JoLaCE

Journal of Language and Cultural Education, 2020, 8(1)
ISSN 1339-4584

sciendo

DOI: 10.2478/jolace-2020-0003

Meta-analysis of studies on the acquisition of receptive skills and vocabulary in CLIL

Ivana Cimermanová
University of Presov, Slovakia
ivana.cimermanova@unipo.sk

Abstract

Content and language integrated learning (CLIL) is a major area of interest within the field of formal education. There are numerous studies presenting data and results of CLIL implementation. The positive impacts have been reported in building positive attitudes to language learning, to content subject learning, increasing efficacy of language learning. Questions have been raised about the factors that (may) affect research results and their interpretation. Many small studies bring statistically non-significant data as they use small convenience samples. Meta-analyses enable the researchers to synthesise data from research with the same characteristics. The present article analyses the studies that focus on CLIL implementation at primary and secondary schools with special focus on receptive skills and vocabulary gains. Out of 385 selected studies were 9 included and applying randomised-effect model evaluated. The analysis found no statistically significant differences between the CLIL and EFL groups in listening and reading performance. Concerning vocabulary the statistically significant difference in favour of CLIL ($p<0,0001$) with overall estimate effect 0,84 and confidence interval ranging from 0,56 to 1,11 was observed.

Keywords: CLIL, receptive skills, meta-analysis, statistics, forest plot, confidence interval

1 Introduction

Foreign language teaching is compulsory in almost all European countries. It might seem to be surprising that almost in half of the European countries a mandatory foreign language is specified (European Commission/EACEA/Eurydice, 2017). To increase the exposure to foreign (or minority) languages, to enhance students' proficiency, make foreign language learning more meaningful and to prepare students to be able to communicate the content the Content and language integrated learning (CLIL) has been applied in selected schools.

CLIL is "a dual-focused educational approach in which an additional language is used for the learning and teaching of both content and language. That is, in the teaching and learning process, there is a focus not only on content, and not only on language. Each is interwoven, even if the emphasis is greater on one or the other at a given time" (Coyle, Hood & Marsh, 2010). In some European countries (e.g. Italy, Cyprus, Liechtenstein, Austria) at least one year (at least) one non-language subject has to be taught in a foreign language (see European Commission/EACEA/Eurydice, 2017).

CLIL has been the subject of much systematic investigation. Various aspects of the linguistic potential of CLIL have been extensively discussed similarly as different factors influencing the effectiveness of CLIL programmes.

During its development and implementation it has been compared and/or considered as a synonym to integrated thematic instruction (school model designed by Kovalik, see Kovalik & Olsen, 1993), immersion (Lasagabaster & Sierra, 2009), content-based instruction (Cenoz, 2015), task-based language teaching (Ortega, 2015), English for specific purpose (Yang, 2016; Taillefer, 2013; Tzoannopoulou, 2015) or bilingual education (see e.g. Nikula, 2018). The studies have been focussed on key actors: pupils, teachers, parents and school management. The impact on foreign language performance, content knowledge and also the impact on mother tongue, classroom interaction (Pastrana, Llinares & Pascual, 2018), the **influence of affective factors** (Lasagabaster & Doiz, 2015; Otwinowska & Foryś, 2017), **household structure** (Mensel, Hilligsmann, Mettewie & Galand, 2020), **time and intensity** (Surmont, et al., 2016; Merino & Lasagabaster, 2017), **age** (Roquet, 2015), **motivation** (Fontecha & Canga Alonso, 2014), **gender** (Canga Alonso, 2016; Fontecha & Canga Alonso, 2014), strategies (Zarobe, 2017; Straková, 2020) and other aspects has been the subject of numerous studies.

It has been already mentioned that in CLIL we deal with dual-focused teaching, unfortunately, it seems that it is mostly the domain of language teaching and language teachers who are actively involved in CLIL implementation. It similarly seems that the majority of studies focus on language aspects. The studies focus not only on general foreign **language performance** (Isidro & Lasagabaster, 2018a; Salamanca & Montoya, 2018; Merino & Lasagabaster, 2017; Pérez-Cañado, 2018; Dallinger, et al., 2016) but also on possible growth of **vocabulary** (Castellano-Risco, et al., 2020; Moghadam & Fatemipour, 2014), specifically of **receptive vocabulary** (Castro-García, 2017; Augustín-Llach & Canga Alonso, 2016; Canga Alonso, 2016,), development of **listening skills** (Diemaz, 2018; Pladevall-Ballester & Vallbona, 2016; Dallinger, et al., 2016; Mattheoudakis, 2014), **reading skills** (Hamidavi, Amiz & Gorjan, 2016; Pladevall-Ballester & Vallbona, 2016; Chostelidou & Griva, 2014; Mattheoudakis, et al., 2014), **receptive skills** (Prieto-

Arranz, et al., 2014), **writing** (Lo & Jeong, 2018; Gené-Gil, et al., 2015), **grammar** (Pérez-Cañado & Basse, 2015), **quality of argumentations** (Myskow & Ono, 2018; Morton & Llinares, 2016), **density and type of errors** (Pérez-Cañado & Basse, 2015; Pérez-Vidal & Roquet, 2015), **pronunciation and fluency** (Zarobe, 2008**), code-switching** (Isidro & Lasagabaster, 2018b) and other aspects of language teaching and learning.

2 The effect of CLIL on foreign language proficiency

There are a plethora of CLIL studies that focus on linguistic gains. They focus on different aspects of language learning, they are realised in different contexts, they work with different sample sizes and their significance differ. CLIL is studied across all age levels, still, it seems that most studies concentrate at primary level. CLIL is widely spread in Spain and this is also reflected in the volume of published studies. Bruton (2011) interpreting CLIL results in Spain stated that "the results do not show that the CLIL group's performance was better on most counts and even if they did with such a small sample, the results would be of dubious validity" (p. 525).

Pérez-Cañado (2018a) notes there is the controversy affecting CLIL characterization, she stresses the problem with "no uniform teaching methodology" and "lack of a coherent methodology". She further highlights the necessity of the revision of the role of teacher and the need to realise that "communicative teaching should underpin CLIL, and fluency and oracy should be awarded primacy over accuracy and written skills" (p. 372).

The present study tries to synthesise the selected studies that focus on the development of receptive skills and vocabulary. The focus is on the groups of students aged 11-17. Isidro and Lasagabaster (2018b) studied the influence of CLIL in groups of 14-15 years olds and the gains in global competence and all 4 language skills (2nd years study). Both groups improved but the CLIL group reached statistically significant better results. The researchers realised 2 measurements (after the 1st year and after the second one) and the results were not the same. No progress was recorded in the non-CLIL group between the 1st and 2nd measurements in listening, writing and speaking and the progress in global competence and reading was observed only after the 2nd year. Concerning the CLIL group, the positive gain in listening and speaking was observed only after one year. They stress (similarly as Pérez-Cañado indicated above) that further research should "pay special attention to the pedagogical features and the methodology (that is, the set of methodological practices related to the key pedagogic approaches that make CLIL identifiable as classroom practice) employed in the CLIL contexts under scrutiny, because this will help researchers to reach more robust conclusions on the impact of this particular approach" (p.16).

Nietro Moreno de Diezmas (2018a) divided the group of CLIL students according to the grade they studied in and studied the influence of CLIL on listening skills and vocabulary development. Concerning global comprehension, CLIL students scored significantly higher, however, non-CLIL students outperformed non-CLIL group in vocabulary (p=0,000) and understanding of space-time relations.

Merino and Lasagabaster (2017) also studied individual language skills and total results in a year study with a sample of 11-13 years olds. They divided CLIL groups according to the intensity what enabled them not only to compare EFL and CLIL groups but also to consider the importance of CLIL intensity. They indicate that "CLIL will only produce a significant EFL improvement when it is part of a high intensity programme" (p. 27).

Even though there is a high number of studies realised at the primary and a secondary level, university level is also in the focus of researchers. Chostelidou and Griva (2013) studied the development of reading skills at tertiary level and they found statistically significant differences in the post-intervention measurements and similarly Gorjian and Hamidavi (2017) focussed their attention on university students and their results showed statistically significant differences in the post-tests focussed on vocabulary development.

Sylvén described contextual differences of 4 European countries and analysed possible factors that may influence the success of CLIL. She mentions policy, teacher (education), age (and cognitive development) and extramural English (and the amount of exposure) as the key factors that may influence the result. She states that regarding the amount of exposure to English outside of school there are huge differences and Sweden, being at the top of the countries compares also reached the highest scores in English language skills. This might be also one of the reasons why in some countries is CLIL not so successful.

The controversy in CLIL application needs further study with the focus on the factors that influence the results and the impact on both, foreign language and content learning, and whatmore, mother tongue (see e.g. Pérez-Cañado, 2017; Nieto Moreno de Diezmas, 2018).

The systematic review and meta-analysis that allow the research to synthesise data and interpret them are also one of the possibilities of how to include studies with non-significant results, small samples, etc.

3 Methods

There are numerous studies that from various reasons work with small samples and thus their results are not considered as reliable. Systematic reviews and meta-analysis can be used to aggregate the effect size by integrating the results of different studies (selected based on defined criteria according to the research

question). The heterogeneity of the studies can influence the results and their interpretation.

Review question

This review explores the effectiveness of a CLIL on receptive skills and vocabulary gains in a foreign language in the group of students aged 10-17.

Selection of the studies

To identify the relevant information the Web of Science Core Collection was used as a source of high-quality peer-reviewed studies, namely 4 databases. Databases searched included (1) Science Citation Index Expanded (SCI-EXPANDED), (2) Social Sciences Citation Index (SSCI), (3) Arts & Humanities Citation Index (A&HCI) (4) Emerging Sources Citation Index (ESCI). The conference Proceedings were intentionally omitted. The timespan was limited to the studies published not sooner than 2010.

The basic search (looking for "CLIL research" studies) resulted in 385 studies (see figure 1) out of which 332 texts written in English, 6 in Russian and 3 in German were selected.

Fig. 1: WOS categories statistics

Inclusion criteria

The abstracts of 341 studies were carefully read and to be included those studies the studies had to (1) apply quantitative research methods, (2) the sample

age corresponds to the research question (10-16years) and (3) possibly provide statistical data (n, mean, SD) (4) comparing intervention and conventional group (5) with the focus on foreign language development. The list of studies was reduced to 41 out of which after full text reading the following studies were selected as the subject of the present analysis (see the following table).

Tab. 1: Details of included studies

Study ID	Study authors(s)	year	source	participants	used in the subgroup		
					listening	reading	vocabulary
Agudo 2019	Agudo, J. D.	2019	Which instructional programme (EFL or CLIL) results in better oral communicative competence? Updated empirical evidence from a monolingual context. *Linguistics and Education, 51,* 69-78.	318	●		
Agustín-Llach 2014	Agustín-Llach, M. P., & Alonso, A. C.	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.	107			●
Cañado 2018	Cañado, M. L.	2018	CLIL and Educational Level: A Longitudinal Study on the Impact of CLIL on Language Outcomes. *Porta Linguarum, 29* (January), 51–70.	2024	●	●	●
Castellano-Risco 2020	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.	138			●
Castro-García 2017	Castro-García, D.	2017	Receptive vocabulary measures for EFL Costa Rican high school students. *International Journal of English Studies, 17* (2), 81-99.	85			●
Dallinger 2016	Dallinger, S., Jonkmann, K., Hollm, J., & Fiege, C.	2016	The effect of content and language integrated learning on students' English and history competences – Killing two birds with one stone? Learning and Instruction, 41, 23-31.	837	●		
Hamidavi 2016	Hamidavi, N., Amiz, M., & Gorjian, B.	2016	The Effect of CLIL Method on Teaching Reading Comprehension to Junior High School Students. *Modern Journal of Language Teaching Methods, 6* (9), 64-73.	60		●	
Mattheoudakis 2014	Mattheoudakis, M., Alexiou, T., & Laskaridou, C.	2014	To CLIL or Not to CLIL? The Case of the 3rd Experimental Primary School in Evosmos. *Major Trends in Theoretical and Applied Linguistics. vol.* 3. Berlin, Boston: De Gruyter.	31	●	●	
Pladevall-Ballester 2016	Pladevall-Ballester, E., & Vallbona, A.	2016	CLIL in minimal input contexts: A longitudinal study of primary school learners' receptive skills. *System, 58* , 37-48.	287	●	●	

After obtaining and reading the full texts of the selected studies a total of 32 studies were excluded as the data did not follow the design or there were no comparisons of experimental and control groups, the age of the respondents did not fall to the set limit or they did not focus on L2. For example, the research conducted by Canga Alonso published in 2013 focused on the receptive vocabulary of Spanish 6th grade primary school CLIL students (n=79, aged 11-12 years old), however, he worked only with one CLIL group and studied the sex-based differences related to receptive vocabulary size and comprehension abilities. Similarly, e.g. the research reported by Zarobe (2017) was excluded as the tool used to map the reading abilities (application of the reading strategies) was an interview, results of which were statistically elaborated.

In some cases (Agudo, 2019; Pérez-Cañado, 2018; Hamidavi, 2016) data provided were split by the researchers into two groups, e.g. according to sex or school attended. In those cases, we combined the reported subgroups into a single group. The combined mean was computed as the weighted mean across groups:

$$\bar{x}_1 = \frac{n_{11}\bar{x}_{11} + n_{12}\bar{x}_{12}}{n_{11} + n_{12}}$$

and the combined standard deviation was computed as

$$S1 = \sqrt{\frac{(n_{11}-1)S_{11}^2 + (n_{12}-1)S_{12}^2 + \frac{n_{11}n_{12}}{n_{11}+n_{12}}(\bar{x}_{11} - \bar{x}_{12})}{n_{11} + n_{12} - 1}}$$

where $\bar{x}_{11}$, $\bar{x}_{12}$ are the means in subgroups 1 and 2 of treatment group; S_{11}, S_{12} the standard deviations, and n_{11}, n_{12} the sample sizes; of subgroups (Borenstein, et al 2009, p. 222).

Comprehensive meta analysis (CMA version 3.3.070, trial/evaluation version) and RevMan (Review Manager 5) software were used to conduct a meta-analysis. The measures of the effect of the intervention were generally continuous data based on results obtained in a test and we used mean and standard deviation to compare the effect. Even though we tried to select the studies that met set criteria the effect size could vary according to the not controlled variables, or the broader set limit (e.g. age, different populations) and thus we applied random-effects model. The level of statistical significance was set at $p<0,05$. As the studies in the analysis did not use the same scale it would be not appropriate to use raw differences in means and thus to assess the outcome the standardised mean difference (δ) and the unbiased estimate of δ (Hedges' g) were used. Glenn (2016) introduces three levels or categories of effects (a) small effect (cannot be discerned by the naked eye) = 0,2, (b) medium effect = 0,5 and (c) large effect = 0,8.

The studies were evaluated separately according to the data they provide – listening, reading and vocabulary. Even though the studies met the set criteria for inclusion there were still variables that may influence the interpretation of data. It has to mentioned that the national policies influence the significance and application of CLIL. In some countries (e.g. Spain) it is substantially supported by the government, in other countries it is realised systematically but offered as an option (e.g in Germany students apply to secondary schools with CLIL programme) and there are countries where it is implemented not systematically but rather randomly depending on the capacities and willingness of teachers and approvement of the school management and parents. (In 2015, e.g. Pokrivčáková

stated about the status of CLIL in Slovakia: "The initiative to start CLIL mostly comes from "below", i.e. school managements or individual teachers. Many schools have started it through various school projects…" (p. 17).) Thus, the studies realised in different settings can differ in a number of instruction hours, number of content subjects, foreign language exposure time.

4 Results

4.1 Listening

The following table brings the basic information on the studies focussing on listening. The research design of all studies was an experiment based on pre-test/post-test. Two studies (Agudo, 2019, Cañado 2018) present also data on delayed testing as they study also the sustainability of the experiment results or impact.

Tab. 2: Details of included studies – part Listening

Study ID	country	sample grade (age)	no of CLIL/ non CLIL	research design / instrument	length	favor; effect size; p
LISTENING						
Agudo 2019	Spain	6th grade PE (11-12) 4th grade SE (15-16)	156/162	The listening test was designed to assess partic-ipants' oral comprehension in the target language in which they must deduce meanings and draw inferences from brief dialogues. The test consisted of different dialogues containing true/false, matching and mult. choice questions.	3 years	CLIL;-0,122; p=0,443 CLIL; -0,786; p<0,001*
Cañado 2018	Spain	6th grade PE (11-12) 4th grade SE (15-16)	1033/991	Specifically designed English language tests (validated for the study); comprised the use of English, vocabulary, reading, writing and speaking parts with a total score of 100 points	4 ac. years	CLIL; -0,223; p<0,001* CLIL; -0,873; p<0,001*
Dallinger 2016	Germany	8th grade (13,5 avg age)	483/354 (reduced)	C test consisting of 159 items (3 texts out of 7 used); the listening comprehension test consisted of 9 tasks with 4 open or colose items.	data collected autumn 2012-summer 2013, CLIL applied since 5th grade	CLIL
Mattheoudakis 2014	Greece	6th grade (11-12)	26/25	A language test was designed by the English language teachers of the school aiming to test CLIL and non CLIL learners' reading and listening skills in English (the same test before and after intervention)	9 months, 1 ac. year	
Pladevall-Ballester 2016	Spain	5th-6th grades	138/149	Cambridge young learners test	20 months, 2 ac. years	non-CLI; p=0,0078*

Agudo (2019) in his longitudinal research focused his attention on the possible development of oral communicative competence in a CLIL context. The study also brings data on language gains in time and Agudo states that the results indicate that the positive effects of CLIL on oral competence are visible with time. Concerning listening skills, he observed similar results (comparing EFL and CLIL groups) at the end of PE (with low Cohen's d = -0,122). When finishing their

Compulsory Secondary Education studies, a statistically significant difference was observed (p<0,001) in favour of CLIL students with Cohen's d= 0,659. As the data for both measurements were available (PE and SE separately) we also tried to evaluate the data separately but the summary result was not significant (p=0,92) with small effect size (0,12) and CI from -0,19 to 0,43). The present study used combined data from Agudo's study.

The delayed post-test was also applied by Pérez-Cañado (2018). She presents the data of the broad study and she describes the system of CLIL application in details. CLIL teaching in Spain takes place 5 hours a week and is compounded with foreign language instruction (3-4 hours a week). This may suggest (and this is not just the case of the Spanish system) that the number of language classes increases and thus language gains are the logical result and frequently it is not just the language but also the content subject(s) that is (are) observed and evaluated. In her study the researchers devoted a year to make sure they work with homogeneous groups. The research considered verbal intelligence, motivation, socioeconomic status, type of school, setting, exposure to English outside school as moderating variables. Concerning language testing the researcher used specifically designed language test that was validated for the study. It comprised the use of English, vocabulary, reading, writing and speaking parts with a total score of 100 points. Cohen's d was small at PE (-0,223) and large at SE (-0,873) in favour of CLIL group with p<0,001 in both cases and Pérez-Cañado (2018) similarly to Agudo (2019) states: "the longer the students have been benefitting from bilingual education, the greater the differences with their non-bilingual counterparts".

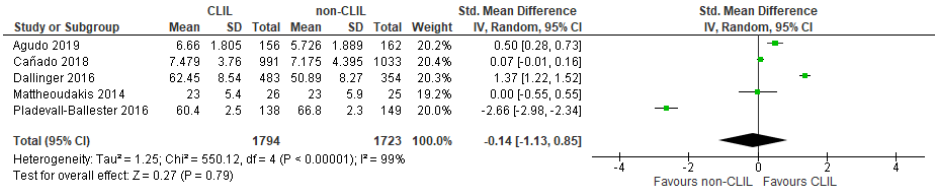
The special attention on content was paid in the study conducted by Mattheoudakis, et al (2014). Both, experimental and control groups were instructed 8 lessons per week and the 2 Geography lessons were taught in English in a CLIL group. The researchers used content (Geography) tests and language test to answer their research questions about the content and language gains and the possible correlation between the content and language achievements. The language of content test differed in the groups. The experimental (CLIL) group was tested in English while the control (non-CLIL) group was tested in the Greek language. Even though both groups' language performance was improved, there were no statistically significant differences observed.

Out of 5 included studies focussed on listening there is just 1 study with the results that favour non-CLIL group. Pladevall-Ballester and Vallbona (2016) realised the research with the sample of 287 5th and 6the graders (138 in CLIL group and 149 in non-CLIL group). Similarly as Mattheoudakis, et al (2014), their groups did not differ statistically significantly but the control (non-CLIL) group outperformed the experimental one. The researchers studied not only the impact

of CLIL on language gains but also the influence of time (the tests were realised after the first and after the second year of the experiment).

The following figure (with forest plot) and table summarise the data from the included studies.

Fig. 2: Forest plot illustrating the results in Listening using random effect model



Tab. 3: Statistics for the studies and summary

Study ID	Hedges's g and 95% CI						
	Hedges's g	Stnd. error	Variance	Lower limit	Upper limit	Z-Value	p-Value
Agudo, 2019	0,504	0,114	0,013	0,281	0,727	4,435	0,000*
Cañado, 2018	0,074	0,044	0,002	-0,013	0,161	1,668	0,095
Dallinger, 2016	1,371	0,078	0,006	1,219	1,522	17,681	0,000*
Mattheoudakis, 2014	0,000	0,276	0,076	-0,541	0,541	0,000	1,000
Pladevall Ballester, 2016	-2,662	0,162	0,026	-2,979	-2,344	-16,435	0,000*
Random model	-0,137	0,506	0,256	-1,129	0,855	-0,270	0,787

The confidence interval of 2 studies, similarly as the summary result include zero what means that the differences are not statistically significant (applying the fixed model the total result would be statistically significant). The study conducted by Pladevall-Ballester & Vallbona (2013) shows a statistically significant negative effect, the confidence interval is entirely on the negative side of zero [-2,98;-2,34]; studies conducted by Agudo (2019) and Dallinger, et al (2016) show a statistically significant positive effect. The highest effect size (g=-2,66) was recorded in the study by Pladevall-Ballester & Vallbona (2013).

The results show that the effect sizes are not consistent from study to study, they fall in the range of -2,66 to 1,37; the proportion of observed variance (I^2) is

very high (99%) what means we deal with substantial heterogeneity. The combined effect size is -0,14 (what can be evaluated as small effect) with a 95% confidence interval of -1,13 to 0,85. Confidence intervals are broader as we deal with the random model. The p-value for the summary effect is 0,79. The variance of dispersion (τ^2) that reflects the variance of the true effect is 1,25 what is also high with the standard deviation 1,12.

4.2 Reading

Four included studies presented the information on the impact of CLIL on development of reading skills. The research design of all studies was an experiment based on pre-test/post-test.

Tab. 4: Details of included studies – part Reading

Study ID	country	sample grade (age)	no of CLIL/ non CLIL	research design / instrument	length	favor; effect size; p
READING						
Cañado 2018	Spain	6th grade PE (11-12) 4th grade SE (15-16)	1033/991	Specifically designed English language tests (validated for the study); comprised the use of English, vocabulary, reading, writing and speaking parts with a total score of 100 points	4 ac. years	CLIL; -0,525; p<0,001* CLIL; -0,755; p<0,001*
Hamidavi 2016	Iran	(12,14 avg age)	15/15 high achievers 15/15 low ach.	Oxford quick placement test, pre-test/post-test; 50 multiple choice items	pre-test + 10 sessions + post-test	CLIL; p<0,05* CLIL; p<0,05*
Mattheoudakis 2014	Greece	6th grade (11-12)	26/25	A language test was designed by the English language teachers of the school aiming to test CLIL and non CLIL learners' reading and listening skills in English (the same test before and after intervention)	9 months, 1 ac. year	non-CLIL not statistically significant
Pladevall-Ballester 2016	Spain	5th-6th grades	138/149	Cambridge young learners test	20 months, 2 ac. years	non-CLI; p=0,3521

The figure and table below summarise the data from the included studies that presented data about the influence of CLIL on development of reading skills.

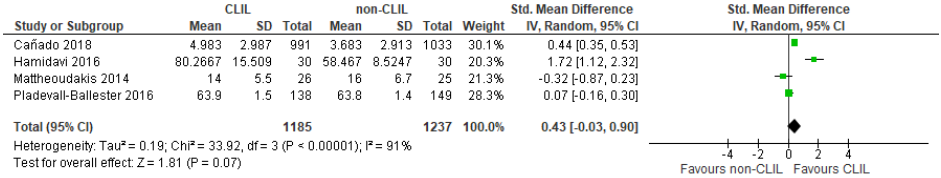
Hamidavi, et al (2016) in their research realised experiment with the group of students (n=60) divided into high achievers and low-achiever. The sample was equally divided into 4 subgroups (15 CLIL high achievers, 15 non-CLIL high achievers and 15 CLIL low achievers, 15 non-CLIL low achievers). The pre-tests were not statistically different. For the need of the study the groups (low achievers and high achievers) were combined. Compared to other studies the length of the research was relatively short (10 sessions of treatment). The authors mention

images and pictures that enrich CLIL texts as a possible important factor that may positively influence the effect of CLIL.

The study conducted by Mattheoudakis, et al (2014), similarly as in the listening evaluation, presented the data where non-CLIL students outperformed CLIL students. The recorded difference was not statistically significant (see the forest plot and CI). Similarly, data presented from the Pladevall-Ballester and Valbona (2016) were not statistically significant and the results of the control and experimental groups were remarkably close.

Statistically significant differences were observed in two studies, Hamidavi 2016 and Cañado 2018.

Fig. 3: Forest plot illustrating the results in Reading using random effect model



Tab. 5: Statistics for the studies and summary

Study ID	Hedges's g and 95% CI						
	Hedges' g	Stnd. error	Variance	Lower limit	Upper limit	Z-Value	p-Value
Cañado, 2018	0,441	0,045	0,002	0,352	0,529	9,794	0,000*
Hamidavi, 2016	1,719	0,299	0,090	1,133	2,306	5,745	0,000*
Mattheoudakis, 2014	-0,322	0,278	0,077	-0,866	0,222	-1,159	0,246
Pladevall, Ballester, 2016	0,069	0,118	0,014	-0,162	0,300	0,584	0,559
Random model	0,435	0,240	0,058	-0,035	0,905	1,812	0,070

The table above shows that research results have been contradictory. Two studies have shown statistically significant positive effects. Two other studies have shown statistically non-significant effects, in one case negative effect was observed.

The summary result suggests the students in the experimental group demonstrated larger gains in reading comprehension than the non-CLIL group in the regular classroom (see also the effect size Hegde's g = 0,43 what is a medium effect). Still, the progress in reading comprehension made by the intervention group was not statistically

significant (p = 0,07; 95% CI[-0,03; 0,9]). The heterogeneity (I^2) is, similarly as in the part Listening, very high (91%) what means we deal with considerable heterogeneity.

4.3 Vocabulary

Four studies out of those included dealt with learning vocabulary. Three of them applied the same 2k Vocabulary level test (Schmitt, Schmitt & Clapham, 2001).

Tab. 6: Details of included studies – part Vocabulary

Study ID	country	sample grade (age)	no of CLIL/non CLIL	research design / instrument	length	favor; effect size; p
VOCABULARY						
Agustín-Llach 2014	Spain	4th grade PE (9-10) 5th grade PE (10-11) 6th grade PE (11-12)	58/49	2k Vocabulary Levels Test (VLT)	3 years	p=0,024*
Cañado 2018	Spain	6th grade PE (11-12) 4th grade SE (15-16)	1033/991	Specifically designed English language tests (validated for the study); comprised the use of English, vocabulary, reading, writing and speaking parts with a total score of 100 points	4 ac. years	CLIL; -0,619; p<0,001* CLIL; -0,940; p<0,001*
Castellano-Risco 2020	Spain	3rd grade SE (14-15)	82/56	2k VLT	CLIL applied differently	CLIL; d=2,07; p<0,0004*
Castro-García 2017	Costa Rica	11th grade (16,6 avg age)	55/30	2k VLT	CLIL applied since the 7th grade	CLIL; p<0,001*

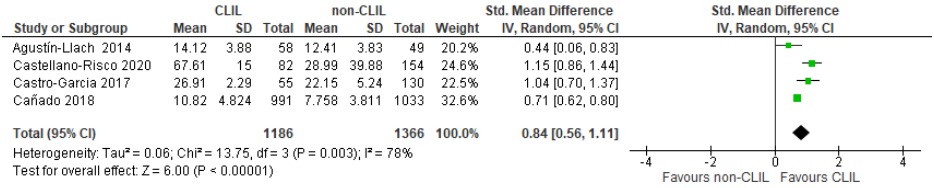
Castro Garcia (2017) in his research used 2K vocabulary level test. Students in his sample were slightly older than those in other studies. The results show that CLIL students (n=55) mastered more words than those in EFL groups (n=130). The context for CLIL application in content-based schools in the sample is „CB school students have received approximately 1,368 hours, and the mainstream EFL school students have received approximately 1,140 hours of class. ... The hours of instruction mentioned above for the CB school include 3 hours a week in a one-subject course that varies from one level to the next: Ecology, Social Studies, and Biology in 7th, 8th, and 9th school years, respectively“. Terrazas Gallego and Agustín Llach's (2009) state that there is an increase of passive vocabulary gain in time, i.e. “students move up a grade and become more proficient in the foreign language, they show receptive knowledge of significantly more words than years before“.

In the included study Llach and Alonso (2016) used also 2k VLT. The research lasted for 3 years and they studied the learners from the 4th till 6th grade who in their last year have received 944 hours of exposure to English in CLIL groups compared to 629 hours in non-CLIL groups. Both groups received approximately

105-110 hours of English years since the 1st grade and CLIL groups received extra hours in CLIL science. The study presents the results for 3 years successively. The research results from third year of the research were used in the present study. The results indicate the increase of vocabulary from the year to the next one, all vocabulary gains are in favour of CLIL groups. The inferential statistics researchers present shows there was not a statistically significant difference between the groups in the 4th grade, however, it became significant in the 5th grade and even more in the 6th grade. The authors also focussed their attention on the progression differences and they state the CLIL learners show slightly higher growth rates compared to the non-CLIL, however, there are "no significant differences in the number of words incorporated to the lexicons of CLIL and non-CLIL learners" (Agustin-Llach, Canga-Alonso, 2016). The authors stress the possible significance of time factor that can influence students' receptive vocabulary acquisition.

Vocabulary Levels Test was also used in the study of Castellano-Risco et al (2020) who divided their sample (n=138) into four strands, CLIL1 (n=23, EFL hours 1300, CLIL lesson 1700, CLIL from the first grade, 2-3 content subjects), CLIL2 (n=25, EFL hours 2400, CLIL lesson 1300), CLIL3 (n=34, EFL hours 1300, CLIL lesson 700, started with the CLIL at the SE) and EFL (n=56, EFL hours 1200). The CLIL learners in the study almost doubled EFL learner's knowledge of non-academic vocabulary.

Fig. 4: Forest plot illustrating the results in Vocabulary using random effect model



Tab. 7: Statistics for the studies and summary - part vocabulary

Study ID	Hedges's g and 95% CI						
	Hedges's g	Stnd. error	Variance	Lower limit	Upper limit	Z-Value	p-Value
Agustín-Llach, 2014	0,440	0,195	0,038	0,058	0,822	2,257	0,024*
Castellano-Risco, 2020	1,819	0,204	0,042	1,418	2,219	8,906	0,000*
Castro-Garcia, 2017	1,037	0,169	0,029	0,706	1,368	6,135	0,000*

Cañado, 2018	0,706	0,046	0,002	0,616	0,796	15,407	0,000*
Random model	0,985	0,238	0,057	0,519	1,452	4,140	0,000*

When comparing CLIL substituted for a traditional monolingual education (figure 4, table 6), there were significant enhancements in vocabulary gains observed (none of the 95% confidence intervals of the studies overlap 0). In case of vocabulary gains, all included studies presented statistically significant differences in favour of CLIL at level p<0,0001. Overall estimate effect is 0,84 with the CI [0,56; 1,11]. Concerning heterogeneity, I^2 is relatively high (78%) what means we deal with substantial heterogeneity.

5 Discussion

Three aspects of language performance were the subject of the present analysis of the selected CLIL studies, listening, reading and vocabulary. Results of nine studies were synthesised and analysed within three individual subgroups. There were 3517 participants in 5 studies with the focus on listening (1794 participants in CLIL and 1723 students in non-CLIL groups). The random-effect model was applied and the weight of studies ranged from 19,2% to 20,4%. The effect size of the Pladevall-Ballester 2016 was relatively high (-2,66) compared to the other 4 studies. It was the only study where the observed results favoured the non CLIL group (statistically significant difference). In the applied model the results of three studies were not statistically significant, similarly as the summary result. The confidence interval was [-1,13; 0,85] what confirms the null hypothesis that there is no statistically significant difference between the CLIL and non-CLIL groups. These results are on agreement with e.g Pérez-Vidal and H. Roquet (2015), Nieto-Moreno-de-Diezmas (2016).

Data from 4 studies focused on reading (n=2422) were synthesised. Similarly, as in the previous subgroup, there was one study where the result favoured non-CLIL group (Mattheoudakis 2014). The highest weight had Cañado 2016 study with the results favouring CLIL (with the effect size 0,44). Two of the studies brought non statistically significant results similarly as the summary result with the CI [-0,03; 0,9] and p=0,07. This finding is consistent with that of e.g. Nieto-Moreno-de-Diezmas (2016).

The vocabulary subgroup (n=2552) is the only one where the statistically significant difference was observed (p=0,003). All studies included into the analysis presented statistically significant differences in favour of CLIL. The effect size of the summary result is 0,84 and the CI ranges from 0,56 to1,11.

Generally, a low number of studies where synthesised and the heterogeneity was very high (99%, 91% and 78% successively). Due to a small number of studies,

the publication bias was not discussed. Identifying and accesing e.g. the unpublished studies, theses might be one of the ways for further objectivisation of the data synthesised.

Limitations

Meta-analysis is an observational study of selected studies. The method is considered to be very useful as it allows to synthesise data from different (even small samples where the results can even be from various reasons not statistically significant). On the other hand, there are aspects that can be understood or perceived as threat, risks or drawbacks. Not all the studies are realised in the same conditions and do not control all the effects. The selection of the studies can also be understood as a limitation as "some studies have not been published, or have been published in a form to which the researcher has no access, or have been published in a language that the researcher cannot read, etcetera" (Hak, Rhee, & Suurmond, 2016). The authors (ibid) also mention the problem with probability sampling, missing cases, the problem with pre-test and post-design and test differences.

The important discussion is about the possibility to combine and estimate the different outcomes that measure the same concept. Latest trends are to use not more than two different instruments.

Due to high heterogeneity the random-effects model was applied. It is important to realise that if compared to fixed-effect model, "the random effects model may apply too much weight to small studies, which are often poorly done and biased" (Schroll, et al, 2011). Applying the fixed-effect model the results would be a statistically significant difference in favour of CLIL groups in all three parts – listening, reading and vocabulary.

Model	Effect size and 95% confidence interval						Test of null (2-Tail)		Heterogeneity			
Model	Number Studies	Point estimate	Standard error	Variance	Lower limit	Upper limit	Z-value	P-value	Q-value	df (Q)	P-value	I-squared
Listening												
Fixed	5	0,254	0,035	0,001	0,185	0,323	7,184	0,000	553,584	4,000	0,000	99,277
Random effects	**5**	**-0,138**	**0,507**	**0,257**	**-1,132**	**0,856**	**-0,271**	**0,786**				
Reading												
Fixed	4	0,403	0,041	0,002	0,322	0,484	9,784	0,000	34,890	3,000	0,000	91,402
Random effects	**4**	**0,438**	**0,241**	**0,058**	**-0,035**	**0,911**	**1,815**	**0,070**				
Vocabulary												
Fixed	4	0,762	0,042	0,002	0,679	0,844	18,049	0,000	33,646	3,000	0,000	91,084
Random effects	**4**	**0,985**	**0,238**	**0,057**	**0,519**	**1,452**	**4,140**	**0,000**				

We also have to mention publication bias that was not estimated in the present study as the number of studies was low. We realise that the unpublished studies were omitted similarly as we searched only WOS databases to ensure the quality

of studies. However, this can also mean that we missed important data that can significantly influence the summary result.

Conclusion

Goris et al (2019) realised a systematic review of longitudinal studies and it is evident how much research is done in the field. However, as we deal with the process of education, we deal with many factors that are difficult to control and thus the results differ, similarly as the interpretation.

One of the crucial moments is the problem with the strict definition of CLIL methodology what successively influence the research and what is labelled as CLIL. It would be useful to conduct the studies that would carefully reflect the methodology and the effects of language but also the content gains. Similarly, it seems it might be useful to focus the attention on the extramural exposure to foreign language and the efficiency of CLIL.

The results of the present meta-analysis show that with respect to language learning skills the application of CLIL significantly increases the gains in foreign language vocabulary. The important factors that affect the results can be the more intensive exposure to a foreign language, meaningful context, association with visual or multimodal context. Still, it is necessary to compare those results with the impact on content learning.

As the key actors of CLIL have been identified, teachers, school management and parents. Here we should also mention how important teacher preparation is and it should be present as soon as in pre-service preparation (Sepešiová, 2019) and it should be subject not only of language but also content preparation. The significance of setting the aims and planning lessons for the CLIL success is indisputable, similarly as teacher training. Whatmore, as Graddol (2006) explains CLIL teachers must be in a position to "convey not only the subject content and disciplinary language but also the practical problem-solving, negotiations, discussions and classroom management in ways that characterise disciplinary pedagogic practices"(p. 86). This should be supported by school management and done systematically as the research indicates the success is influenced by time and intensity of exposure. Still, it must be mentioned, it is one side of a coin. The results must be carefully evaluated along with the data about the content subject knowledge and gains.

Acknowledgements

I wish to express my sincere gratitude and appreciation to María Luisa Pérez Cañado who willingly shared data from her research.

This study is a part of a research project financed by the Slovak Ministry of Education, Science, Research and Sport (KEGA 032PU-4/2019, project: Creating teaching materials for content and language integrated learning at elementary schools).

References

Agudo, J. D. (2019). Which instructional programme (EFL or CLIL) results in better oral communicative competence? Updated empirical evidence from a monolingual context. *Linguistics and Education*, 51, 69-78. doi:10.1016/j.linged.2019.04.008

Agustín-Llach, M. P., & Canga Alonso, 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. doi:10.1111/ijal.12090

Borenstein, M., Hedges, L. V., Higgins, J. P., & Rothstein, H. R. (2009). *Introduction meta-analysis*. Chichester: John Wiley & Sons.

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

Canga Alonso, A. (2013). The receptive vocabulary of Spanish 6th-grade primary-school students in CLIL instruction: A preliminary study. *Latin American Journal of Content and Language Integrated Learning, 6*(2), 22-41. doi:10.5294/laclil.2013.6.2.2

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. doi:10.1016/j.system.2020.102263

Castro-García, D. (2017). Receptive vocabulary measures for EFL Costa Rican high school students. *International Journal of English Studies, 17*(2), 81-99. doi:10.6018/ijes/2017/2/265681

Cenoz, J. (2015). Content-based instruction and content and language integrated learning: The same or different? *Language, Culture and Curriculum, 28*(1), 8-24. doi:10.1080/07908318.2014.1000922

Chostelidou, D., & Griva, E. (2014). Measuring the Effect of Implementing CLIL in Higher Education: An Experimental Research Project. *Procedia - Social and Behavioral Sciences, 116*, 2169-2174. doi:10.1016/j.sbspro.2014.01.538

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

Dallinger, S., Jonkmann, K., Hollm, J., & Fiege, C. (2016). The effect of content and language integrated learning on students' English and history competences –

Killing two birds with one stone? *Learning and Instruction, 41*, 23-31. doi:10.1016/j.learninstruc.2015.09.003

European Commission/EACEA/Eurydice. (2017). *Key Data on Teaching Languages at School in Europe – 2017 Edition. Eurydice Report*. Luxembourg: Publications Office of the European Union. doi:10.2797/839825

Fontecha, A. F., & Canga Alonso, A. (2014). A preliminary study on motivation and gender in CLIC and non- CLIL types of instruction. *International Journal of English Studies, 14*(1), 21-36. doi:10.6018/ijes/14/1/156681

Gené-Gil, M., Juan-Garau, M., & Salazar-Noguera, J. (2015). Development of EFL writing over three years in secondary education: CLIL and non-CLIL settings. *The Language Learning Journal, 43*(3), 286-303. doi:10.1080/09571736.2015.1053278

Glen, S. (2016). Hedges' g: Definition, Formula. From *StatisticsHowTo.com: Elementary Statistics for the rest of us!* Retrieved August 08, 2020, from https://www.statisticshowto.com/hedges-g/

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. doi:10.1177/1474904119872426

Gorjian, B., & Hamidavi, N. (2017). Using Clil Method in Teaching Vocabulary to Intermediate Efl Learners. *Modern Journal of Language Teaching Methods, 7*(3), 13-23.

Hak, T., Rhee, H. V., & Suurmond, R. (2016, updated 2018). How to Interpret Results of Meta-Analysis. *SSRN Electronic Journal*. doi:10.2139/ssrn.3241367, Available at SSRN: https://ssrn.com/abstract=3241367

Hamidavi, N., Amiz, M., & Gorjian, B. (2016). The Effect of CLIL Method on Teaching Reading Comprehension to Junior High School Students. *Modern Journal of Language Teaching Methods*, 6(9), 64-73. doi:10.26655/mjltm.2016.12.1

Hamidavi, N., Amiz, M., & Gorjian, B. (2016). The Effect of Clil Method on Teaching Reading Comprehension to Junior High School Students. *Bulletin De La Société Royale Des Sciences De Liège, 85*, 1642-165.

Harrer, M., Cuijpers, P., Furukawa, T., & Ebert, D. (2019). Doing Meta-Analysis in R: A Hands-on Guide. Retrieved August 08, 2020, from https://bookdown.org/MathiasHarrer/Doing_Meta_Analysis_in_R/, doi:10.5281/zenodo.2551803

Isidro, X. S., & Lasagabaster, D. (2018a). Code-switching in a CLIL multilingual setting: A longitudinal qualitative study. *International Journal of Multilingualism, 16*(3), 336-356. doi:10.1080/14790718.2018.1477781

Isidro, X. S., & Lasagabaster, D. (2018b). The impact of CLIL on pluriliteracy development and content learning in a rural multilingual setting: A longitudinal study. *Language Teaching Research, 23*(5), 584-602. doi:10.1177/1362168817754103

Kovalik, S., & Olsen, K. (1993). *ITI: The model: Integrated thematic instruction*. United States: S. Kovalik & Associates.

Lasagabaster, D., & Doiz, A. (2015). A Longitudinal Study on the Impact of CLIL on Affective Factors. *Applied Linguistics*. doi:10.1093/applin/amv059

Lasagabaster, D., & Sierra, J. M. (2009). Immersion and CLIL in English: More differences than similarities. *ELT Journal, 64*(4), 367-375. doi:10.1093/elt/ccp082

Lo, Y. Y., & Jeong, H. (2018). Impact of genre-based pedagogy on students' academic literacy development in Content and Language Integrated Learning (CLIL). *Linguistics and Education, 47*, 36-46. doi:10.1016/j.linged.2018.08.001

Madrid, D., & Barrios, E. (2018). A Comparison of Students' Educational Achievement across Programmes and School Types with and without CLIL Provision. *Porta Linguarum: Revista Internacional De Didáctica De Las Lenguas Extranjeras,* (29), 29-50.

Mattheoudakis, M., Alexiou, T., & Laskaridou, C. (2014). To CLIL or Not to CLIL? The Case of the 3rd Experimental Primary School in Evosmos. *Major Trends in Theoretical and Applied Linguistics Volume 3*. doi:10.2478/9788376560915.p13

Mensel, L. V., Hiligsmann, P., Mettewie, 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. doi:10.1080/07908318.2019.1571078

Merino, J. A., & Lasagabaster, D. (2017). 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. doi:10.1111/ijal.12177

Moghadam, N. Z., & Fatemipour, H. (2014). The Effect of CLIL on Vocabulary Development by Iranian Secondary School EFL Learners. *Procedia - Social and Behavioral Sciences, 98*, 2004-2009. doi:10.1016/j.sbspro.2014.03.635

Morton, T., & Llinares, A. (2016). Students' use of evaluative language in L2 English to talk and write about history in a bilingual education programme. *International Journal of Bilingual Education and Bilingualism, 21*(4), 496-508. doi:10.1080/13670050.2016.1192101

Myskow, G., & Ono, M. (2018). A matter of facts: L2 writers' use of evidence and evaluation in biographical essays. *Journal of Second Language Writing, 41*, 55-70. doi:10.1016/j.jslw.2018.08.002

Nieto Moreno de Diezmas, E. (2016). The impact of CLIL on the acquisition of L2 competences and skills in primary education. *International Journal of English Studies, 16*(2), 81-101. https://doi.org/10.6018/ijes/2016/2/239611

Nieto-Moreno de Diezmas, E. (2018a). The Acquisition of L2 Listening Comprehension Skills In Primary And Secondary Education Settings: A Comparison Between Clil And Non-Clil Student Performance. *RLA. Revista De Lingüística Teórica Y Aplicada, 56*(2), 13-34. doi:10.4067/s0718-48832018000200013

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. doi: http://dx.doi.org/10.18239/ocnos_2018.17.1.1471

Nikula, T. (2016). CLIL: A European Approach to Bilingual Education. *Second and Foreign Language Education,* 1-14. doi:10.1007/978-3-319-02323-6_10-1

Ortega, L. (2015). Researching CLIL and TBLT interfaces. *System, 54*, 103-109. doi:10.1016/j.system.2015.09.002

Otwinowska, A., & Foryś, M. (2015). They learn the CLIL way, but do they like it? Affectivity and cognition in upper-primary CLIL classes. *International Journal of Bilingual Education and Bilingualism, 20*(5), 457-480. doi:10.1080/13670050.2015.1051944

Pastrana, A., Llinares, A., & Pascual, I. (2017). Students' language use for co-construction of knowledge in CLIL group-work activities: A comparison with L1 settings. *Zeitschrift Für Erziehungswissenschaft, 21*(1), 49-70. doi:10.1007/s11618-017-0802-y

Pérez-Cañado, M. & Basse, R. (2015). Analysing Errors of CLIL and non-CLIL Primary School Students in their Written and Oral Productions: A Comparative Study. *Procedia - Social and Behavioral Sciences, 173*, 11-17. doi: 10.1016/j.sbspro.2015.02.023

Pérez-Cañado, M. L. (2017). The effects of CLIL on L1 and content learning: Updated empirical evidence from monolingual contexts. *Learning and Instruction, 57*, 18-33. doi: 10.1016/j.learninstruc.2017.12.002

Pérez-Cañado, M. L. (2018). CLIL and Educational Level: A Longitudinal Study on the Impact of CLIL on Language Outcomes. *Porta Linguarum, 29*(January), 51–70.

Pérez-Cañado, M. L. (2018a). CLIL and pedagogical innovation: Fact or fiction? *International Journal of Applied Linguistics, 28*(3), 369-390. doi:10.1111/ijal.12208

Pérez-Vidal, C., & Roquet, H. (2015). The linguistic impact of a CLIL Science programme: An analysis measuring relative gains. *System, 54*, 80-90. doi: 10.1016/j.system.2015.05.004

Pladevall-Ballester, E., & Vallbona, A. (2016). CLIL in minimal input contexts: A longitudinal study of primary school learners' receptive skills. *System, 58*, 37-48. doi: 10.1016/j.system.2016.02.009

Pokrivčáková, S. (2015). CLIL in Slovakia: Projects, research, and teacher training (2005-2015*). CLIL in Foreign Language Education: E-textbook for Foreign Language Teachers*, 17-29. doi:10.17846/clil.2015.17-29

Prieto-Arranz J.I., Rallo Fabra L., Calafat-Ripoll C., Catrain-González M. (2015) Testing Progress on Receptive Skills in CLIL and Non-CLIL Contexts. In: *Juan-Garau M., Salazar-Noguera J. (eds) Content-based Language Learning in Multilingual Educational Environments. Educational Linguistics*, vol 23. 123-137, doi:10.1007/978-3-319-11496-5_8

Roquet, H., Llopis, J., & Pérez-Vidal, C. (2015). Does gender have an impact on the potential benefits learners may achieve in two contexts compared: Formal instruction and formal instruction + content and language integrated learning? *International Journal of Bilingual Education and Bilingualism, 19*(4), 370-386. doi:10.1080/13670050.2014.992389

Salamanca, C., & Montoya, S. I. (2018). Using CLIL Approach to Improve English Language in a Colombian Higher Educational Institution. *English Language Teaching, 11*(11), 19. doi:10.5539/elt.v11n11p19

Schmitt, N., Schmitt, D. & Clapham, C. (2001). Developing and exploring the behavior of two new versions of the Vocabulary Level Test. *Language Testing, 18*(1), 55–88.

Schroll, J. B., Moustgaard, R., & Gøtzsche, P. C. (2011). Dealing with substantial heterogeneity in Cochrane reviews. Cross-sectional study. *BMC Medical Research Methodology, 11*(1), 22. doi:10.1186/1471-2288-11-22

Sepešiová, M. (2019). New Perspectives in Pre-Service Teacher Training in CLIL via VLE. In C. Can, P. Patsala, & Z. Tatsioka (Eds.), *Contemporary Means and Methods in ELT and Applied Lingustics* (pp. 515-533). Tallin, Estonia: LIF - Language in Focus.

Straková, Z. (2020). CLIL and Global Education: A Meaningful Match. *SOCIETY. INTEGRATION. EDUCATION. Proceedings of the International Scientific Conference, 5*, 546-557. doi:10.17770/sie2020vol5.4819

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

Sylvén, L. K. (2013). CLIL in Sweden – why does it not work? A metaperspective on CLIL across contexts in Europe. *International Journal of Bilingual Education and Bilingualism, 16*(3), 301-320. doi:10.1080/13670050.2013.777387

Taillefer, G. (2013). CLIL in higher education: The (perfect?) crossroads of ESP and didactic reflection. *ASp,* (63), 31-53. doi:10.4000/asp.3290

Terrazas Gallego, M. & Agustín Llach, M. P. (2009). Exploring the increase of receptive vocabulary knowledge in the foreign language: A longitudinal study. *International Journal of English Studies, 9*(1), 113–133, Available at https://revistas.um.es/ijes/article/view/90681/87481

Tzoannopoulou, M. (2015). Rethinking ESP: Integrating Content and Language in the University Classroom. *Procedia - Social and Behavioral Sciences, 173*, 149-153. doi:10.1016/j.sbspro.2015.02.045

Yang, Wen-hsien. (2016). ESP vs. CLIL: A coin of two sides or a continuum of two extremes*?. ESP Today*. *4*(1). 43-68.

Zarobe, Y. R. (2008). CLIL and Foreign Language Learning: A Longitudinal Study in the Basque Country. *International CLIL Research Journal, 1*(1), 60-73.

Zarobe, Y. R. (2017). Improving reading strategy knowledge in young children: What self-report questionnaires can reveal. *Elia,* (17), 15-45. doi:10.12795/elia.2017.i17.02

Contact

Ivana Cimermanová
Faculty of Arts
University of Presov
ivana.cimermanova@unipo.sk
17. novembra 1
08001 Prešov
Slovakia