The Journal of Language Teaching and Learning™
2021 Volume 11/Issue 1 Article 3

An integrative research study on the impact of CLIL programs on writing skills: The case of age and exposure duration

Özgür Çelik, School of Foreign Languages, Balıkesir University, Balıkesir, Turkey,
ozgurcelik@balikesir.edu.tr

Recommended Citations:

APA
Çelik, Ö. (2021). An integrative research study on the impact of CLIL programs on writing skills: The case of age and exposure duration. *The Journal of Language Teaching and Learning*, 11(1), 29-39

MLA
Özgür Çelik. "An integrative research study on the impact of CLIL programs on writing skills: The case of age and exposure duration." The Journal of Language Teaching and Learning 11.1 (2021): 29-39.

www.jltl.org

The Journal of Language Teaching and Learning, 2021(*1*), pp.29-39

An integrative research study on the impact of CLIL programs on writing skills: The case of age and exposure duration

Özgür Çelik[1]

ARTICLE INFO

Article History:
Received 09 April, 2020
Revisions completed 07 December, 2020
Published 1 January, 2021

Key Words:
Writing skills
CLIL
Age
Exposure duration
Integrative research

ABSTRACT

Language is at the core of Content and Language Integrated Learning (CLIL) programs along with culture, environment, content, and learning (Marsh, Maljers & Hartiala, 2001). Among other linguistic skills, writing has a unique role and purpose in that it acts as a bridge between content and language. Thus, it is crucial to reveal all aspects of the interactive relationship between writing skills and CLIL programs. In this respect, this study set out to conduct an integrative research study on the impact of CLIL programs on writing skills. Rather than a holistic approach, this study adopted a component analysis that emerged in the studies found, and the independent variables were refined to age and exposure duration. Before the data collection process, specific criteria were set for the studies: the studies would be empirical with a comparison group (CLIL vs Non-CLIL). The focus of the studies would be on the components of writing skills, and the studies would include the age of the participants and the exposure duration to CLIL. Within this framework, 15 studies were found, and the results were analysed. The results indicated that age is not a definite determiner in overall writing production except for lexical complexity. Also, longitudinal exposure to CLIL was found to be slightly more effective than in non-longitudinal studies. Similarly, lexical complexity was improved better in a non-CLIL setting for longitudinal studies.

Improving writing skills has always been one of the focal points of language learning research. A keyword search on Google Scholar (search date:10.05.2019) reveals that the keyword "writing skills" yields 320.000 results after "reading skills" (326.000) and followed by "listening skills" (121.000) and "speaking skills" (60.000). Accordingly, writing skills are prominent in Content and Language Integrated Learning (CLIL) programs in that writing activities act as a bridge between content and the language. However, there is no consensus on the effect of CLIL programs on writing skills in the literature. While some studies find evidence for the success of CLIL programs to improve writing skills (Gené-Gil et al., 2015; Pérez-Vidal & Roquet, 2015; Lahuerta, 2017), some indicate no effect (Roquet & Pérez-Vidal, 2015; de Zarobe, 2010;

[1]School of Foreign Languages, Balıkesir University, Balıkesir, Turkey, ozgurcelik@balikesir.edu.tr

 ISSN: 2146-1732

Vázquez, 2014). Moreover, Dalton-Puffer's (2008) review study unveils the fact that writing skills are less likely to improve in CLIL settings. This suggests that the relationship between CLIL and writing skills requires a more in-depth analysis focusing on sub-components of writing and the features of CLIL programs because both constructs (writing skills and CLIL programs) affect, and are affected, by one another.

Writing is not a unidimensional skill but has multiple sub-components such as complexity, accuracy, fluency, vocabulary, organisation, grammar, and so on. These components constitute the writing skill as a whole, and they do not follow linear progress and may address different age groups and different levels. For example, complexity may not be the case for beginner learners but is a criterion to be achieved for advanced learners. Similarly, certain components of writing such as fluency, accuracy or complexity do not develop in parallel for all age groups (Torras and Celaya, 2001). When assessing writing, a component evaluation would make sense rather than the assessment of writing as a whole. In this respect, this study aims to make integrative research of the experimental studies on the effect of CLIL programs on improving writing skills by addressing the age of the participants and the duration of the intervention with a special focusing on sub-components of writing skills. In order to collect data, some refinements were applied. First, experimental studies with CLIL and Non-CLIL groups were preferred so that the results would be based on empirical data. Second, the studies that focus on the components of writing were selected and third, the studies in which age and treatment duration are independent variables were addressed.

This research is guided by the following research questions:

1. Does the success of CLIL programs in writing skills change according to the age of the participants?
2. Is the duration of intervention a factor affecting the success of CLIL programs in writing skills?

2. CLIL and Writing Skills

CLIL emerged as an innovative way of teaching several converging dimensions, among which are culture, environment, language, content, and learning (Marsh, Maljers & Hartiala, 2001). Language is the prominent dimension along with the content in CLIL programs. From a linguistics perspective, writing has a vital role in CLIL programs in that content knowledge is largely gained through writing activities (Papaja, 2014). Thus, writing skills deserve a more comprehensive approach in CLIL settings.

In the literature, rather than a holistic approach, a component analysis was adopted to analyse writing skills in CLIL programs. These analyses are made in two categories: quantitative analysis (complexity, accuracy, fluency) and qualitative analysis (Task fulfilment, organisation, grammar, vocabulary). Complexity is measured in two ways, as syntactic complexity and lexical complexity. Syntactic complexity refers to how varied and complex the production units or grammatical structures (Wolfe-Quintero et al., 1998). Although there are several different ways to measure the syntactic complexity, the featured method is the coordination index which is measured by dividing the independent clause coordination by the number of combined clauses (Wolfe-Quintero et al., 1998). The mean values of this index can be between 0 and 1. A lower score means higher competence. Lexical complexity is related to the lexical richness of the texts and in most cases measured by dividing the total amount of lexis to the total amount of words or dividing the total amount of lexis by the square root of the total amount of tokens. The former is known as Type/Token Ration (Johnson, 1944), and the latter as Guiraud's index (Guiraud, 1954).

To eliminate the validity problems with these methods, Malvern and Richards (2002) proposed a new way of measuring lexical complexity by using a random selection of tokens to produce a curve for type/token ratio. Another quantitative measurement method is accuracy, which is defined as 'the conformity of second language knowledge to target language norms' by Wolfe-Quintero et al. (1998, p.4). In a general sense, accuracy is measured by identifying the errors per word, but many different accuracy measurement methods were proposed, some of which are morphological errors per clause, sentence or T-

unit; syntactic errors per clause, sentence or T-unit or by dividing the correct number of words in a category to the total number of words (Roquet, 2011). The last quantitative measurement is fluency that refers to more words or structures accessed in a limited time (Wolfe-Quintero et al., 1998). One common way of measuring written fluency is to calculate the number of words in a specified time. The higher the number of words used in a limited time, the more fluent the writer is.

As for the qualitative measures, Friedl & Auer (2007) suggested four types: task fulfilment, organisation, grammar, and vocabulary. Each component is measured with a scoring scale from one to five. A higher score indicates a higher competency. Task fulfilment aims to measure to what extent the task is fully achieved. Specific points to be covered in task fulfilment are content, relevance, text format, length, and register. A competent text in the task fulfilment is described as "Task fully achieved, content entirely relevant; appropriate format, length and register." Organisation assessment is made with a special focus on structure, paragraphing, cohesion and coherence, editing and punctuation. The scale identifies a sound organisation as "clear overall structure, meaningful paragraphing; very good use of connectives, no editing mistakes, conventions of punctuation observed." For the grammar assessment, errors, the variety of structures, and readiness to use complex structures were measured. Other than the quantitative measurement, grammar assessment in this model is based on human checking. Five points are given if the text has "accurate use of grammar and structures, hardly any errors of agreement, tense, word order, articles, pronouns, etc.; meaning clear, a great variety of structures, frequent use of complex structures." Vocabulary measurement comprises the analysis of the range and choice of words, accuracy, spelling, and comprehensibility. For the vocabulary analysis, the text is expected to have a "wide range of vocabulary; very good choice of words; accurate form and usage; hardly any spelling mistakes; meaning clear."

3. Methodology

3.1. Data Collection Procedure

This study is based on the premises of integrative research, which aims to make a synthesis of empirical findings into a coherent whole (Cooper, 1982). According to Cooper (1982), the focus of integrative research is twofold: (1) to replace papers that have fallen behind the research front (Price, 1965) and (2) direct future research so that it yields a maximum amount of new information. It can be understood that when compared to the literature review, integrative research is more sensitive in data collection in that it requires empirical and ultimately selective data to yield new information. For this reason, the data collection procedure was handled with the utmost care and followed specific steps. In the first step, certain characteristics of the studies to be found were drawn as the studies would be empirical with a comparison group (CLIL vs Non-CLIL). The focus on the studies would be on the components of writing skill. Finally, the studies would include the age of the participants and the duration of the intervention. Within this framework, the databases of the following sources were inspected; Google Scholar, Web of Science, Sage Journals, and Taylor and Francis. In the second step, the aforementioned databases were searched by using the keywords "CLIL, writing skills." The potential studies were pinned, and a new search was made with "Content and Language Integrated Learning, writing skills" keywords. This search procedure was made in all databases mentioned. After the searching process, the studies were skimmed to see if they were appropriate for this study. All the related studies (N = 23) were listed. All the studies found were in line with the characteristics identified above. After this step, for the purpose of ensuring the validity of the data, a new criterion was added as opting-out the studies that may be published in predatory journals. To achieve this, the studies that were published in the journals not indexed by SSCI, AHCI, SCOPUS, ERIC, Web of Science were opted out. After this refinement, 15 studies remained. The publication years of the selected studies range from 2010 to 2018.

3.2. ***Data Analysis***

Having determined the studies to be evaluated (N = 15), data were prepared for analysis. First, the age range or the age mean of the participants in the selected studies was identified. A similar process was conducted for the duration of the intervention. Also, which components of writing the studies focused on was listed along with the main findings of the studies with respect to the components.

The age and duration of treatment were set as independent variables. For the age group, the following coding scheme was adopted, and studies were categorised into three groups: age < 13 = young learner; age between 13- 18 = teenage learners; age > 18 = adult learners. However, no studies that comply with the criteria were found for adult learners. A similar coding was made for the duration variable as duration < one term or year = Non-longitudinal; duration > one term or year = Longitudinal. For the next step, the emerging components of writing skills in the selected studies were identified: Lexical Complexity, Syntactic Complexity, Accuracy, Fluency, Task Fulfilment, Organization, Vocabulary, Content, Mechanics, Spelling. For this variable, symbol-coding was used. The following symbols were added to each component based on the effect of CLIL programs: ☑ symbol indicates that there is an improvement; ☒ symbol indicates no improvement; ↑ means higher than the other group, and ↓ means lower than the other group (See Table 2). Then a scoring table was formed, and the studies were scored according to the scoring table.

Table 1
The Scoring Table

Point	Assigned to
0	the studies that showed no improvement in the writing skills
1	the studies that showed an improvement, but mean scores were lower than the other group
2	the studies that showed an improvement and mean scores were higher than the other group

The scoring and the descriptive analysis were made with SPSS (Version 21) software. An overall analysis of the selected studies is presented in Table 2.

Table 2
The Overall Analysis of the Selected Studies

	Study	Age Group	Duration	Focus	
				CLIL	Non-CLIL
1	Gené-Gil, Juan-Garau & Salazar-Noguera (2015a)	Teenage Learners	Longitudinal	Lexical Complexity☑↓ Syntactic Complexity☑↑ Accuracy☑↓ Fluency ☑↑	Lexical Complexity☑↑ Syntactic Complexity☑↓ Accuracy☑↑ Fluency ☑↓
2	Roquet & Pérez-Vidal (2015)	Teenage Learners	Longitudinal	Syntactic Complexity☑↑ Lexical Complexity☑↓ Accuracy☑↑ Fluency☒ Task Fulfilment☑↑ Organization☑↑ Vocabulary☑↑	Syntactic Complexity☑↓ Lexical Complexity☑↑ Accuracy☑↓ Fluency☒ Task Fulfilment☑↓ Organization☑↓ Vocabulary☑↓
3	Pérez-Vidal & Roquet (2015)	Teenage Learners	Longitudinal	Fluency☒ Accuracy☑↑ Syntactic Complexity☑ Lexical Complexity☑↓ Task Fulfilment☑↑ Organization☑↑ Vocabulary☑↑	Fluency☒ Accuracy☑↓ Syntactic Complexity☒ Lexical Complexity☑↑ Task Fulfilment☑↓ Organization☑↓ Vocabulary☑↓
4	de Zarobe (2010)	Teenage Learners	Longitudinal	Content☑↑ Organization ☑↑ Vocabulary ☑↑ Accuracy ☑↑ Mechanics☑↑	Content☑↓ Organization ☑↓ Vocabulary ☑↓ Accuracy ☑↓ Mechanics☑↓
5	Olsson (2015)	Teenage Learners	Longitudinal	Vocabulary ☑↑	Vocabulary ☑↓
6	Gené-Gil, Juan-Garau & Salazar-Noguera (2015b)	Teenage Learners	Longitudinal	Syntactic Complexity☑↑ Lexical Complexity☑↑ Accuracy☑↑	Syntactic Complexity☑↓ Lexical Complexity☑↓ Accuracy☑↓

Table 2 (Continued)

7	Lahuerta (2017)	Teenage Learners	Non-Longitudinal	Accuracy☑↑	Accuracy☑↓
8	Jexenflicker & Dalton-Puffer (2010)	Teenage Learners	Non-Longitudinal	Task Fulfilment☑↑ Organisation☑↑ Accuracy☑ Vocabulary☑↑	Task Fulfilment☑↓ Organisation☑↓ Accuracy☒ Vocabulary☑↓
9	Klampfl (2010)	Teenage Learners	Non-Longitudinal	Lexical Complexity☑↑ Syntactic Complexity☑↑ Vocabulary☑↑ Spelling☑↓ Accuracy☑↑	Lexical Complexity☑↓ Syntactic Complexity☑↓ Vocabulary☑↓ Spelling☑↑ Accuracy☑↓
10	Corral-Robles & González-Gijón (2018)	Teenage Learners	Non-Longitudinal	Fluency☑↑ Accuracy☑↑ Syntactic Complexity☑↑ Lexical Complexity☑↑	Fluency☑↓ Accuracy☑↓ Syntactic Complexity☑↓ Lexical Complexity☑↓
11	Lasagabaster (2011)	Teenage Learners	Non-Longitudinal	Accuracy☑↑	Accuracy☑↓
12	Vázquez (2014)	Young Learners	Longitudinal	Accuracy☑↑	Accuracy☑↓
13	De Diezmas (2016)	Young Learners	Non-Longitudinal	Organization☑↑ Vocabulary☑↑ Fluency☑↑ Spelling☑↑ Accuracy☑↓	Organization☑↓ Vocabulary☑↓ Fluency☑↓ Spelling☑↓ Accuracy☑↑
14	Agustín Llach (2017)	Young Learners	Non-Longitudinal	Lexical Complexity☑↓ Spelling☑↑	Lexical Complexity☑↑ Spelling☑↓
15	Pérez & Basse (2015)	Young Learners	Non-Longitudinal	Accuracy ☑↑	Accuracy ☑↓

**☑ symbol indicates that there is an improvement; ☒ symbol indicates no improvement; ↑ means higher than the other group; and ↓ means lower than the other gro*

5. Results

In the data analysis section, 15 studies were reviewed and classified based on the variables of this study (age, duration, and components of writing). The classification revealed that of 15 studies, six of them were longitudinal, and five of them were non-longitudinal studies on teenage learners. There were four studies on young learners, three of which were non-longitudinal, and one was longitudinal. The data were analysed for each research question.

RQ1- Does the success of CLIL programs in writing skills change according to the age of the participants?

For the age variable, two age groups (young and teenage) were formed, and the studies were classified based on these groups. The classification yielded 11 studies for teenage learners, four studies for young learners. The studies were scored based on the scoring table (See Table 1), and the results are presented in Table 3.

Table 3
Mean Scores of the Components of Writing Based on Age

	Young Learners		*Teenage Learners*	
	CLIL	*Non-CLIL*	*CLIL*	*Non-CLIL*
	M	*M*	*M*	*M*
Lexical Complexity	1	2	1.50	1.50
Syntactic Complexity	-	-	2	0.83
Accuracy	1.67	1.33	1.89	1
Fluency	2	1	1	0.50
Task Fulfilment	-	-	2	1
Organization	2	1	2	1
Vocabulary	2	1	2	1
Content	-	-	2	1
Mechanics	-	-	2	1
Spelling	2	1	1	2
Total	10.67	7.33	17.39	10.83

The overall table suggests that CLIL programs are effective in improving writing skills for young and teenage learners. With a more detailed inspection, specific conclusions can be drawn. For the young learners, CLIL programs are effective in accuracy, fluency organisation, vocabulary, and spelling while Non-CLIL groups outperform the CLIL groups in lexical complexity. For teenage learners, apart from spelling and lexical complexity, all other writing skills were found to be improved in CLIL groups. It can be inferred that lexical complexity is the writing skill that is affected much by the age of the learners. Additionally, Table 3 reveals the need for more research on the effect of CLIL programs for young learners with a special focus on syntactic complexity, task fulfilment, content and mechanics which have never been studied before.

RQ-2. Is the duration of the intervention a factor affecting the success of CLIL programs in writing skills?

The duration of the intervention was analysed under two categories; non-longitudinal studies and longitudinal studies. The studies whose treatment lasted for one term or less than one year were labelled as non-longitudinal studies while longitudinal studies comprise those the treatment of which lasts for more than one year. Bearing on this categorisation, eight studies were labelled as longitudinal and seven studies as non-longitudinal. Table 4 demonstrates the mean scores of writing components based on the duration of the treatment.

Table 4
The Mean Scores of the Components Based on the Duration of Treatment

	Non-Longitudinal		*Longitudinal*	
	CLIL	*Non-CLIL*	*CLIL*	*Non-CLIL*
	M	*M*	*M*	*M*
Lexical Complexity	1.67	1.33	1.25	1.75
Syntactic Complexity	2	1	2	0.75
Accuracy	1.83	1	1.83	1.17
Fluency	2	1	0.67	0.33
Task Fulfilment	2	1	2	1
Organization	2	1	2	1
Vocabulary	2	1	2	1
Content	-	-	2	1
Mechanics	-	-	2	1
Spelling	1.67	1.33	-	-
Total	15.17	8.66	15.75	9

Table 4 offers strong evidence that CLIL programs are effective in writing skills regardless of the duration of treatment. However, longitudinal programs are slightly more effective than non-longitudinal programs (M_L = 15.75 > M_{NL} = 15.17). In non-longitudinal programs, all components of writing improved in CLIL groups while in longitudinal studies, all programs apart from lexical complexity improved better in CLIL groups than Non-CLIL ones.

6. Discussion

The results of this study provided significant takeaways regarding the relationship between CLIL programs and writing skills with a special focus on the age of the students and exposure duration to CLIL instruction. From the overall perspective, the findings of the 15 studies suggest that CLIL programs are effective in improving the writing skills of the learners. The age of the participants is not a strong determiner in the success of CLIL programs on writing skills from a holistic perspective, but for lexical complexity, Non-CLIL groups have a higher mean score than the CLIL group for younger learners and the equal mean score for teenage learners. Roquet's (2011) robust study revealed that young CLIL learners are only better at syntactic complexity but also adds that when the exposure hours were kept constant, an older age counterbalances the positive effects of CLIL programs due to the role of cognitive development. On the other hand, Lasagabaster (2008) hypothesised that younger CLIL learners (14-15) would catch up with older (16) Non-CLIL learners in language skills. His findings suggested that younger learners not only caught up but also surpassed the older Non-CLIL learners. Similarly, Navés and Victori (2010) concurred a similar finding that younger CLIL learners outperformed the older (2 years ahead) Non-CLIL learners. However, it is worth considering that the age difference in the studies of Lasagabaster (2008) and Navés and Victori (2010) is two years maximum which may not be a sufficient interval to make a judgement on the effect of age in CLIL programs.

The second focus of this study was on the effect of treatment duration in CLIL programs. The overall results indicated a slightly better effect of longitudinal treatments on writing skills than the Non-Longitudinal ones. Again, lexical complexity acted differently from the other skills, and Non-CLIL groups were better in lexical complexity in longitudinal treatments. Zarobe (2010) reported a similar result in his study claiming that when the hours of exposure were kept constant, the Non-CLIL group outperformed the CLIL group in lexical complexity. The study of Gené-Gil, Juan-Garau & Salazar-Noguera (2015a) reached a

similar conclusion, and they reported that when the exposure hours were kept constant Non-CLIL group showed a considerable increase in accuracy compared to the CLIL group. It can be inferred that the duration of treatment in CLIL programs should be inspected at the syntactic (complexity and accuracy) level.

7. **Conclusion**

This study set out to make integrative research on the effect of CLIL programs on writing skills in general and the role of age and duration of treatment in particular. To do so, an extensive review of the literature was made by confining the studies to experimental ones with comparison groups (CLIL vs Non-CLIL) which focus on components of writing skills. After the first round, some refinements were made to ensure the validity of the data by opting out the studies that are not indexed by SSCI, AHCI, SCOPUS, ERIC, and Web of Science. Following this stage, the emerging components in the studies were identified, and the studies were classified based on the age (young and teenage learners) and the duration of treatment (Non-Longitudinal and longitudinal studies). At the final stage, the studies were scored based on the scoring table generated, and the results were analysed and interpreted.

This study attempted to analyse the studies showing the effect of CLIL programs on writing skills in a systematic way and draw some conclusions. First, the age of the participants is not a strong determiner in the success of CLIL programs on overall writing production. However, complexity and accuracy are found to be affected by the age of the participants, and this also indicates the need for more research on the relationship between age, complexity, and accuracy in CLIL programs. Second, longitudinal treatment programs yield more effective results for overall writing production than the non-longitudinal programs.

Again, accuracy and complexity domains deserve a particular focus in further studies.

In conclusion, CLIL programs and writing skills have an interactive relationship. When designing writing tasks or instruction in CLIL programs, it will be useful to consider the components of writing, age group and the duration of the program in order to exploit the program effect to the full.

References

Agustín Llach, M. P. (2017). The effects of the CLIL approach in young foreign language learners' lexical profiles. *International Journal of Bilingual Education and Bilingualism*, *20*(5), 557-573. https://doi.org/10.1080/13670050.2015.1103208

Cooper, H. M. (1982). Scientific guidelines for conducting integrative research reviews. *Review of educational research*, *52*(2), 291-302. https://doi.org/10.3102/00346543052002291

Corral-Robles, S., & González-Gijón, G. (2018). Effect of CLIL and non-CLIL approaches on the written competence of upper-secondary students. *International Journal of Pedagogy and Curriculum*, *25*(2), 31-43. doi:10.18848/2327-7963/CGP/v25i02/31-43 https://doi.org/10.18848/2327-7963/cgp/v25i02/31-43

Dalton-Puffer, C. (2008). Outcomes and processes in Content and Language Integrated Learning (CLIL): current research from Europe. *Future perspectives for English language teaching* (pp. 139-158). Heidelberg: Winter.

De Diezmas, E. N. M. (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

de Zarobe, Y. R. (2010). Written production and CLIL: An empirical study. *Language use and language learning in CLIL classrooms*, 191-212. https://doi.org/10.1075/aals.7.10rui

Friedl, G., & Auer, M. (2007). Rating scale used for assessment of the writing task. *Erläuterungen zur Novellierung der Reifeprufungsverordnung fur AHS, lebende Fremdsprachen*.

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

Gené-Gil, M., Juan-Garau, M., & Salazar-Noguera, J. (2015b). Writing development under CLIL provision. In *Content-based language learning in multilingual educational environments* (pp. 139-161). Springer, Cham. https://doi.org/10.1007/978-3-319-11496-5_9

Guiraud, P. (1954). Les Charactères Statistiques du Vocabulaire. Essai de méthodologie. Paris: Presses Universitaires de France.

Jexenflicker, S., & Dalton-Puffer, C. (2010). Comparing the writing of CLIL and non-CLIL students in higher colleges of technology. *Language use and language learning in CLIL classrooms*, *7*, 169. https://doi.org/10.1075/aals.7.09jex

Johnson, W. (1944). Studies in language behavior: A program of research. *Psychological Monographs*, *56*(2), 1-15.

Klampfl, A. (2010). *A comparative study of writing proficiency between an Austrian CLIL and mainstream EFL class with regard to vocabulary* (Doctoral dissertation, uniwien).

Lahuerta, A. (2017). Analysis of accuracy in the writing of EFL students enrolled on CLIL and non-CLIL programmes: the impact of grade and gender. *The Language Learning Journal*, 48(2), 121–132. https://doi.org/10.1080/09571736.2017.1303745

Lasagabaster, D. (2008). Foreign language competence in content and language integrated courses. *The Open Applied Linguistics Journal*, 1(1), 30–41. https://doi.org/10.2174/1874913500801010030

Lasagabaster, D. (2011). English achievement and student motivation in CLIL and EFL settings. *Innovation in language Learning and Teaching*, *5*(1), 3-18. https://doi.org/10.1080/17501229.2010.519030

Loranc-Paszylk, B. (2009). Enhancing the effectiveness of CLIL through writing. In D. Newby & H. Penz, Languages for social cohesion (pp. 125-131). European Centre for Modern Languages of the Council of Europe.

Malvern, D., & Richards, B. (2002). Investigating accommodation in language proficiency interviews using a new measure of lexical diversity. *Language testing*, 19(1), 85-104. https://doi.org/10.1191/0265532202lt221oa

Marsh, D., Maljers, A., & Hartiala, A. K. (2001). *Profiling European CLIL Classrooms: Languages Open Doors*, European Platform for Dutch Education.

Navés, T., & Victori, M. (2010). CLIL in Catalonia: An overview of research studies. In *CLIL in Spain: Implementation, results and teacher training* (pp. 30-54). Cambridge Scholars Publishing.

Olsson, E. (2015). Progress in English academic vocabulary use in writing among CLIL and non-CLIL students in Sweden. *Moderna språk*, *109*(2), 51-74.

Papaja, K. (2014). *Focus on CLIL: A Qualitative Evaluation of Content and Language Integrated Learning (CLIL) in Polish Secondary Education* (p. 33). Newcastle: Cambridge Scholars Publishing.

Pérez, M. C., & Basse, R. (2015a). 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. https://doi.org/10.1016/j.sbspro.2015.02.023

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

Price, D. J. D. S. (1965). Networks of scientific papers. *Science*, 149(3683), 510–515. https://doi.org/10.1126/science.149.3683.510

Roquet i Pugès, H. (2011). *A study of the acquisition of English as a foreign language: integrating content and language in mainstream education in Barcelona*. Universitat Pompeu Fabra.

Roquet, H., & Pérez-Vidal, C. (2015b). Do productive skills improve in content and language integrated learning contexts? The case of writing. *Applied Linguistics*, *38*(4), 489-511. https://doi.org/10.1093/applin/amv050

Torras, M. R., & Celaya, M. L. (2001). Age-related differences in the development of written production. An empirical study of EFL school learners. *International Journal of English Studies*, *1*(2), 103-126.

Vázquez, B. M. (2014). Lexical transfer in the written production of a CLIL group and a non-CLIL group. *International Journal of English Studies*, *14*(2), 57-76. https://doi.org/10.6018/j.166251

Wolfe-Quintero, K., Inagaki, S., Kim, H. Y., Kim, H. Y., & Inagaki, S. (1998). *Second language development in writing: Measures of fluency, accuracy, & complexity* (p. 187). Honolulu: University of Hawai'i Press.