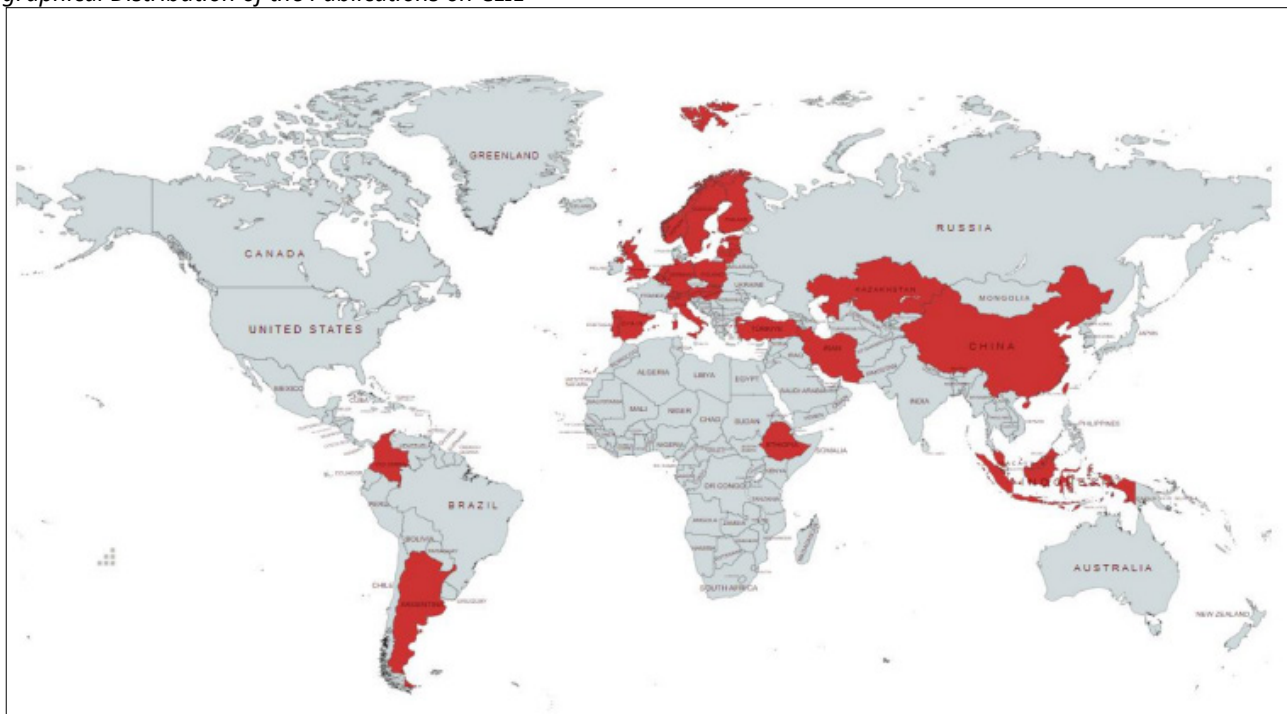
Considering the eligibility criteria, an evaluation of the selected articles was performed independently. In case of disagreement, judgment was sought from a third person. Both the JBI Critical Appraisal Tools (Lockwood & Tricco, 2020) and the Checklist for Qualitative Research³ were used specifically. Therefore, the selected articles were found to satisfactorily meet the inclusion criteria (Figure 1).

Based on the number of final records, which met the criteria of inclusion and exclusion, a database was elaborated (Antropova & Poveda, 2023). To proceed with the codification and the analysis of the publications, a series of variables were established: (1) extrinsic to the scientific process (country, year of publication, educational stage); (2) methodological (type of research, data collection tool, area of research⁴, population group); (3) content-based (4 Cs, analysis of abstracts).

Finally, a detailed qualitative analysis (Stimson, 2014) of the abstracts of each of the articles was conducted to explore content-based variables. The aim was to find "units of registration" (Díaz Herrera, 2018, p. 127) common to all the pub-

³ Checklist for Qualitative Research. <https://bit.ly/2YFoTWH>

⁴ Area of research: performance evidence, process evidence, affective evidence, material and task evidence (see Portfolio of Evaluation Measures).

Figure 2*Geographical Distribution of the Publications on CLIL*

lications, and thus, to establish the central categories and subcategories.

RESULTS

Analysis of Extrinsic Variables

Considering the variables extrinsic to the scientific process, the country of research, year of publication, and educational stage were analysed.

As for the geographical location, a great diversity was observed in terms of the origin of the study sample (Figure 2). Spain ($n=70$) stood out as the country with the highest concentration of research, followed by other countries: Germany ($n=10$); Finland ($n=9$), the Netherlands ($n=7$), Austria ($n=7$). Belgium, Portugal, Italy, and the UK had 5 publications. Authors from Colombia and Kazakhstan published 4 articles, followed by scientific research in Switzerland ($n=3$) and in Argentina, China, Hungary, Indonesia, Iran, Slovakia ($n=2$). In the rest of the countries (Estonia, Ethiopia, Latvia, Lithuania, Malaysia, Norway, Poland, Sweden, Taiwan, Turkey) there was 1 publication in the time interval analysed.

Regarding the evolution over time (2018 - 2022), starting with a small number in 2018 ($n=11$), the number of publications displayed a steady growth until 2019 ($n=32$). In 2020 this type of publication decreased ($n=24$), albeit in 2021 the publications resumed their gradual increase, reaching the highest number ($n=38$) and in 2022 ($n=36$).

Concerning the educational stage of the empirical studies, secondary schools accumulated the highest number of publications ($n=82$), followed by the studies in primary ($n=33$), several educational stages simultaneously ($n=26$), and pre-primary ($n=1$).

Analysis of Methodological Variables

Regarding the analysis of the methodological variables with respect to the type of research, both qualitative and quantitative studies accumulated 36,6%, mixed method studies - 16%, and feasibility studies - 3,5%.

In terms of data collection strategies, the following tools were used by the researchers (Table 2).

Regarding the analysis according to area of research (Coyle et al., 2010), it was performance evidence (PE) which accrued the highest number of publications ($n=80$), followed at a considerable distance by material and task evidence (MTE) ($n=19$) and affective evidence (AE) ($n=15$). Process evidence (PrE) was the area with the least scientific research ($n=13$). Some publication explored various areas of research concomitantly in their studies: AE and MTE ($n=5$), AE/PE ($n=13$), and PE/MTE/PrE ($n=1$).

Concerning the population groups, students were the major target subjects ($n=74$); followed by the studies focused on teachers ($n=42$). Some researchers enriched their analyses of data collection from mixed population groups of parents, teachers, and students ($n=19$). The rest of the studies

Table 2*Data Collection Tools according to the Type of Research*

Data Collection Tool	Quantitative	Qualitative	Mixed Method	Feasibility	Total
Questionnaire	34		30	1	65
Interview		31	19		50
Observation/audio-video recordings	2	23	6		31
Focus group		9	3		12
Test	12		2		14
Standardized test	13		1		14
Verbal interactions	1	5		1	7
Reports/field notes/essays/written reflections/portfolios	6	4	1		11
Didactic resources analysis/tool proposal and/or validation	2	2		2	6
Task	4	9	4		17
Accelerometer	1		1		2
Didactic proposal				3	3
Pedagogical intervention	4	1		1	6

focused on obtaining information from other types of informants: CLIL programme coordinators (n=2), CLIL experts (n=2), directors (n=2), and inspectors (n=1).

Analysis of Content-Based Variables

For a more in-depth content-based analysis, the following variables were explored: 4Cs framework, abstract.

Delving into the 4Cs framework, the studies dealing with the Communication principle (n=72) were the most numerous, followed by Content (n=31), Culture/Community (n=15) and Cognition (n= 12). Exploring studies which focused on Communication, there was scientific research which only dealt with this principle (n=46); however, Communication was also addressed in concert with other variables including Content (n=12), Community/Culture (n=4), Cognition (n=2) highlighting the combination of Communication and Content, with respect to the others.

Furthermore, abstract analysis was performed. As a result, 13 central categories together with their corresponding subcategories (n=70) emerged (Table 3). This categorisation streamlined content analysis with the focus on the issues currently being researched.

The categories with the highest percentage of appearance were Communication principle (26.83%), Focus on teacher (17.49%), and Focus on student II (11.72%).

In Communication, the subcategory with the highest percentage was L2 development/L2 proficiency/L2 competence (20,74%), followed by Vocabulary/lexical richness (14,81%), written production/literacy skills (13,33%), and Oral production/speaking (11,85%). The rest of the subcategories obtained less than 10%.

In the analysis of the category Focus on teacher, the subcategories that accrued the highest percentages were Pedagogies/didactic competence/teaching experience (26,13%), Professional development/training (25%), Identity/personal beliefs/perspectives (20,45%), and Coordination/organization (11,36%). The rest of the subcategories scored less than 10%.

Apropos of the category Focus on students II, it was Motivation that scored the highest (47,45 %) followed by Satisfaction/enjoyment (16,84%), and Effort/attitude/willingness to communicate (13,55%). Other subcategories obtained less than 10%.

DISCUSSION

This systematic review aimed to analyse CLIL-related scientific research published in English and Spanish in school settings worldwide in 2018-2022. Having explored 142 scientific publications, the findings are the following.

Table 3*Emerging Categories and Subcategories in the Qualitative Analysis with Cumulative Percentages*

Categories	Subcategories	N	Percentage	
Socio-educational context	International/national educational policies	Language assistants	16	3,20%
	Exchange programmes			
L1 and L2, L3	Translanguaging	L2-L1 influence	27	5,36%
	Bilingualism/multilingualism	L1 as language of instruction		
Curriculum	Didactic application/innovation		11	2,18%
	Evaluation			
	Non-specified			
Communication	L2 development/L2 proficiency/L2 competence	Pronunciation/phonological awareness	135	26,83%
	Reception/listening	Vocabulary/lexical richness		
	Oral production/speaking	Grammar/morphology		
	Reception/reading	Fluency		
	Written production/literacy skills	Non-verbal language		
	Students' interaction/conversation analysis	Teacher's language (verbal/non-verbal)		
	Teacher-Students interaction	Teacher's corrections		
Content	Subjects		44	8,78%
	Other content			
Community/ Culture	L2 culture	Students' collaboration/cooperation	16	3,18%
	Intercultural awareness/competence	Class community		
	International orientation			
Cognition	LOTs	Metacognition/learning strategies	21	4,17%
	HOTs	Non-specified		
Teaching methodologies/ strategies	CBI; TBLT; PLTL; CLL; IBSE; GTM; CLT; PBL; Gamification		23	4,57%
	Scaffolding; questioning; games; peer-tutoring			
	Other			
Design I: grouping, time organization	Grouping patterns		5	0,99%
	Study/task time organization			
Design II: instructional materials	Graphic organizers	Textbook/books	33	6,58%
	Digital resources	School facilities		
	Audio-visual aids	Augmented reality		
	Non-specified			
Focus on Teacher	Professional development/training	Collaboration	88	17,49%
	Coordination; organization	Ideology		
	Pedagogies/didactic competence/teaching experience	Affective factors		
	Identity/personal beliefs/perspectives			

Categories	Subcategories	N	Percentage
Focus on student I: socio-demographic and personal characteristics	Socio-economic level	23	4,57%
	Age		
	Gender		
	Ideology		
	Physical activity		
Focus on student II: affective factors	Motivation	59	11,72%
	L2 anxiety		
	Satisfaction/enjoyment		

Note. CBI=Content Based Instruction; TBL=Task-Based Language Teaching; PTL=Peer-Led Team-Learning; CLL=Cooperative Language Learning; IBSE; GTM=Grammar Translation Method; CLT=Communicative Language Teaching; PBL=Problem-Based Learning.

Gradual Growth of CLIL-related Research in School Settings over the World

As for the analysis according to countries and year of publication, the results show that although most scientific research has been performed in the EU, CLIL has transcended European borders as other countries (e.g., Iran, Taiwan, Indonesia) have gradually introduced this method in their classrooms. On a global scale, there is undeniable interest in CLIL, as the number of publications has been on a continuous rise, albeit there are countries (e.g., Brazil, Russia, India) where it is unclear if there has been any scientific research on the matter at all, perhaps due to the language bias. Since the languages of this study were English and Spanish, scientific research in other languages was not considered.

As CLIL was initially designed for the EU, all the countries which have accumulated the highest number of articles on this topic are in the EU: Spain, Germany, Finland, the Netherlands, and Austria. Nevertheless, other European membership countries have not been prolific CLIL research wise. Therefore, the question arises: are there any factors which have influenced a CLIL implementation in school settings in the countries with the greatest quantity of articles? An example of Spain, which has accrued the highest number of scientific publications (Cimermanová, 2021; Goris et al., 2019) may serve as a model to delve into this question. Several factors have affected the rapid acceptance of CLIL in Spain: (1) bilingualism with minority languages; (2) top-down implementation; (3) low-proficiency in L2 skills due to the lack or inconsistency of SLT instruction (Goris et al., 2019). In the analysis of the countries with more prominent research in CLIL, it seems that traditional multilingual idiosyncrasy has been a mighty factor to promote CLIL in the Netherlands, Finland, Germany, and Austria; perhaps, due

to the urgency to introduce teaching of content in other languages. However, not all multilingual/bilingual countries have demonstrated scientific interest in CLIL (e.g., Switzerland, Malta).

Furthermore, regarding CLIL implementation, both top-down implementation and grassroots initiatives have proved efficient. Whilst educational laws streamlined CLIL in Finland (Roiha, 2019), grassroots initiatives gave an impulse to CLIL at schools in the Netherlands (Mearns & Graaff, 2018), Germany (Siepmann et al., 2021), and Austria (Bauer-Marschallinger et al., 2021). For instance, in the Dutch context parents and teachers promoted this method as “a pedagogical principle” in bilingual schools (Mearns & Graaff, 2018, p. 125). The importance of grassroots initiatives can be viewed as evidence of high expectation of CLIL and its easy acceptance on the part of stakeholders (Hüttner et al., 2013).

Regarding low proficiency L2 skills, this is true only for Spain. In 2018, the starting point of the current study, the rest of the abovementioned countries fell into the very high proficiency ranking⁵: the Netherlands - 2, Finland - 8, Austria - 12, and Germany - 10. Therefore, on a global scale, this factor has not been by far the most influential. As no clear pattern emerges to explain the scientific interest in specific countries, further research is required to gain an insight into the factors which can boost CLIL implementation. In the analysis of the articles from the countries with the highest percentage of research, a new ecological factor has emerged in line with the one proposed by Dalton-Puffer et al. (2022): internationalisation. To comply with the educational demands, more specifically plurilingual competence in the era of globalization, the knowledge of English as a lingua franca is given priority. Therefore, CLIL is viewed by many as “a form of extended language policy” (Hüttner et al., 2013) and a tool

⁵ EF English First. (2018). The world's largest ranking of countries and regions by English skills. <https://www.ef.com/assetscdn/WIBIwq6Rd-jcD9bc8RMd/cefcom-epi-site/reports/2018/ef-epi-2018-english.pdf>

to promote global citizenship and future job opportunities, hence its importance.

Concerning educational stage, secondary school has attracted more scientific interest in compliance with other systematic reviews (Cimermanová, 2021; Gil-López et al., 2021; Goris et al., 2019; Li et al., 2020), whilst in the pre-primary stage the scientific research has been scarce. Though the pre-primary stage was included in CLIL programmes (Marsh, 2002), due to low L2 proficiency and a small number of subjects in this educational stage, other SLT methodologies are applied. While educational laws promote CLIL in pre-primary stage in Austria and Belgium (Bauer-Marschallinger et al., 2021; Van Mensel et al., 2020), the major bulk of research has explored the outcomes of CLIL in primary and secondary education in these countries. There is still a debate in the European educational context whether an early start of CLIL is beneficial for L2 learning (Goris et al., 2019) or whether L2 training should precede CLIL programmes. In the longitudinal study (Pfenninger, 2020) on the comparison of L2 development in a CLIL group to its mainstream counterpart, the results revealed that there were no meaningful differences in L2 acquisition between early starters of CLIL and slightly later CLIL beginners. Some countries (e.g., the Netherlands, Germany) initiate CLIL only in secondary education with self-selection criteria based on L2 proficiency to be applied to the candidates (Mearns & Graaff, 2018; Siepmann et al., 2021). Thus, verbal intelligence and academic ability have become core elements to join a CLIL programme.

Methodological Approaches to Explore CLIL Efficiency

Regarding methodological variables, the type of research, research instrument, and population groups of the studies were explored. The results of both methodological and content-based variables are contrasted with the portfolio of evaluation measures (Coyle et al., 2010).

There is a similar proportion between qualitative and quantitative studies in comparison to Li et al.'s (2020) systematic review, which highlighted the predominance of quantitative-driven studies. According to the results of this research, the gap between quantitative and qualitative designs has disappeared. Mixed method studies have not been abundant albeit recommended by Coyle et al. (2010). Hence, the importance of a mixed method design has been seemingly overlooked by the researchers. Molina-Azorin and Fetters (2019) highlighted the special value of mixed-method research, since it not only engages stakeholders in the creation of knowledge, but it also helps to evaluate and disseminate the impacts of the study. Thus, a mixed method design can enable us to obtain new empirical insights as well as to

provide compelling evidence that can contribute to CLIL's improvement. For instance, in a mixed method study on the impact of CLIL on motivation in the UK (Bower, 2019a), contrastive results to the questionnaire led to a series of questions for interviews and focus groups to explore the reasons for those differences. Moreover, different stakeholders were involved which helped to interpret the findings and to propose a solid theoretical framework for motivation in CLIL.

In terms of data collection instruments, questionnaires have been a predominant tool (Goris et al., 2019; Li et al., 2020). Although Coyle et al. (2010) proposed the use of questionnaires to delve exclusively into affective factors of CLIL participants, in the scientific literature analysed this tool has been used to explore both PE and MTE. Whilst questionnaires have many benefits towards carrying out research, such as possibility of administration to a wide sample of participants and little complexity in analysis and comparison of the results (Hernandez-Sampieri & Mendoza, 2018, p. 263), their disadvantages must not be neglected. Among them are assessment of attitudes, but not behaviours; they do not provide information about the individual, except for the variables measured, and the use of language can be a source of bias and can influence responses. Therefore, the peculiarities of CLIL in different socio-cultural contexts, such as different ages of implementation, various models for CLIL, compulsory/voluntary status of CLIL, not to mention personal experiences, could give rise to a controversy of stakeholders' perspectives which should not be equated on a global level.

Despite Coyle et al.'s (2010) recommendation to use a wide array of data collection instruments, such as portfolios, standardised tests, etc., these instruments have been used with less frequency. As a result, there is a disproportion between qualitative instruments and quantitative techniques. In the literature revised there have been a plethora of voices which have called for more solid evidence to prove CLIL's success in comparison to national and universal standards of education (Coyle et al., 2010; Goris et al., 2019). Even though PE has been by far the most researched area, standardised tests, which supply norm-referenced inferences and provide the results with more rigorous evidence, have been used to a lesser degree. Most of the empirical research which applied standardized tests dealt with English language proficiency. On the other hand, there have been few empirical studies which have used standardized tests to verify L2 proficiency according to CEFR (Common European Framework for Languages), which is nowadays becoming a standardized framework for language abilities on an international scale. In her recent study in Castilla-La Mancha, Ruiz Cordero (2019) applied PET (Preliminary English Test) exams, which helped her not only to compare the L2 writing

skills progression to the mainstream counterparts but also to contrast the results to the EU standards. Hence, to consolidate the findings on CLIL efficiency, standardized tests in compliancy with CEFR are recommended to compare the levels of L2 acquisition with CLIL in different educational environments on a global scale.

Even though CLIL teachers have specified a lack of didactic resources in CLIL (Lorenzo & Granados, 2020; Nieto Moreno de Diezmas, 2019; Siepmann et al., 2021), which has also been stated by Marsh (2002) as an apparent weakness in any CLIL programme, surprisingly, MTE has been scanty. However, it is noteworthy to point out that new types of resources in the era of digitalization, such as augmented reality, have emerged (Çelik & Yangin Ersanlı, 2022). In this experimental study with the use of virtual objects in a physical environment of a CLIL high school classroom in Turkey, the results showed that not only the students improved their L2, but their motivation increased (Çelik & Yangin Ersanlı, 2022). Thus, CLIL didactic resources can be a powerful ally for the teachers, which not only serve as a guide for the correct implementation of this methodology, but they can also boost the learners' communicative competence and motivation. It would be highly advisable to carry out more research on the analysis of didactic resources in a CLIL classroom.

As for population groups, Coyle et al. (2010) mainly proposed students, teachers, and parents. Along the same line, most empirical research has focused on students, either alone or in concert with teachers and/or parents. Moreover, new stakeholders have emerged such as coordinators (Fernández-Barrera, 2019), CLIL specialists (Nieto Moreno de Diezmas, 2019), and school leaders (Bower, 2020). The fact that new stakeholders have been involved allows for a greater enrichment of CLIL studies, as well as providing the scientific community with a broader spectrum of research and a more panoramic view.

Scientific Interest in the 4 C's

Regarding the results of the content-based analysis aimed at exploring which of the 4Cs receives more and less scientific interest, it can be stated that Communication has by far been the most researched principle. There is undeniable interest in the results of CLIL apropos L2 acquisition. The launch of CLIL "was accompanied by hopes of it heralding in a period of change in education, with the most prominent expectations revolving around the desire for change in foreign language learning" (Dalton-Puffer et al., 2022, p. 183). Since then, this methodology has been viewed more as a linguistic phenomenon within a plurilingual educational agenda worldwide. The results of the empirical studies have demonstrated close relation between the instruction in a foreign language and enhanced L2 proficiency skills. It seems that in comparison to their mainstream counterparts CLIL learners have better linguistic performance in listen-

ing (Morilla García & Pavón Vázquez, 2018), oral production (Ruiz Cordero, 2022), reading, especially a better comprehension of lexical items (Nieto Moreno de Diezmas, 2018), writing (Ruiz Cordero, 2019), and higher levels of receptive vocabulary and vocabulary size (Castellano-Risco et al., 2020). However, most comparative studies in L2 proficiency were performed in Spain. Therefore, it can be recommended to carry out this type of research on a global level.

Moreover, the least studied principle in CLIL is Cognition. Notwithstanding the fact that creative and critical thinking is one of the educational demands of the 21st century, the question whether HOTS are worked with in a CLIL classroom remains open. In the mixed method study in Spain, the results showed that the teachers tended to lapse into working with LOTs, understanding being the predominant cognitive skill in a CLIL classroom (Campillo-Ferrer & Miralles-Martínez, 2022; Valverde Caravaca, 2019). Thus, through methodological intervention focused on the optimized use of questions, Valverde Caravaca (2019) demonstrated it as a valid strategy to foster critical thinking. More scientific research on the matter is required not only to demonstrate the development of critical thinking but also to propose effective strategies which can boost HOTS.

Emerging Areas of Scientific Research on CLIL

As for the analysis of recent areas of scientific research in CLIL, three areas have emerged: Communication, Focus on teacher, and Focus on student (affective factors).

Within Communication, L2 proficiency/development attracts a major part of scientific interest (Martínez Agudo & Fielden Burns, 2021). It is also noteworthy that another area of emerging research within PrE is in-class procedure in line with the one proposed by Coyle et al. (2010). Discourse analysis with transcripts of verbal reports of students' in-class interaction and teacher-students interaction has been the object of study to explore not only the teachers' language, for instance, the use of open questions as a scaffolding technique to foster interaction and science content in CLIL classes (Tagnin & Ní Ríordáin, 2021), but also to pair dynamics in task-based interaction (Basterrechea & Gallardo-del-Puerto, 2020), collaboration in written tasks (Jakonen, 2019).

Furthermore, since bilingualism is becoming a norm in our globalized world, there is a growing scientific interest in the interaction between the students' L1 and L2, as well as the impact of L2 on mother tongue development in compliance with Coyle et al.'s (2010) recommendations. In the experimental research on the impact of L2 on L1 written production in secondary schools in Spain (Nieto, 2020), the results showed that learners under CLIL instruction outperformed their mainstream counterparts despite limited L1 exposure. On a different note, several studies delved into the strategic use of L1 in a CLIL classroom, specifically translanguaging

(Nieto Moreno de Diezmas, 2018; Nikula & Moore, 2019). Considering that bilingualism is a relatively novel phenomenon on a global educational scope, more research is advised to delve into L1 via L2 development with CLIL instruction to pinpoint benefits and drawbacks.

Notwithstanding the fact that Coyle et al. (2010) proposed to explore affective factors related to teachers in terms of research with a Focus on teachers, it is the area of Pedagogies/didactic competence that has attracted more scientific interest (Ljalikova et al., 2021; Martínez Agudo & Fielden Burns, 2021). Since it is teachers who contribute to the shaping of the CLIL method and as such have become key players, the pedagogic quality of teachers has been put in the spotlight. Both linguistic proficiency and subject knowledge were stated as areas for improvement (Dvorjaninova & Alas, 2018; Huertas-Abril & Shashken, 2021; Lorenzo & Granados, 2020; Pappa et al., 2019), which may be one of the reasons why the synergy of Content and Communication has not been achieved. Consequently, in most research dealing with teachers' perspectives the need for a better professional development was stated (Dvorjaninova & Alas, 2018; Lorenzo & Granados, 2020; McDougald & Pissarello, 2020; Mearns & Graaff, 2018). CLIL training is not obligatory in many countries, which can lead to a misunderstanding of the method and its mediocre implementation. In a case study performed in Columbia with in-service teachers (McDougald & Pissarello, 2020), the results demonstrated that after receiving training in CLIL, the teachers improved significantly not only in the subject knowledge but also in CLIL strategies and lesson planning. In-service opportunities for professional development as well as exchange programmes were some of the improvements highlighted by the teachers in Finland (Pappa et al., 2019).

In the early implementation of CLIL, certain drawbacks were observed, such as the lack of collaboration and administrative support⁶ (Mehisto et al., 2008), which continue to be the major challenge shared by CLIL teachers (Lorenzo & Granados, 2020; Mearns & Graaff, 2018; Nieto Moreno de Diezmas, 2019; Pappa et al., 2019). In the qualitative analysis of CLIL teachers' narratives in relation to their teaching experiences in Estonia (Ljalikova et al., 2021), relative loneliness and the extreme importance of administrative support and collaboration were shared by all the stakeholders. This CLIL teachers' loneliness seems to be a common feature in other educational environments. Administrative tensions and apparent lack of administrative support were pinpointed by other studies in Spain (Nieto Moreno de Diezmas, 2019), Kazakhstan (Huertas-Abril & Shashken, 2021), the UK (Bower, 2020), Colombia (McDougald & Pissarello, 2020), etc. Even though Coyle et al. (2010) and Mehisto et al. (2008) stressed

the importance of collaboration at all levels of CLIL implementation, it is still a major weakness.

Finally, in relation to Affective factors (students), most studies explored the students' motivation in CLIL. There is a consensus that instruction in a vehicular language boosts students' motivation and increases their linguistic competence (Bower, 2019b; Mearns & Graaff, 2018; Roiha, 2019). Furthermore, it seems that instruction in CLIL has influenced students' personality traits, such as extraversion and agreeableness (Bowers, 2020). However, considering Coyle et al.'s (2010) proposal which stated the need to explore affective factors (AE) in students, parents, and teachers in equal measure, it would be advised to include teachers and parents in the research of affective factors. Furthermore, the joint evaluation of PE and AE has been scarce, albeit recommended by Coyle et al. (2010).

Limitations

As for the limitations of this study, the shortened period (2018-2022) selected by the researchers due to the vast quantity of publications on CLIL, if prolonged it might have strengthened the findings on the expansion of CLIL worldwide. Moreover, since only two databases (Wos and Scopus) were explored, other scientific studies, which could have contributed to the current research, were not considered. Another limitation was the language of the current study. Since only articles in English and Spanish were explored, 84 articles written in other languages were excluded. Thus, even though researchers from other countries had worked with this methodology, they were not visible in this study.

CONCLUSION

We consider that this systematic review has accomplished its pivotal objective to analyse CLIL-related research in school settings around the world. Even though CLIL was started in the EU, it has transpired European borders and there is a growing scientific interest in this method worldwide. Initially, the researchers expected CLIL to be studied in a wider spectrum of countries due to its importance to multilingualism and globalization, however, there are still countries which have not been prolific research wise. Thus, the findings suggest that further research is required to delve into possible factors which can booster CLIL implementation in different countries.

As for methodologies used in the studies, both quantitative and qualitative studies have been applied on an equal basis. However, to enhance the quality of the results, mixed method studies are recommendable. Despite a wide spec-

⁶ Eurydice Report. (2006). Content and Language Integrated Learning (CLIL) at school in Europe (Y. European Commission, Directorate-General for Education Sport and Culture, Ed.). Publications Office. <https://op.europa.eu/en/publication-detail/-/publication/756eb-daa-f694-44e4-8409-21eef02c9b9b>

trum of research tools applied, a questionnaire has been favoured by the researchers. On the other hand, the use of standardized tests in the quantitative studies, especially the ones related to measure students' linguistic competences, has been scarce. Moreover, there has been a disproportion between different areas of research: Performance Evidence, which is more concerned with evaluating learners' outcomes in CLIL, has emerged as the most prominent area of research. Other areas of evaluation portfolio have been less studied notwithstanding their importance to evaluating in-class procedures and affective factors of CLIL stakeholders.

Since in a globalized world the knowledge of English is an essential requirement, Communication principle has attracted the major scientific interest. As a result, other CLIL principles have been overlooked. Cognition has been the less studied principle, albeit the development of HOTS is one of the educational goals of this century. On the other hand, it is very positive that other fields of research in the study of CLIL have emerged, such as bilingualism, and CLIL impact on L1 development. These findings open the possibility of studying the interaction between the two languages

and their effect on better implementation of bilingual programmes in education.

We believe that the findings of this study will appeal not only to CLIL researchers but to a wider community of CLIL stakeholders, such as policy makers, CLIL coordinators, and teachers.

DECLARATION OF COMPETING INTEREST

None declared.

AUTHORS' CONTRIBUTION

Belen Poveda García-Noblejas: methodology, data curation, writing review.

Svetlana Antropova: conceptualization, writing – original draft, editing.

REFERENCES

- Antropova, S., & Poveda García-Noblejas, B. (2023), Database: The worldwide expansion of CLIL in school settings. *Mendeley Data*, V1. <https://doi.org/10.17632/db8xc2gd6y.1>
- Basterrechea, M., & Gallardo-del-Puerto, F. (2020). Language-related episodes and pair dynamics in primary school CLIL learners: A comparison between proficiency-matched and student-selected pairs. *Studies in Second Language Learning and Teaching*, 10(3), 423–447. <https://doi.org/10.14746/ssl.t.2020.10.3.2>
- Bauer-Marschallinger, S., Dalton-Puffer, C., Heaney, H., Katzinger, L., & Smit, U. (2021). CLIL for all? An exploratory study of reported pedagogical practices in Austrian secondary schools. *International Journal of Bilingual Education and Bilingualism*, 1–16. <https://doi.org/10.1080/13670050.2021.1996533>
- Bower, K. (2019a). Explaining motivation in language learning: A framework for evaluation and research. *The Language Learning Journal*, 47(5), 558–574. <https://doi.org/10.1080/09571736.2017.1321035>
- Bower, K. (2019b). 'Speaking French alive': Learner perspectives on their motivation in Content and Language Integrated Learning in England. *Innovation in Language Learning and Teaching*, 13(1), 45–60. <https://doi.org/10.1080/17501229.2017.1314483>
- Bower, K. (2020). School leaders' perspectives on content and language integrated learning in England. *Language, Culture and Curriculum*, 33(4), 351–367. <https://doi.org/10.1080/07908318.2019.1667367>
- Byram, M. (1997). *Teaching and assessing intercultural communicative competence*. Multilingual Matters.
- Byram, M. (2012). Language awareness and (critical) cultural awareness – relationships, comparisons and contrasts. *Language Awareness*, 21(1–2), 5–13. <https://doi.org/10.1080/09658416.2011.639887>
- Campillo-Ferrer, J.-M., & Miralles-Martínez, P. (2022). Primary school teachers' perceptions of the level of development of low-order cognitive skills under the Content and Language Integrated Learning Approach. *Frontiers in Education*, 7. <https://doi.org/10.3389/educ.2022.815027>
- Castellano-Risco, I., Alejo-González, R., & Piquer-Píriz, A. M. (2020). The development of receptive vocabulary in CLIL vs EFL: Is the learning context the main variable? *System*, 91, 102263. <https://doi.org/10.1016/j.system.2020.102263>
- Çelik, F., & Yangın Ersanlı, C. (2022). The use of augmented reality in a gamified CLIL lesson and students' achievements and attitudes: a quasi-experimental study. *Smart Learning Environments*, 9(1), 9–30. <https://doi.org/10.1186/s40561-022-00211-z>
- 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>

- Cimermanová, I. (2021). A review of European research on Content and Language Integrated Learning. *Integration of Education*, 25(2), 192–213. <https://doi.org/10.15507/1991-9468.103.025.202102.192-213>
- 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>
- Coyle, D., Hood, P., & Marsh, D. (2010). *CLIL*. Cambridge University Press. <https://doi.org/10.1017/9781009024549>
- Dalton-Puffer, C., Hüttner, J., & Llinares, A. (2022). CLIL in the 21st century. *Journal of Immersion and Content-Based Language Education*, 182–206. <https://doi.org/10.1075/jicb.21021.dal>
- Díaz Herrera, C. (2018). Qualitative research and thematic content analysis. Intellectual orientation of *Universum* journal. *Revista General de Información y Documentación*, 28(1). <https://doi.org/10.5209/RGID.60813>
- Dvorjaninova, A., & Alas, E. (2018). Implementing Content and Language Integrated Learning (CLIL) in Estonia: Subject and language teacher perspective. *Eesti Rakenduslingvistika Ühingu Aastaraamat. Estonian Papers in Applied Linguistics*, 14, 41–57. <https://doi.org/10.5128/ERYa14.03>
- Fernández-Barrera, A. (2019). Doing CLIL in the Science classroom: A critical sociolinguistic ethnography in La Mancha secondary schools. *Foro de Educación*, 17(27), 37–63. <https://doi.org/10.14516/fde.712>
- Gil-López, V., González-Villora, S., & Hortigüela-Alcalá, D. (2021). Learning foreign languages through content and language integrated learning in physical education: A systematic review. *Porta Linguarum*, 35, 165–182. <https://doi.org/10.30827/portalin.v0i35.15785>
- Goris, J., Denessen, E., & Verhoeven, L. (2019). Effects of content and language integrated learning in Europe: A systematic review of longitudinal experimental studies. *European Educational Research Journal*, 18(6), 675–698. <https://doi.org/10.1177/1474904119872426>
- Graham, K. M., Choi, Y., Davoodi, A., Razmeh, S., & Dixon, L. Q. (2018). Language and content outcomes of CLIL and EMI: A systematic review. *Latin American Journal of Content & Language Integrated Learning*, 11(1), 19–38. <https://doi.org/10.5294/laclil.2018.11.1.2>
- Hernandez-Sampieri, R., & Mendoza, C. (2018). *Metodología de la investigación: Las rutas cuantitativa, cualitativa y mixta*. McGraw Hill.
- Huertas-Abril, C. A., & Shashken, A. (2021). Exploring the potential of CLIL in Kazakhstan: A qualitative study. *Revista Complutense de Educación*, 32(2), 261–271. <https://doi.org/10.5209/rced.68345>
- Hüttner, J., Dalton-Puffer, C., & Smit, U. (2013). The power of beliefs: Lay theories and their influence on the implementation of CLIL programmes. *International Journal of Bilingual Education and Bilingualism*, 16(3), 267–284. <https://doi.org/10.1080/13670050.2013.777385>
- Jakonen, T. (2019). The integration of content and language in students' task answer production in the bilingual classroom. *International Journal of Bilingual Education and Bilingualism*, 22(4), 428–444. <https://doi.org/10.1080/13670050.2016.1267694>
- Kovacikova, E. (2019). Development of speaking at primary schools through CLIL. *XLinguae*, 12(2), 17–26. <https://doi.org/10.18355/XL.2019.12.02.02>
- Li, L., Huang, F., Chen, S., Pan, L., Zeng, W., & Wu, X. (2020). Exploring the curriculum development in content and language integrated learning: A systematic review. *International Journal of Evaluation and Research in Education*, 9(4), 1102–1113. <https://doi.org/10.11591/ijere.v9i4.20705>
- Llinares, A., & Morton, T. (Eds.). (2017). *Applied linguistics perspectives on CLIL* (vol. 47). John Benjamins Publishing Company. <https://doi.org/10.1075/llt.47>
- Ljalikova, A., Meristo, M., Alas, E., & Jung, M. (2021). Narrative analysis as a means of investigating CLIL teachers' meaningful experiences. *Qualitative Research in Education*, 10(3), 228–259. <https://doi.org/10.17583/qre.7511>
- Lockwood, C., & Tricco, A. C. (2020). Preparing scoping reviews for publication using methodological guides and reporting standards. *Nursing & Health Sciences*, 22(1), 1–4. <https://doi.org/10.1111/nhs.12673>
- Lorenzo, F., & Granados, A. (2020). One generation after the bilingual turn: Results from a large-scale CLIL teachers' survey. *Elia*, 20, 77–111. <https://doi.org/10.12795/elia.2020.i20.04>
- Marsh, D. (2002). CLIL/EMILE- The European dimension: Actions, trends and foresight potential. European Commission.
- Martínez Agudo, J. de D., & Fielden Burns, L. V. (2021). What key stakeholders think about CLIL programmes: Commonalities and differences of perspective. *Porta Linguarum*, 35, 221–237. <https://doi.org/10.30827/portalin.v0i35.15320>
- McDougald, J. S., & Pissarello, D. (2020). Content and Language Integrated Learning: In-service teachers' knowledge and perceptions before and after a professional development program. *Íkala*, 25(2), 353–372. <https://doi.org/10.17533/udea.ikala.v25n02a03>

- Mearns, T., & Graaff, R. de. (2018). Bilingual education and CLIL in the Netherlands. *Dutch Journal of Applied Linguistics*, 7(2), 122–128. <https://doi.org/10.1075/dujal.00002.int>
- Mehisto, P., Marsh, D., & Frigols, M. J. (2008). *Uncovering CLIL. Content and Language Integrated Learning in bilingual and multi-lingual education*. Macmillan Publishers.
- Merino, J. A., & Lasagabaster, D. (2018). The effect of content and language integrated learning programmes' intensity on English proficiency: A longitudinal study. *International Journal of Applied Linguistics*, 28(1), 18–30. <https://doi.org/10.1111/ijal.12177>
- Molina-Azorin, J. F., & Fetters, M. D. (2019). Building a better world through mixed methods research. *Journal of Mixed Methods Research*, 13(3), 275–281. <https://doi.org/10.1177/1558689819855864>
- Morilla García, C., & Pavón Vázquez, V. (2018). Psychopedagogical factors that affect L2 listening acquisition in diverse Spanish bilingual and non-bilingual instructional settings: Multiple intelligences influence. *Journal of English Studies*, 16, 185. <https://doi.org/10.18172/jes.3550>
- Nieto, E. (2020). Mother tongue development in bilingual programs type CLIL in secondary school: A comparative study on written production. *RLA. Revista de Lingüística Teórica y Aplicada*, 58(2), 117–136. <https://doi.org/10.29393/RLA58-11MTEN10011>
- Nieto Moreno de Diezmas, E. (2018). Acquisition of reading comprehension in L1 in bilingual programmes of primary education. A comparative study. *Ocnos*, 17(1), 43–54. https://doi.org/10.18239/ocnos_2018.17.1.1471
- Nieto Moreno de Diezmas, E. (2019). Students, teachers and management teams in bilingual programmes: Shared perceptions and areas for improvement. *Journal of English Studies*, 17, 277. <https://doi.org/10.18172/jes.3564>
- Nikula, T., & Moore, P. (2019). Exploring translanguaging in CLIL. *International Journal of Bilingual Education and Bilingualism*, 22(2), 237–249. <https://doi.org/10.1080/13670050.2016.1254151>
- Palacios-Hidalgo, F. J., Huertas-Abril, C. A., & Gómez-Parra, M. E. (2021). Foreign and bilingual language education in the UK and Spain: A study of similarities and differences. *Journal of Language and Education*, 7(2), 243–255. <https://doi.org/10.17323/jle.2021.11938>
- Pappa, S., Moate, J., Ruohotie-Lyhty, M., & Eteläpelto, A. (2019). Teacher agency within the Finnish CLIL context: Tensions and resources. *International Journal of Bilingual Education and Bilingualism*, 22(5), 593–613. <https://doi.org/10.1080/13670050.2017.1286292>
- Pfenninger, S. E. (2020). The dynamic multicausality of age of first bilingual language exposure: Evidence from a longitudinal Content and Language Integrated learning study with dense time serial measurements. *The Modern Language Journal*, 104(3), 662–686. <https://doi.org/10.1111/modl.12666>
- Roiha, A. (2019). Investigating former pupils' experiences and perceptions of CLIL in Finland: A retrospective analysis. *Nordic Journal of Studies in Educational Policy*, 5(2), 92–103. <https://doi.org/10.1080/20020317.2019.1586514>
- Ruiz Cordero, M. B. (2019). Assessing English writing skills of students from bilingual and non-bilingual schools in Castilla-La Mancha, Spain. A comparative study. *Theory and Practice of Second Language Acquisition*, 5(2), 95–113. <https://doi.org/10.31261/tapsla.7658>
- Ruiz Cordero, M. B. (2022). The oral skill English level in bilingual and non-bilingual educational centres of Castilla-La Mancha. *Revista Complutense de Educación*, 33(2), 201–213. <https://doi.org/10.5209/rced.73909>
- Siepmann, P., Rumlich, D., Matz, F., & Römhild, R. (2021). Attention to diversity in German CLIL classrooms: Multi-perspective research on students' and teachers' perceptions. *International Journal of Bilingual Education and Bilingualism*. <https://doi.org/10.1080/13670050.2021.1981821>
- Stimson, R. (2014). *Handbook of research methods and applications in spatially integrated social science*. Edward Elgar Publishing. <https://doi.org/10.4337/9780857932976>
- Tagnin, L., & Ní Riordáin, M. (2021). Building science through questions in Content and Language Integrated Learning (CLIL) classrooms. *International Journal of STEM Education*, 8(34). <https://doi.org/10.1186/s40594-021-00293-0>
- Valverde Caravaca, R. (2019). Effective questioning in CLIL classrooms: Empowering thinking. *ELT Journal*, 73(4), 367–376. <https://doi.org/10.1093/elt/ccz030>
- Van Mensel, L., Hiligsmann, P., Mettwie, L., & Galand, B. (2020). CLIL, an elitist language learning approach? A background analysis of English and Dutch CLIL pupils in French-speaking Belgium. *Language, Culture and Curriculum*, 33(1), 1–14. <https://doi.org/10.1080/07908318.2019.1571078>