Studies in Second Language Learning and Teaching
Department of English Studies, Faculty of Pedagogy and Fine Arts, Adam Mickiewicz University, Kalisz
SSLLT 13 (2). 2023. 373-397
https://doi.org/10.14746/ssllt.38279
http://pressto.amu.edu.pl/index.php/ssllt

The impact of EMI on student English writing proficiency in a Spanish undergraduate engineering context

María del Mar Sánchez-Pérez ✉
University of Almería, Spain
https://orcid.org/0000-0002-6550-9030
mmar.sanchez@ual.es

Abstract

This study analyzes the influence of English medium instruction (EMI) on student English writing development in a Spanish undergraduate context, a language skill that remains underexplored in EMI research at tertiary level. Through a longitudinal pre- and post-test method, it explores the writing progress of EMI engineering students as compared to the achievement of a group of English as a foreign language (EFL) learners, according to different quantitative and qualitative writing measures. Results show a positive impact of EMI on student writing development at the levels of lexical accuracy and vocabulary. However, other writing areas, such as syntax, grammar, organization, or fluency, appeared unaffected. These findings challenge, to some extent, the widespread assumption that EMI contributes to improving student English language proficiency. This study suggests that sole exposure to the language in EMI contexts does not suffice to improve students' English writing abilities beyond the area of vocabulary. Therefore, it argues for the inclusion of language skill support within EMI programs to provide students with opportunities to achieve a comprehensive development of their English language and writing competence.

Keywords: English medium instruction (EMI); integrating content and language in higher education (ICLHE); language outcomes; writing; proficiency

1. Introduction

The number of English medium instruction (EMI) programs has undergone an unprecedented rise in the last two decades as a result of the internationalization process experienced by higher education institutions worldwide. The rationale behind the implementation of EMI programs includes the desire to attract international students to enhance the international profile and thus the prestige of the institutions, to increase student and faculty international mobility to foster international collaboration and competitiveness, and to provide future graduates with opportunities to improve their English language skills to help them operate successfully in an increasingly global labor market (Briggs et al., 2018; Galloway et al., 2017).

EMI is defined as an educational model in which the English language is used "to teach academic subjects (other than English itself) in countries or jurisdictions where the first language of the majority of the population is not English" (Macaro, 2018, p. 1). In EMI higher education programs, the primary and virtually exclusive goal is to learn disciplinary content. English language instruction is rarely considered explicitly in EMI courses (Dafouz & Smit, 2020; Pecorari & Malmstrom, 2018; Rose et al., 2021; Unterberger, 2018). This differs substantially from other similar educational environments in Europe where academic content is taught through the medium of a foreign (FL) or a second language (L2) with an explicit dual focus on content and language acquisition, such as, content and language integrated learning (CLIL) or integrating content and language in higher education (ICLHE).

While the pursuits for the implementation and development of EMI programs seem to be shared by the majority of EMI stakeholders, the actual effectiveness of this educational approach to achieve the aforementioned goals remains unconfirmed in the research literature, in particular, as far as student English language development is concerned (Briggs et al., 2018; Graham et al., 2018). Notwithstanding the increasing number of studies on EMI conducted in the last two decades, the swift propagation of EMI programs has preceded and outpaced empirical research on the real language skill gains that students achieve in EMI environments (Galloway et al., 2017). The present study intends to fill this gap by analyzing the impact of EMI on students' English language proficiency in a Spanish undergraduate engineering setting. In particular, it focuses on the development of writing, a language skill that has generally received little attention in EMI research, especially in Southern Europe. Through a longitudinal pre- and post-test analysis, this study explores the English writing progress of a group of EMI students as compared to the gains achieved by a group of students following a traditional English as a foreign language (EFL) course. This study may have potential implications for EMI stakeholders, curriculum planners, and university administrators by providing insights into the real effect of EMI on student language

skill development, a matter that is often assumed in these contexts (McKinley & Rose, 2022; Wilkinson, 2018;).

2. Literature review

The burgeoning of EMI programs at tertiary level has triggered a growing number of studies attempting to shed light on a varied range of topics. Research into stakeholders' attitudes toward EMI has attracted much scholarly attention (Aguilar & Rodriguez, 2012; Dearden & Macaro, 2016; Kamaşak et al., 2020; Lasagabaster, 2018; Tatzl, 2011). By and large, despite the numerous challenges and resistances reported in these studies, most lecturers and students display positive attitudes toward this educational model, and they are confident about its potential for student language development.

Notwithstanding the lack of provision of explicit language instruction in EMI settings (Pecorari & Malmström, 2018), student English language improvement stands as one of the estimated benefits when it comes to implementing EMI programs (Galloway et al., 2017; Galloway & Rose, 2021; Graham et al., 2018). Indeed, many students and faculty members report this expectation as one of the primary motivations to engage in EMI practices (Salaberri-Ramiro & Sánchez-Pérez, 2018; Tatzl, 2011). The assumption that English language competence improves as an indirect outcome of receiving instruction through English is widespread, especially among learners (Kamaşak et al., 2020; Wächter & Maiworm, 2014). Studying through the medium of English is perceived as a valuable opportunity for future graduates to improve their linguistic competence and, thus, to get access to better academic and/or professional opportunities (Ament & Pérez-Vidal, 2015). Nevertheless, the dearth of empirical research on language learning outcomes in EMI contexts and the contradictory data available to date have not confirmed the actual effectiveness of EMI for the development of student English language abilities (Graham et al., 2018; Macaro et al., 2018).

Research into specific language abilities is rather limited in the EMI literature (Graham & Eslami, 2020). Writing in higher education is an essential skill through which most of the assessment takes place; thus it can be considered a vehicle for academic success (Sánchez-Pérez, 2021). Additionally, unlike other language skills such as reading, speaking and listening, students are less likely to develop academic writing out of the classroom, thus it is an appropriate skill via which to explore gains that can be more closely attributed to EMI. Notwithstanding the above, writing has generally received scant focus in EMI research, especially in Southern European countries such as Spain, the context of this study, where, unlike many universities from Northern Europe, academic writing teaching programs are infrequently found within academic curricula, especially in STEM areas (Dafouz, 2020).

Among the studies addressing writing outcomes in EMI contexts, most empirical investigations are centered on comparing achievement between writing and other language skills, and they do not offer detailed accounts of the development of this specific language ability. In general, these studies show a tendency toward advantages in the gains of receptive skills over productive skills (Aguilar & Muñoz, 2014; Ament & Perez-Vidal; 2015, Hernández & Jiménez; 2017). However, the holistic scores employed in these studies do not allow for capturing students' specific achievement in the development of their writing skills. Studies including different measures to analyze students' writing ability are, therefore, urgently needed, since learners may not necessarily develop all areas of the writing competence at the same level (Hughes, 2003; Jexenflicker & Dalton-Puffer, 2010).

Studies that examine students' writing gains according to different areas of writing in EMI higher education settings are even scarcer. Indeed, most of the research available to date in educational contexts where disciplinary content is taught through the medium of English is found in earlier educational phases, for example, in primary and secondary education, or in CLIL environments, where some focus on language learning is considered (Gené-Gil et al., 2015; Jexenflicker & Dalton-Puffer, 2010; Lasagabaster, 2008; Roquet & Pérez Vidal, 2015; Ruiz de Zarobe, 2010; Whittaker et al., 2011). Overall, these studies report CLIL advantages for writing development, especially at the levels of grammar and lexicon. As for the studies based on EMI higher education settings, only a few are found, yet yielding a variety of results. For example, in a study conducted in Thailand with undergraduate students of agriculture, Chansri and Wasanasomsithi (2016) reported that they improved their writing ability at the levels of vocabulary, language use, and mechanics when completing a one-semester EMI course. However, this study used only a one-group pre- and post-test analysis with a rather small sample size (N = 27), which challenges the reliability of the results. In the cross-sectional studies conducted by Salaberri-Ramiro and Sánchez-Pérez (2015) and by Sánchez-Pérez and Salaberri-Ramiro (2015) with Spanish EMI undergraduate students, whilst the authors did not explore the evolution over time, they found that participants showed writing deficiencies at the levels of pragmatics and discourse, but performed better as for lexicogrammar. In another study conducted in Spain, Vidal and Jarvis (2020) examined undergraduate students' English language proficiency and lexical diversity in English-medium writing after three years of instruction. They found higher writing holistic scores, but no significant gains in lexical diversity. However, the sample of this study pursued a degree in English Studies; thus there are serious doubts whether this can be considered an EMI learning context, since the purpose of the degree is to train future English language specialists and participants might have been influenced by further language-related training amid the course of the three-year investigation. In a more recent study conducted

in China, Zhang and Pladevall-Ballester (2021) found almost no progress in grammar proficiency but identified statistically significant gains in the three writing measures analyzed, namely, task achievement, general English vocabulary, and discipline-specific vocabulary, with higher gains being observed in the latter.

As can be observed, evidence about the effect of EMI on student writing development is scarce and inconclusive; thus there is a need for further research into the writing outcomes in EMI and, more specifically, into the specific writing areas that may be potentially affected by EMI instruction at tertiary level (Graham & Eslami, 2020). The present study aims to fill this gap by analyzing the English writing progress of two groups of undergraduate engineering students following an EMI and an EFL course, respectively, in a Spanish university setting. The differences between both groups' achievement will allow us to ascertain the impact of the EMI treatment on students' English writing proficiency as compared to another language learning context where development in English language and writing proficiency is expected. For this purpose, different measures of the writing ability are examined both quantitatively (complexity, accuracy, and fluency) and qualitatively (task fulfillment, organization, grammar, and vocabulary), which, to the author's knowledge, have not been considered jointly in previous research into English writing development in an EMI higher education context. With such aim in mind, this study addresses the following research question:

> Are there any differences in the development of the writing competence in English between undergraduate engineering students following an EMI and an EFL course, respectively, over the course of one semester?

3. Method

3.1. Setting and participants

This study was conducted at a state-run higher education institution in Southern Spain. Within its internationalization strategy, the university runs a multilingualism promotion plan through which different courses from diverse areas of knowledge are taught through the medium of English. Alternatively, in some areas, traditional EFL courses are offered for those students who do not opt for completing courses taught entirely in English, yet they want to improve their English language competence. The courses involved in this study belong to these two learning contexts, namely an EMI chemistry course and an EFL course. The EMI course was a compulsory[1] subject for

[1] While the course was compulsory according to the syllabus of the degrees involved, there were different groups taught in Spanish (L1) and in English (EMI). Student enrollment in the

students enrolled in chemical and agricultural engineering, whereas the EFL course was an elective module available for students enrolled in any science or engineering-related degree. Both were 6-ECTS courses (60 hours) that were taught during the fall semester of 2018-2019. The EMI course was taught by a Spanish-native content lecturer with more than 12 years of experience in EMI, whereas the EFL course was taught by a bilingual (English-Spanish) language specialist with more than 15 years of experience in EFL teaching. Students in the EFL group were also enrolled in the same chemistry (compulsory) course, but in a Spanish-medium taught group; thus both groups were expected to learn the same disciplinary content throughout the semester.

The sample of this study consisted of two groups of non-native English-speaking undergraduate engineering students following an EFL and an EMI course, respectively. A total of 129 students (82 males and 47 females) volunteered to participate in the study of a total population of 143 students enrolled officially in the course (92 males and 51 females). They were initially asked to sign a written consent and to complete a preliminary questionnaire about social, academic and language background details (i.e., gender, age, major, nationality, L1, certified English language level and exposure to English outside the classroom). In order to get a homogeneous sample regarding English proficiency level and as for language exposure in both groups, only students with a B1 (independent user) certified English language level according to the CEFR who reported not being enrolled in any other EMI or English language learning course that semester were selected for analysis. Hence, potential distorting results related to differences in language proficiency levels or to the amount of exposure to the language were minimized (Rose et al., 2019). This resulted in a final total sample of 98 students (62 males and 36 females), with ages ranging from 20 to 25, split into two independent groups, the EFL group (N = 51) and the EMI group (N = 47). As the setting of the study was a Spanish-speaking monolingual region where English is considered a foreign language (FL), exposure to the language and communication through English in both groups was restricted to the EFL and the EMI classrooms, respectively.

3.2. Data collection

To explore and compare the English writing progress achieved by both groups, students' written samples were elicited and gathered at two different points in time, at the beginning (pre-test) and at the end (post-test) of the respective courses. The writing tasks administered were the same for both groups. Compositions were

EMI group was made on a voluntary basis, as long as they certify a minimum B1 language level according to the *Common European Framework of Reference for Languages* (CEFR).

written in a classroom environment in an exam-like situation with a time limit of 30 minutes. In the pre-test, students were asked to describe a picture that showed a man conducting an experiment in a laboratory. In particular, they were asked to (1) describe the picture in as much detail as possible, and (2) explain the possible reason(s) why the man in the picture was carrying out such action. The writing task in the post-test was of a very similar nature, but in that case, the picture showed a team of scientists gathering samples from a crop. The writing tasks were, therefore, of a descriptive and argumentative nature, two genres commonly found in engineering-related areas (Parkinson, 2017). The topics were sufficiently general but related to the participants' area of knowledge to facilitate and inspire students' writing (Foster & Skehan, 1996).

3.3. Data analysis

Students' compositions were analyzed both quantitatively and qualitatively. The quantitative measures included complexity, accuracy, and fluency (CAF), three linguistic dimensions that have been largely considered in L2 writing performance (Hidalgo & Lázaro-Ibarrola, 2020; Norris & Ortega, 2009; Wolfe-Quintero et al., 1998). Regarding complexity, we included both syntactic and lexical complexity. To measure syntactic complexity, we used the coordination index (CI), obtained by dividing the "total number of independent clause coordination by the total number of combined clauses (clauses minus sentences)" (Wolfe-Quintero et al., 1998, p. 95). A decrease in the ratio of this measure indicates a lower use of coordinate clauses and, therefore, a higher use of subordinate and dependent clauses, which has been considered a signal of successful L2 writing development at the level of complexity in previous research (Ortega, 2003; Wolfe-Quintero et al., 1998). Lexical complexity was measured by means of the Guiraud's index (GI) of lexical richness, obtained by dividing the total number of lexical types by the square root of the total number of lexical tokens. As for accuracy, we used the error-per-word ratio. To obtain a comprehensive picture of this domain, different types of errors were considered, namely, grammatical, lexical and mechanical errors. Grammatical errors included students' inaccuracies concerning the choice of verb form and tense, subject-verb agreement, use of gerunds and infinitives, etc. Lexical errors included mistakes related to inadequate vocabulary choice, and mechanical errors included issues such as spelling or punctuation. Finally, fluency was measured by calculating the total number of words produced per minute, following the concept that more fluent writers produce longer pieces of writing within a certain time limit than less fluent counterparts, who tend to produce shorter and simpler ones (Norris & Ortega, 2009).

As for the qualitative measures, students' written texts were analyzed according to a rating scale adapted from Friedl and Auer (2007) considering four equally weighted aspects of the writing skill, namely, task fulfillment, organization, grammar, and vocabulary. Scores ranged according to a six-point Likert-type scale, from 0 (*not enough to evaluate*) to 5 (*very good*), in each category. Regarding task fulfillment, texts were assessed in terms of content relevance, format, length, and register. The dimension of organization focused on text structure, meaningful paragraphing, cohesion and coherence, use of connectives, editing and punctuation. The grammar category addressed morpho-syntax accuracy, variety of grammatical structures, appropriate use of subject-verb agreement, verb tense, word order, pronouns, articles and use of complex structures; and the vocabulary dimension addressed lexical variety, word choice, accurate form and usage, spelling mistakes and clear meaning.

3.4. Procedure

A total of eight writing measures were selected for analysis (Table 1). The measures were identified and coded in each student's written composition both in the pre-test and the post-test, and the indices, ratios, and scores for each measure were calculated. The mean results obtained per group in the pre-test underwent a series of independent-samples *t*-tests to detect potentially significant differences in the measures analyzed between the EFL and the EMI groups at the onset. Subsequently, the average findings per group in the pre- and post-tests were compared statistically with a mixed analysis of variance (ANOVA) to identify significant differences between both groups' achievements after completing their respective

Table 1 Quantitative and qualitative measures of analysis

Type of analysis	Measure	Description
Quantitative	Syntactic complexity	Coordination index (CI) = Independent clause coordination/ Combined clauses
	Lexical complexity	Guiraud's Index (GI) = WTypes/√Wtokens
	Accuracy	Total errors per word (E/W); Grammatical errors per word (GE/W); Lexical errors per word (LE/W); Mechanical errors per word (ME/W)
	Fluency	Words per minute (W/M)
Qualitative	Task fulfillment	Content relevance, format, length, and register
	Organization	Text structure, meaningful paragraphing, cohesion and coherence, use of connectives, editing and punctuation
	Grammar	Accuracy and variety of morpho-syntactic structures, subject-verb agreement, verb tense, word order, pronouns, articles, clear meaning, and complex structures
	Vocabulary	Lexical variety, word choice, accurate form and usage, spelling mistakes and clear meaning

courses. The measures that showed statistically significant differences between the EFL and the EMI group in the pre-test were not included in the ANOVA test, since they were not comparable and the differences in the gains between both groups might not be attributable to the treatments (EFL or EMI) (Larson-Hall, 2010). All measures were doubly coded and rated by the author of this paper and by an additional language expert, reaching an average inter-rater concordance of 86.9%. Inter-rater disagreements were solved by consensus.

4. Results

The results are structured according to the research question and the types of measures. We first present the findings of the inter-group comparison for the quantitative measures, and then for the qualitative ones.

4.1. Inter-group comparison for the quantitative measures: Complexity, accuracy and fluency

Our research question was set out to identify the differences in the development of the writing competence in English between undergraduate engineering students following an EMI and an EFL course, respectively. For this purpose, we compared the mean values per group in each quantitative measure both in the pre-test and the post-test. Table 2 shows the results of the *t*-test for all the quantitative measures analyzed at the onset. As can be observed, there were no statistically significant differences between the groups in the pre-test ($p > .05$); thus the progress made by each group in all measures could be safely compared (Larson-Hall, 2010).

Table 2 Pre-test quantitative results for the EFL and EMI groups

Measure	EFL group (*N* = 51)	EMI group (*N* = 47)	*t* (*df*)	*p*
	M (*SD*)	*M* (*SD*)		
Syntactic complexity (CI)	.42 (1.01)	.34 (1.13)	1.16 (96)	.248
Lexical complexity (GI)	5.87 (1.18)	6.33 (1.26)	-1.26 (96)	.093
Accuracy (E/W)	.06 (.06)	.07 (.08)	-.49 (96)	.325
Fluency (W/M)	7.72 (.72)	7.96 (1.83)	-1.58 (96)	.130

The average results of the quantitative measures per group in the pre-test and the post-test were compared by means of a mixed ANOVA to explore whether the differences between the progress achieved by each group were significant (Table 3).

Table 3 Pre- and post-test quantitative results for the EFL and EMI groups

	EFL group (*N* = 51)		EMI group (*N* = 47)	
	Pre-test *M* (*SD*)	Post-test *M* (*SD*)	Pre-test *M* (*SD*)	Post-test *M* (*SD*)
Syntactic complexity (CI)	.42 (1.01)	.37 (.14)	.34 (1.13)	.33 (.91)
Lexical complexity (GI)	5.87 (1.18)	5.98 (1.01)	6.33 (1.26)	7.01 (1.09)
Accuracy (E/W)	.06 (.06)	.05 (.07)	.07 (.08)	.04 (.05)
Fluency (W/M)	7.72 (.76)	7.81 (1.31)	7.96 (1.83)	8.17 (1.55)

Regarding syntactic complexity, results show a tendency towards a decrease in the number of coordinate sentences between the pre-test and the post-test in both groups, indicating a higher use of dependent and subordinate clauses after completing their courses. However, progress was higher in the EFL group than in the EMI group, in which case, the gains were nearly imperceptible (Figure 1). The difference between the achievement of the EMI and the EFL group was not statistically significant [F (1,96) = 3.83, p = .327]; thus no advantage of EMI on student writing proficiency was observed for syntactic complexity.

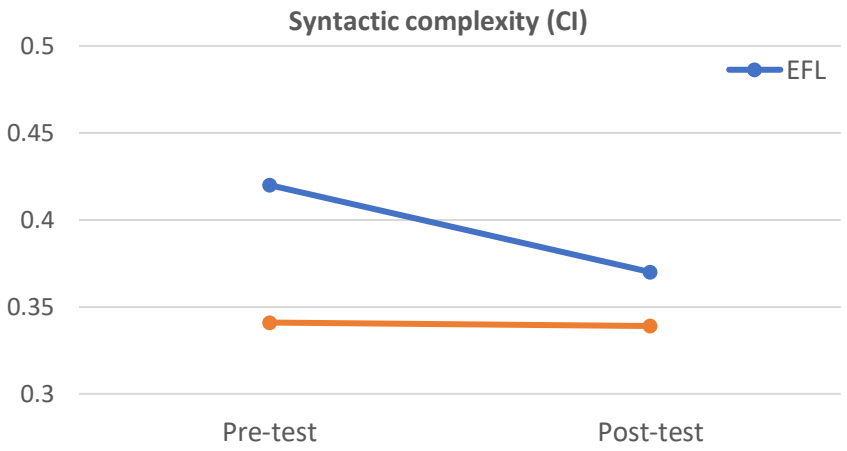

Figure 1 EFL and EMI progress in syntactic complexity (coordination index)

Regarding lexical complexity, results show a higher improvement of the EMI group as compared to the EFL group, which indicates that EMI students' written compositions contain richer and more varied vocabulary than the EFL ones' after the treatment (Figure 2). The mixed ANOVA revealed an almost significant difference between the EFL and the EMI group's gains (F [1,96] = 16.59, p = .061). Therefore, although the difference between both groups' achievement cannot be confirmed, findings show signs of a positive influence of EMI on students' lexical complexity in their writing.

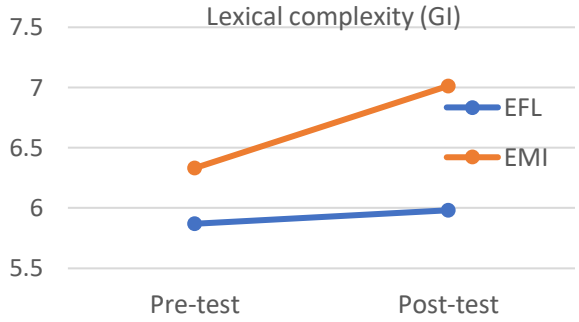

Figure 2 EFL and EMI progress in lexical complexity (Guiraud's index)

Concerning accuracy, the ratios of the overall measure (i.e., total number of errors per word) showed that, even though the EMI group had a higher error ratio than the EFL one at the onset, the former showed greater progress after completing the course (Figure 3). The difference in the global accuracy measure proved not to be statistically significant (F [1,96] = 4.09, p = .132). Nevertheless, the independent analysis per type of error showed statistically significant differences between the EFL and the EMI group in the sub-category of lexical accuracy (F [1, 96] = 1.46; p = .041), in which the EMI group outperformed the EFL one (Table 4). It is noteworthy that the EMI group had a higher lexical error-per-word (LE/W) ratio than the EFL group at the onset; thus it departed with a lower proficiency level than the EFL group for lexical accuracy. However, such initial advantage on the part of the EFL group was not paired with greater progress after completing the EFL course. On the contrary, the EMI group's gains were greater (Figure 4). The sub-domain of grammatical accuracy revealed a smaller progress of the EMI group than that of the EFL group (Figure 5), yet the difference was not statistically significant (F [1, 96] = .97; p = .242). As for mechanical errors (Figure 6), the EMI group showed slightly higher gains than the EFL group, although, again, this difference did not reach a statistically significant value (F [1, 96] = 1.54; p = .193).

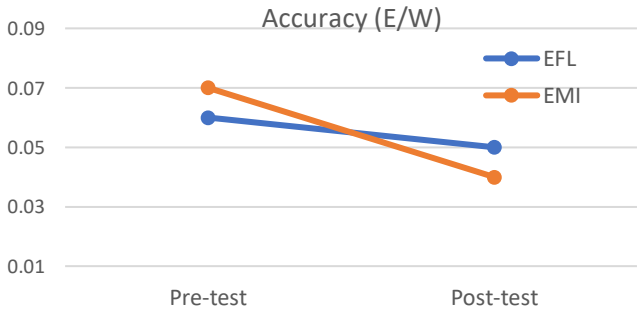

Figure 3 EFL and EMI progress in accuracy (E/W)

Table 4 Pre- and post-test measures of grammatical, lexical and mechanical accuracy for the EFL and EMI groups

	EFL group (*N* = 51)		EMI group (*N* = 47)	
	Pre-test *M* (*SD*)	Post-test *M* (*SD*)	Pre-test *M* (*SD*)	Post-test *M* (*SD*)
Grammatical accuracy (GE/W)	.045 (.08)	.023 (.07)	.057 (.09)	.046 (.08)
Lexical accuracy (LE/W)	.063 (.08)	.056 (.06)	.072 (.07)	.031* (.02)
Mechanical accuracy (ME/W)	.061 (.03)	.044 (.07)	.059 (.07)	.039 (.04)

Note. *Statistically significant difference between EFL and EMI group ($p < .05$)

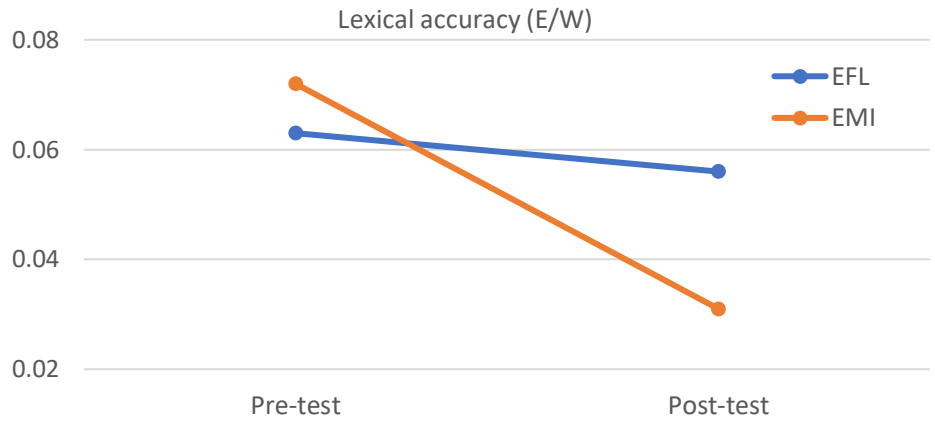

Figure 4 EFL and EMI progress in lexical accuracy (LE/W)

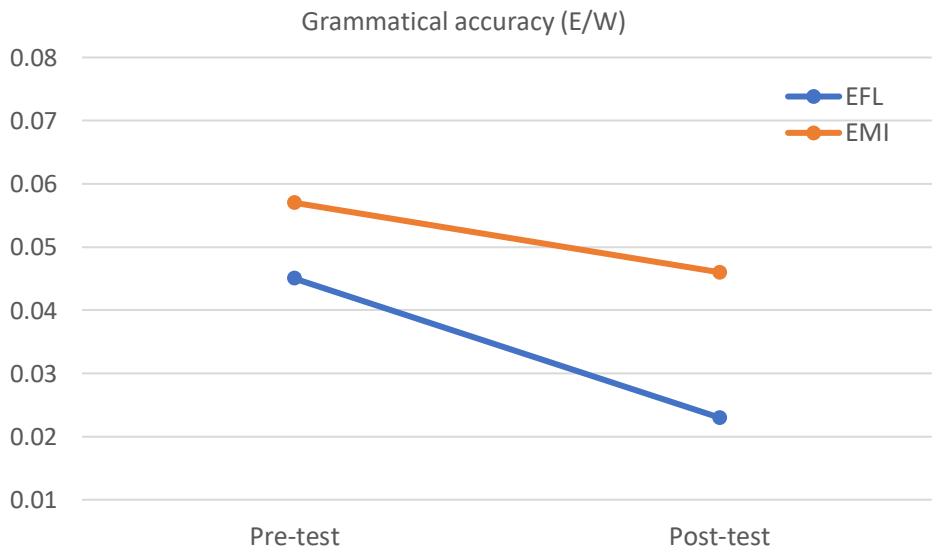

Figure 5 EFL and EMI progress in grammatical accuracy (GE/W)

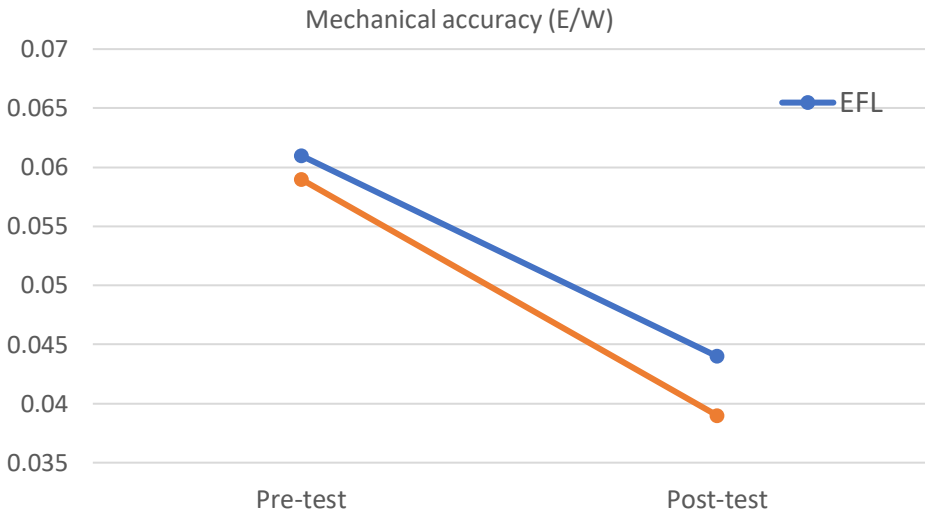

Figure 6 EFL and EMI progress in mechanical accuracy (ME/W)

Finally, regarding fluency, results showed a higher increase in this ratio in the EMI group (Figure 7), indicating that EMI students tend to produce longer pieces of writing, and thus, to be more fluent than their EFL counterparts. However, the difference between both groups' progress was not significant (F [1,96] = 16.59, p = .124). Therefore, the advantage of EMI for fluency cannot be confirmed.

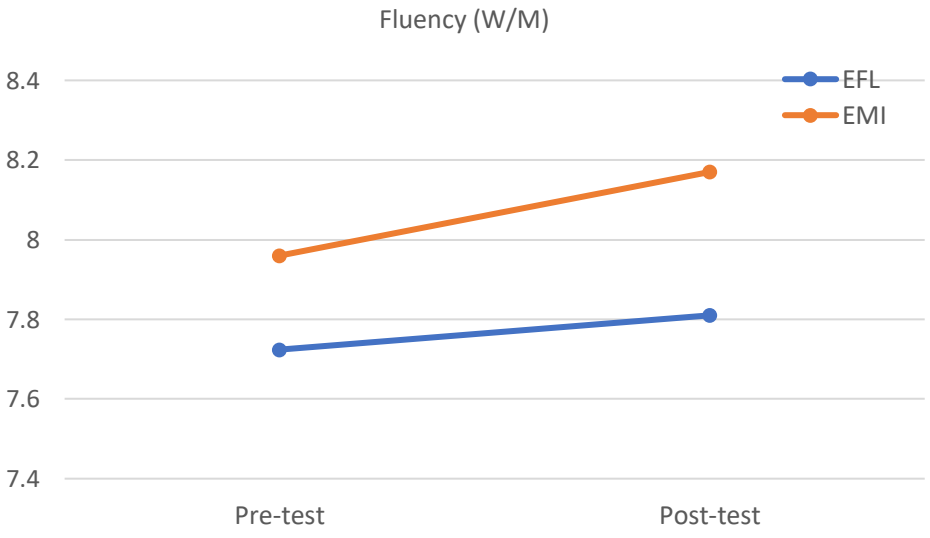

Figure 7 EFL and EMI progress in fluency (W/M)

4.2. Inter-group comparison for the qualitative measures: Task fulfillment, organization, grammar and vocabulary

As with the quantitative measures, independent-samples *t*-tests with the average scores obtained in each dimension per group in the pre-test were conducted to identify potentially significant differences between both groups at the onset (Table 5). The *t*-tests detected a significant difference between both groups in task fulfillment (t [96] = -2.68, p = .039). Therefore, this measure was not included in the subsequent ANOVA as the two groups departed from significantly different initial levels and the reliability of the data arising from the comparison of the results obtained in this category from the two groups after the treatment would be compromised (Larson-Hall, 2010).

Table 5 Pre- and post-test qualitative results for the EFL and EMI groups

Measure	EFL group (N = 51)	EMI group (N = 47)	t (df)	p
	M (SD)	M (SD)		
Task fulfillment	2.31 (.82)	2.69 (1.64)	-2.68 (96)	.039*
Organization	2.45 (.94)	2.47 (.76)	-.34 (96)	.152
Grammar	2.76 (1.51)	2.68 (1.73)	.97 (96)	.484
Vocabulary	2.64 (1.78)	2.71 (1.61)	-.56 (96)	.397

Note. * p < .05

The qualitative measures which did not show significant differences between the EFL and the EMI group in the pre-test underwent the mixed ANOVA analysis (Table 6).

Table 6 Pre- and post-test qualitative results for the EFL and EMI groups

Measure (code)	EFL group (N = 51)		EMI group (N = 47)	
	Pre-test M (SD)	Post-test M (SD)	Pre-test M (SD)	Post-test M (SD)
Organization	2.45 (.94)	2.67 (1.05)	2.47 (.76)	2.49 (1.09)
Grammar	2.76 (1.51)	3.02 (.97)	2.68 (1.73)	2.79 (1.14)
Vocabulary	2.64 (1.78)	2.83 (1.42)	2.71 (1.61)	3.39* (1.15)

Note. *Statistically significant difference between EFL and EMI group (p < .05)

With regard to organization, while both groups had a very similar average score at the onset, the results show a tendency towards higher progress in the EFL group, with the EMI group's progress remaining virtually unaffected (Figure 8). The difference between both groups' gains was not significant (F [1,96] = 2.987; p = .267). It is noteworthy that, unlike the EFL group, the average performance of the EMI group hardly reached the minimum passing score (2.50) according to the six-point rating scale (0-5) in the post-test, revealing the EMI group's

low level of writing performance as for text structure, paragraphing, use of connectives, editing and punctuation after the treatment. Unlike the EFL group, EMI students tended to write in a single paragraph and correlation between sentences was made mostly using simple connectives (*and*, *but*, *also*) both in the pre-test and in the post-test. Additionally, an overall low level of development regarding punctuation was observed both before and after the EMI treatment, as in "They are collecting seeds *and* controlling size *and* weight. *Also*[,] they are cleaning the seeds *and* checking the temperature" (EMI_23_post).

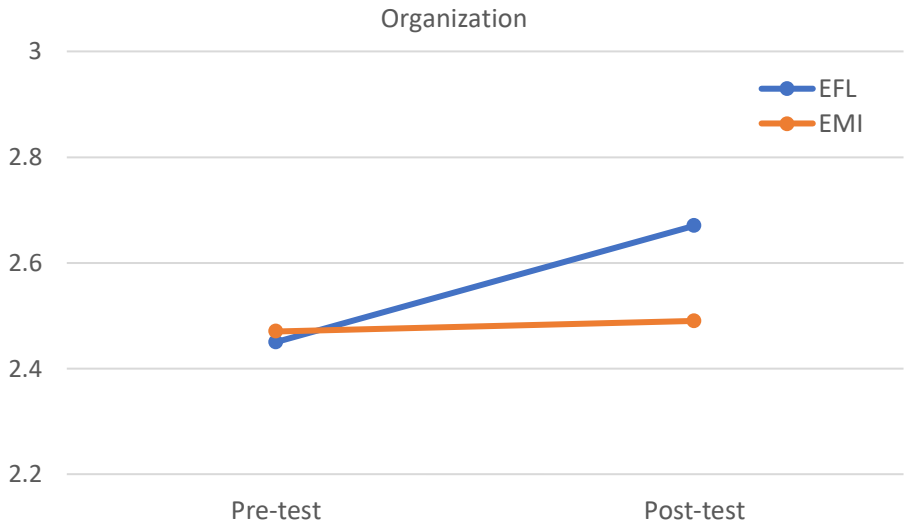

Figure 8 EFL and EMI progress in organization

Concerning grammar, our results point toward improvement in both groups after completing their respective courses, though a greater increase of the average score was found in the EFL group. While ANOVA did not detect statistically significant differences between both groups' progress (F [1,96] = .82; p = .361], the EFL group's compositions tended to include more varied and complex structures after the treatment, such as the use of the passive voice, e.g., "A cage with different seeds *can be observed*" (EFL_09_post). In contrast, the EMI group tended to use and repeat simpler grammar forms both in the pre-test and in the post-test, as in "The men *have* a white shirt, and a watch. They *also have* glasses and a mask" (EMI_47_post). Additionally, the EFL group committed fewer errors of verb tense, subject-verb agreement and word order after completing their course, which were more frequently found in the EMI group in the post-test, as in "The girl clean[*s*] [*is cleaning*] *carefully* the seeds with a [piece of] paper" (EMI_13_post).

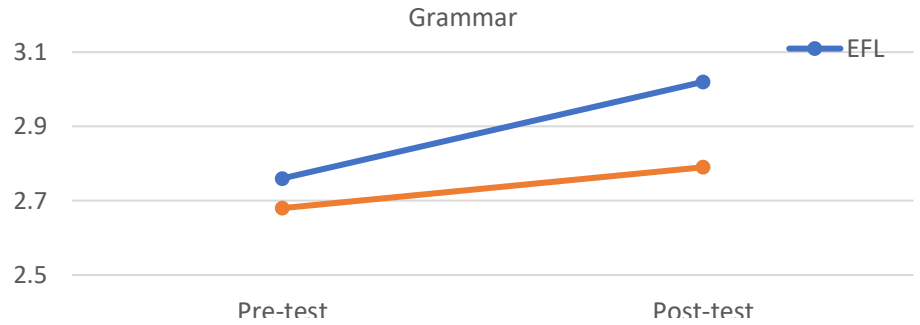

Figure 9 EFL and EMI progress in grammar

As for vocabulary, results also show signs of improvement in both learning environments. In this case, the progress made by the EMI group was remarkably higher than the one achieved by the EFL group. This difference turned out to be statistically significant (F [1,96]) = 24.60; p = .022). These findings reveal a positive influence of EMI on student writing development as for higher lexical variety, appropriate word choice, accurate form and usage, fewer spelling mistakes and clearer meaning. It is worthy of note that, while the majority of EMI students used simpler and less accurate vocabulary, including some spelling mistakes, in the pre-test, as in "*There is* a *man* in a laboratory doing an experiment with a buret[*t*]e and *other materials*" (EMI_24_pre), most of these issues were successfully solved after the treatment, including, in many cases, instances of more complex vocabulary and specialized terms, as in "*We can observe* some *scientists* taking samples from a *crop* . . . . On the table, we can see a *pipette* and a *test tube* with 10 *ml* of H_2O" (EMI_39_post).

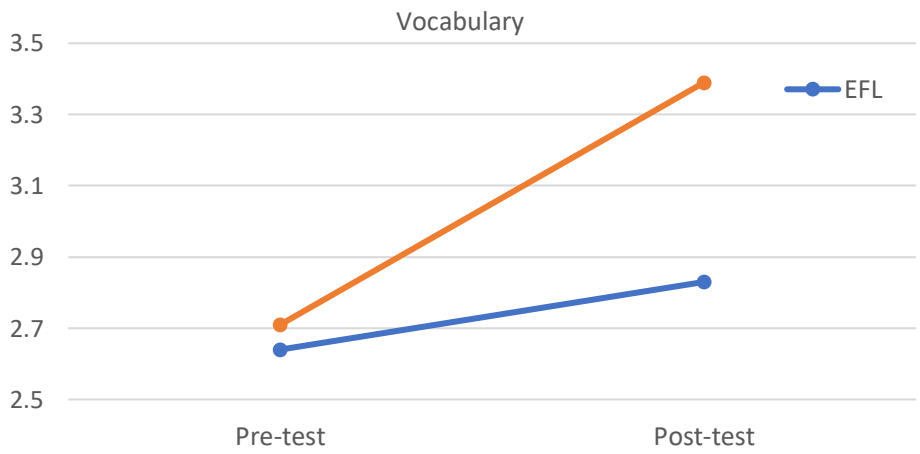

Figure 10 EFL and EMI progress in vocabulary

5. Discussion

According to our research question, this paper set out to explore the differences in the development of the writing competence in English between undergraduate engineering students following an EMI and an EFL course, respectively, in order to explore the impact of EMI on student English writing proficiency. Our findings reveal significant advantages of EMI in the quantitative accuracy variable at the level of lexicon and in the qualitative category of vocabulary, which indicates a positive influence of EMI on student writing proficiency in terms of committing fewer lexical errors, showing greater lexical variety, appropriate word choice, fewer spelling mistakes and clearer meaning.

Regarding the rest of the measures analyzed, while no statistically significant differences were observed between both groups' progress, we found interesting trends worthy of mention. First, the EMI group achieved higher gains than the EFL group in the domain of lexical complexity, where the difference between both groups' gains almost showed a statistically significant level, thus signaling higher progress than the EFL group for lexical richness. Additionally, it showed a tendency toward outrating the EFL group's progress for fluency. Conversely, the EFL group showed higher progress than the EMI group for syntactic complexity, organization, and grammar.

Our findings echo, to a certain extent, previous research conducted in CLIL contexts, where students' language advantage was observed at lexical and vocabulary level (Jexenflicker & Dalton-Puffer, 2010; Lasagabaster, 2008; Ruiz de Zarobe, 2010; Salaberri-Ramiro & Sánchez-Pérez, 2015; Sánchez-Pérez & Salaberri-Ramiro, 2015; Whittaker et al., 2011). As for lexical accuracy, our results concur partially with a study conducted by Roquet and Pérez Vidal (2015) in a Spanish secondary CLIL context where writing was assessed by means of quantitative and qualitative measures, as in our study. In their analysis, the authors found significant advantages in the domain of accuracy in a group of learners who received CLIL plus English language formal instruction (FI) concurrently, as compared to a CLIL-only group. In our study, significance was only found in the accuracy sub-domain of lexicon. However, we must note that, in the EMI context of this study, in contrast to CLIL environments, students did not receive explicit language instruction. These results could be explained by the fact that EMI participants were exposed to reading and/or writing different types of words repeatedly throughout lessons and units, which could have meant that they became more familiar with such words and thus committed fewer errors at the end of the course. Graham et al. (2018, p. 31) exemplify this possibility as follows: "in a science classroom, vocabulary such as experiment or mass will likely be used across units, allowing for multiple encounters." Our findings suggest that

EMI provides opportunities to practice lexicon repetitiously, thus offering multiple exposures to certain vocabulary input. This is less likely to occur in traditional EFL contexts where the teaching and learning scope is generally broader; thus similar vocabulary may not be found across units. Additionally, EMI students are usually more exposed to a wider range of specialized lexicon and complex words than learners of general English courses (Wilkinson, 2013). This could explain the significant outperformance of the EMI group on the vocabulary qualitative measure, and the nearly-significance advantage in lexical complexity with respect to the EFL group in our study. It is noteworthy that our EMI sample showed numerous instances of specialized vocabulary use after completing their course, which was not found in the EFL group. The absence of such instances in the written output of the EFL group was not surprising, since the EFL learning context was not a specialized language course (e.g., English for specific purposes [ESP] or English for academic purposes [EAP]), but a traditional English language course for general communicative purposes. Our results corroborate recent research that argues EMI to be a fertile ground for English lexical development, including both general and specialized vocabulary (Sánchez-Pérez, 2021; Zhang & Pladevall-Ballester, 2021).

Regarding fluency, our results did not show a significant influence of EMI on this measure. While this finding does not concur with some studies conducted in CLIL environments, in which fluency appears as a clearly developed writing area (Gené-Gil et al., 2015; Ruiz de Zarobe, 2010), it should be noted that contradictory findings on this aspect have also been found in other CLIL contexts (e.g., Roquet & Pérez-Vidal, 2015). In our study, the EMI group showed a tendency toward writing longer pieces of texts than EFL students. The fact that the difference between both groups did not reach a statistically significant level may be due to the relatively short duration of the study (i.e., one semester) or to the specificities of classroom tasks. Further research including longer periods of treatment and further research tools, such as classroom observation, would be necessary to dig deeper into this matter.

As for syntactic complexity and grammar, our study did not reveal a significant influence of EMI on these measures as compared to EFL. Indeed, although not to a statistically significant level, the EFL group outperformed the EMI groups' gains in the use of subordination, accurate grammar, variety of forms and complex structures. These results diverge from some previous research on writing development in CLIL contexts which reported CLIL writing advantages at the level of grammar when compared to EFL writers (Jexenflicker & Dalton-Puffer, 2010; Lasagabaster, 2008; Ruiz de Zarobe, 2010). Our findings suggest that in EMI environments, where there is no language instruction, some language guidance and focus on form are needed in order for grammatical and syntactical progress to be seen in higher education learners (Ament & Pérez-Vidal, 2015). The EMI group in our study showed virtually no progress as for syntactic complexity and organization,

which could be interpreted as a form of language "fossilization" (Wilkinson, 2013, p. 19) derived from the repeated use of the language in their writing with no feedback or instruction for improvement. As Wilkinson (2013) claims, if content is taught through an FL with no focus on language, it is unlikely that students will improve the English language skills they already had before engaging in EMI, beyond acquiring new specific vocabulary. This resembles very closely what is observed in this study, where significant gains in EMI were only found at the level of lexicon, and other writing domains, such as syntactic complexity and organization, remained hardly affected. Regarding the latter, the EMI groups' scores even remained disappointing after the treatment. This finding is not surprising, as similar results are found in previous CLIL research (Jexenflicker & Dalton-Puffer, 2010). If this outcome is observed in contexts where some form of language instruction is provided, improvement could hardly be expected in an EMI setting where language instruction is overlooked. Another potential reason for the lack of development at the levels of fluency, syntactic complexity, and grammar may be attributed to the nature of the students' area of knowledge (i.e., engineering). In contrast to other disciplines, such as social sciences or humanities, where students generally need to produce essays, engineering students may not be greatly involved in composition writing, and thus there might be some aspects of the writing competence that are less likely or may take longer to develop. Further empirical research into writing in EMI in other disciplines should be conducted to corroborate this issue.

In view of these results, this study unveils a positive impact of EMI on student writing development at the level of lexical accuracy and vocabulary. However, it suggests that the sole exposure to the language does not suffice to improve other writing areas such as syntax, grammar, or organization. The overall assumption that EMI contributes to improving English language proficiency can, therefore, be confirmed, only partially, as far as writing is concerned.

6. Conclusions

This study has sought to shed some light on the empirically unanswered question concerning the potential of EMI for English language development in higher education settings. In particular, it has focused on writing, a language skill that has been notably underexplored in EMI research (Breeze & Dafouz, 2017; Dafouz, 2020; Graham & Eslami, 2020). Our findings reveal a positive impact of EMI on student writing development at the levels of lexical accuracy and vocabulary, as compared to a traditional English language learning environment. However, other writing areas, such as syntax, grammar, organization, and fluency appear unaffected. These results indicate that simply being exposed to the language in an EMI setting may not necessarily lead to an enhancement of students' writing skills in

English, except for their vocabulary. Our results, therefore, confirm, only partially, the widespread belief that EMI contributes to enhancing English language proficiency as far as writing is concerned.

From a pedagogical standpoint, this study suggests that an approach involving integration of content and language in higher education (ICLHE) could be more effective than a uniquely content-focused EMI model if language skill achievement, beyond specific vocabulary acquisition, is sought within program outcomes at tertiary level (Ament & Pérez-Vidal, 2015; Sánchez-Pérez, 2021; Wilkinson, 2018). The combination of EMI with some form of language skill instruction would allow students to benefit from the EMI advantages observed in this study, enriching and improving their lexical repertoire. This is because of the potential of explicit language skill instruction to "de-fossilize" other language and writing areas that appear to be unaffected in EMI, such as syntax, grammar, cohesion, coherence, paragraphing, and punctuation. This way, a more comprehensive development of students' linguistic, communicative, and disciplinary competence could be achieved.

In order to attain complete advancement of writing abilities in EMI programs, we recommend some pedagogical initiatives, such as implementing EFL, EAP or ESP parallel or in embedded programs (McKinley & Rose, 2022; Sánchez-Pérez, 2020; Unterberger, 2018). Another possibility could be fostering collaborative teaching between content and language specialists (Lasagabaster, 2018), attending closely to the specific language and literacy needs of the EMI course. Additionally, specific L2 academic writing programs addressed both to undergraduate and postgraduate students would be highly beneficial for EMI learners, something that is hardly found in universities from Southern Europe, such as Spain (Dafouz, 2020). University authorities should, therefore, consider such initiatives within their EMI policies if the provision of opportunities for students to improve their English language skills is pursued (Wilkinson, 2018).

Notwithstanding the relevance of this study, we must acknowledge some limitations. First, it was conducted in a particular EMI context with students from a specific disciplinary area (i.e., engineering). Given the highly context-specific nature of EMI (Rose et al., 2019), and the specificity of the field of knowledge explored in this study, caution should be used when generalizing our results. Different findings could be observed in different contexts and/or disciplines (e.g., in more linguistically demanding areas, such as social sciences or humanities). Furthermore, our longitudinal analysis was conducted over a relatively short period (i.e., one semester). While this time sufficed to show significant gains in EMI students' writing competence at the levels of lexical accuracy and vocabulary, it is possible that a more extended period might have yielded significant results in other writing areas that may take longer to develop (Pérez-Vidal, 2007). Finally, whilst the study analyzed different aspects of the writing ability, both quantitatively and

qualitatively, we are aware that the analysis of further quantitative and/or qualitative measures could yield different outcomes. Therefore, prospective studies should include other EMI contexts, address students from different areas of knowledge, include longer periods of treatment, and consider additional writing measures and/or research tools (e.g., classroom observation) to provide further insights into English writing outcomes in the increasing EMI higher education contexts.

References

Aguilar, M., & Muñoz, C. (2014). The effect of proficiency on CLIL benefits in engineering students in Spain. *International Journal of Applied Linguistics, 24*(1), 1-18. https://doi.org/10.1111/ijal.12006

Aguilar, M., & Rodríguez, R. (2012). Lecturer and student perceptions on CLIL at a Spanish university. *International Journal of Bilingual Education and Bilingualism, 15*(2), 183-197. https://doi.org/10.1080/13670050.2011.615906

Ament, J. R., & Pérez-Vidal, C. (2015). Linguistic outcomes of English medium instruction programmes in higher education: A study on economics undergraduates at a Catalan university. *Higher Learning Research Communications, 5*(1), 47-67. https://doi.org/10.18870/hlrc.v5i1.239

Breeze, R., & Dafouz, E. (2017). Constructing complex cognitive discourse functions in higher education: An exploratory study of exam answers in Spanish and English-medium instruction settings. *System, 70*, 81-91. https://doi.org/10.1016/j.system.2017.09.024

Briggs, J. G., Dearden, J., & Macaro, E. (2018). English medium instruction: Comparing teacher beliefs in secondary and tertiary education. *Studies in Second Language Learning and Teaching, 8*(3), 673-696. https://doi.org/10.14746/ssllt.2018.8.3.7

Chansri, C., & Wasanasomsithi, P. (2016). Implementing CLIL in higher education in Thailand: The extent to which CLIL improves agricultural students' writing ability, agricultural content and cultural knowledge. *Journal of Language Teaching and Learning in Thailand 51,* 15-38. https://doi.org/10.14456/pasaa.2016.1

Dafouz, E. (2020). Undergraduate student academic writing in English-medium higher education: Explorations through the ROAD-MAPPING lens. *Journal of English for Academic Purposes, 46*, 1-12. https://doi.org/10.1016/j.jeap.2020.100888

Dafouz, E., & Smit, U. (2020). *ROAD-MAPPING English-medium education in the internationalised university.* Palgrave McMillan.

Dearden, J., & Macaro, E. (2016). Higher education teachers' attitudes towards English medium instruction: A three-country comparison. *Studies in Second Language Learning and Teaching, 6*(3), 455-486. https://doi.org/10.14746/ssllt.2016.6.3.5

Foster, P., & Skehan, P. (1996). The influence of planning and task type on second language performance. *Studies in Second Language Acquisition, 18*, 299-324. https://doi.org/10.1017/S0272263100015047

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

AHS, lebende Fremdsprachen. BIFIE. Retrieved from: http://www.bifie.at/publikationen

Galloway, N., Kriukow, J., & Numajiri, T. (2017). *Internationalisation, higher education and the growing demand for English: An investigation into the English medium of instruction (EMI) movement in China and Japan.* British Council.

Galloway, N., & Rose, H. (2021). English medium instruction and the English language practitioner. *ELT Journal, 75*, 33-41. https://doi.org/10.1093/elt/ccaa063

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. https://doi.org/10.1080/09571736.2015.1053278

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 and Language Integrated Learning, 11*(1), 19-37. https://doi.org/10.5294/laclil.2018.11.1.2.

Graham, K. M., & Eslami, Z. R. (2020). Writing measures and outcomes in CLIL and EMI: A systematic review. *Writing & Pedagogy, 12*(2-3), 341-368. https://doi.org/10.1558/wap.17755

Hernández, N., & Jiménez, A. (2017). English as a medium of instruction: Evidence for language and content targets in bilingual education in economics. *International Journal of Bilingual Education and Bilingualism, 20*(7), 883-896. https://doi.org/10.1080/13670050.2015.1125847

Hidalgo, M. A., & Lázaro-Ibarrola, A. (2020). Task repetition and collaborative writing by EFL children: Beyond CAF measures. *Studies in Second Language Learning and Teaching, 10*(3), 501-522. https://doi.org/10.14746/ssllt.2020.10.3.5

Hughes, A. (2003). *Testing for language teachers* (2nd ed). Cambridge University Press.

Jexenflicker, S., & Dalton-Puffer, C. (2010). The CLIL differential: Comparing the writing of CLIL and non-CLIL students in higher colleges of technology. In C. Dalton-Puffer, T. Nikula, & U. Smit (Eds.), *Language use and language learning in CLIL classrooms* (pp. 169-189). John Benjamins.

Kamaşak, R., Sahan, K., & Rose, H. (2020). Academic language-related challenges at an English-medium university. *English for Academic Purposes, 49*, 100945. https://doi.org/10.1016/j.jeap.2020.100945

Larson-Hall, J. (2010). *A guide to doing statistics in second language research using SPSS*. Routledge.

Lasagabaster, D. (2008). Foreign language competence in content and language integrated courses. *The Open Applied Linguistics Journal, 1*, 31-42. https://doi.org/10.2174/1874913500801010030

Lasagabaster, D. (2018). Fostering team teaching: Mapping out a research agenda for English-medium instruction at university level. *Language Teaching, 51*(3), 400-416. https://doi.org/10.1017/S0261444818000113

Macaro, E. (2018). *English medium instruction: Language and content in policy and practice.* Oxford University Press.

Macaro, E., Curle, S., Pun, J., An, J., & Dearden, J. (2018). A systematic review of English medium instruction in higher education. *Language Teaching, 51*(1), 36-76. https://doi.org/10.1017/S0261444817000350

McKinley, J., & Rose, H. (2022). English language teaching and English-medium instruction: Putting research into practice. *Journal of English-Medium Instruction, 1*(1), 85-104. https://doi.org/10.1075/jemi.21026.mck

Norris, J. M., & Ortega, L. (2009). Towards an organic approach to investigating CAF in instructed SLA: The case of complexity. *Applied Linguistics, 30*(4), 555-578. https://doi.org/10.1093/applin/amp044

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

Ortega, L. (2003). Syntactic complexity measures and their relationship to L2 proficiency: A research synthesis of college-level L2 writing. *Applied Linguistics, 24*(4), 492-518. https://doi.org/10.1093/applin/24.4.492

Parkinson, J. (2017). Teaching writing for science and technology. In J. Flowerdew & T. Costley (Eds.). *Discipline-specific writing: Theory into practice* (pp. 95-112). Routledge.

Pecorari, D., & Malmström, H. (2018). At the crossroads of TESOL and English medium instruction. *TESOL Quarterly, 52*(3), 497-515. https://doi.org/10.1002/tesq.470

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

Rose, H., Macaro, E., Sahan, K., Aizawa, I., Zhou, S., & Wei, M. (2021). Defining English medium instruction: Striving for comparative equivalence. *Language Teaching,* 1-12. https://doi.org/10.1017/S0261444821000483

Rose, H., Curle, S., Aizawa, I., & Thompson, G. (2019). What drives success in English medium taught courses? The interplay between language proficiency, academic skills, and motivation. *Studies in Higher Education, 45*(11), 2149-2161. https://doi.org/10.1080/03075079.2019.1590690

Ruiz de Zarobe, Y. (2010). Written production and CLIL: An empirical study. In C. Dalton-Puffer, T. Nikula & U. Smit (Eds.), *Language use and language learning in CLIL classrooms* (pp. 191-209). John Benjamins.

Salaberri-Ramiro, M. S., & Sánchez-Pérez, M. M. (2015). Analyzing writing in English-medium instruction at university. *Linguarum Arena, 6*, 45-58.

Sánchez-Pérez, M. M., & Salaberri-Ramiro, M. S. (2015). Writing professional genres in a second language: Results from a Spanish university CLIL context. *American Journal of Educational Research, 3*(5), 576-580. https://doi.org/10.12691/education-3-5-7

Salaberri-Ramiro, M. S., & Sánchez-Pérez, M. M. (2018). Motivations of higher education students to enrol in bilingual courses. *Porta Linguarum, 3*, 61-74.

Sánchez-Pérez, M. M. (2020). *Teacher training for English-medium instruction in higher education*. IGI-Global. https://doi.org/10.4018/978-1-7998-2318-6

Sánchez-Pérez, M. M. (2021). Predicting content proficiency through disciplinary-literacy variables in English-medium writing. *System, 97*, 102463. https://doi.org/10.1016/j.system.2021.102463

Tatzl, D. (2011). English-medium masters' programmes at an Austrian university of applied sciences: Attitudes, experiences and challenges. *Journal of English for Academic Purposes, 10*(4), 252-270. https://doi.org/10.1016/j.jeap.2011.08.003

Unterberger, B. S. (2018). The English-medium paradigm: A conceptualisation of English-medium teaching in higher education. *International Journal of Bilingual Education and Bilingualism, 21*(5), 527-539. https://doi.org/10.1080/13670050.2018.1491949

Vidal, K., & Jarvis, S. (2020). Effects of English-medium instruction on Spanish students' proficiency and lexical diversity in English. *Language Teaching Research, 24*(5), 568-587. https://doi.org/10.1177/1362168818817945

Wächter, B., & Maiworm, F. (2014). *English-taught programmes in European higher education. The state of play in 2014.* Lemmens.

Wilkinson, R. (2013). English-medium instruction at a Dutch University: Challenges and pitfalls. In A. Doiz, D. Lasagabaster, & J. Sierra (Eds.). *English-medium instruction at universities* (pp. 3-26). Multilingual Matters.

Wilkinson, R. (2018). Content and language integration at universities? Collaborative reflections. *International Journal of Bilingual Education and Bilingualism, 21*(5), 607-615. https://doi.org/10.1080/13670050.2018.1491948

Whittaker, R., Llinares, A., & McCabe, A. (2011). Written discourse development in CLIL at secondary school. *Language Teaching Research, 15*(3), 343-362. https://doi.org/10.1177/1362168811401154

Wolfe-Quintero, K., Inagaki, S., & Kim, H. (1998). *Second language development in writing: Measures of fluency, accuracy and complexity*. University of Hawaii Press.

Zhang, M., & Pladevall-Ballester, E. (2021). Discipline-specific language learning outcomes in EMI programs in the People's Republic of China. *Language and Education, 35*(4), 357-374. https://doi.org/10.1080/09500782.2021.1895827