COMPLEXITY, ACCURACY, AND FLUENCY (CAF) OF EFL COLLEGE LEARNERS' WRITING: THE CLIL AND NON-CLIL CLASSES

*Masrul[1]; Sharifah Sheha Syed Aziz Baftim[2]; Bayu Hendro Wicaksono[3]
[1]English Language Education, University of Pahlawan Tuanku Tambusai, Indonesia;
[2]English Language Education, Universiti Teknologi MARA (UiTM) Shah Alam, Malaysia;
[3]English Language Education, University of Muhammadiyah Malang, Indonesia
* *masrulm25@gmail.com; sharifah@uitm.edu.my; bayu_hw@umm.ac.id*
(*) Corresponding Author

Abstract: This study was designed to compare the complexity, accuracy, and fluency of EFL written text in CLIL and Non-CLIL classes. The study enrolled two groups of undergraduate students from the State University of Malang, Indonesia: an experimental CLIL class (N = 50 students; 22 males and 28 females) and a non-CLIL as control class (N = 50 students; 19 males and 31 females). Students' essays were evaluated quantitatively using some aspects of linguistic proficiency, such as complexity, accuracy, and fluency. The errors were classified as syntactic, morphological,

lexical, lexicogram, spelling, and punctuation mistakes. The findings indicated that both CLIL and non-CLIL methods produced comparable complexity, accuracy, fluency, syntactic, morphological, lexicogram, and spelling scores in two groups of students. Meanwhile, for complexity and lexical values, the CLIL and Non-CLIL methods produced significantly different average scores, with the application of the non-CLIL method being higher. On the punctuation variable, the CLIL method can significantly improve the assessment. Several possible explanations for the increase in complexity, accuracy, and fluency include a preference for the English standard, the course's assessment criteria, and practice effects. The findings of this study also provide additional pedagogical implications.

Keywords: *accuracy, CLIL, complexity, EFL writing, fluency, non-CLIL*

INTRODUCTION

Writing has long been a critical component of assessing students' language competence in the EFL classroom. Written texts demonstrate the second language (L2) user's active language usage in many attributes, including vocabulary, phrases, verb tenses, sentence structures, and errors (Verspoor et al., 2012). In recent years, there has been a new type of learning method known as Content and Language Integrated Learning (CLIL) spreading throughout European education. This innovative approach has also influenced the way writing is taught, with CLIL emphasizing the integration of language skills within specific content areas, fostering a holistic learning experience that goes beyond traditional language instruction. The CLIL, or English as an instruction medium, is a pedagogical strategy in which specific subjects are taught in languages other than the students' mother tongues (Lyster, 2011). It has grown in popularity as a term for instructional methods used in primary and secondary school programs that utilize a second or foreign language as the medium of instruction for specific curricular subjects (Dalton-Puffer,

2008). The CLIL has been repeatedly recommended by European institutions due to its multifaceted nature, as it aims to promote second or foreign language education, plurilingualism, bilingualism, and learners' worldwide orientation (Cenoz & Gorter, 2014;Pérez-Cañado, 2012;Pérez-Vidal & Roquet, 2015; Temirova & Westall, 2015). It is considered that students' can also absorb the language while they are learning the subject (Krashen, 1981). Several studies have indicated that language growth is viewed as a foregone conclusion after using a CLIL method (Arnó-Macià & Mancho-Barés, 2015) even if there were few clear linguistic instructions (Dalton-Puffer, 2008).

Bilingual programs or the CLIL method have been recommended for various reasons (Coyle, 2005). This approach is advocated due to its potential to enhance language proficiency while simultaneously deepening understanding in specific subject areas. Additionally, CLIL promotes a more authentic and contextualized language learning experience, as it requires students to use the target language in meaningful, real-world situations related to the content being studied. Moreover, research suggests that CLIL can contribute to cognitive benefits, such as improved problem-solving skills and critical thinking, making it a valuable educational strategy.

The CLIL employs a second or foreign language as an instruction medium for subject areas at the elementary and secondary levels. (Pérez-Vidal & Roquet, 2015) emphasized certain essential aspects of these programs, including the global value system that this educational alternative exhibits in the classroom and the fact that these programs would not be possible without the support of a strong policy. This demonstrates unequivocally that CLIL is viewed as an educational method rather than a straightforward methodology. By focusing on meaning and communication, CLIL method is considered to promote implicit and incidental learning and enhance overall linguistic competence in the target language in a CLIL class (Heras & Lasagabaster, 2015).

Although some controversy about scientific evidence for CLIL method has raged unabated, the CLIL research has accelerated over

the recent decade, and numerous research on the effect of the CLIL on language instruction have revealed beneficial outcomes (e.g. Bruton, 2011, 2013). Certain aspects of language proficiency have garnered greater consideration than others. Specifically, the role of CLIL in writing contexts has been greatly dismissed (Llinares et al., 2012). Indeed, in an overview study of CLIL learning outcomes, Dalton-Puffer (2008) identified writing ability as one of the aspects of language skills that are unlikely to be influenced by CLIL method. (Rallo Fabra & Jacob, 2015) , for their part, reviewed research on linguistic competencies and cognitive abilities in CLIL and concluded that while clear improvements were observed in some areas of writing, such as lexical, syntactic complexity, and accuracy, no such improvement was observed in others, such as discourse skills and fluency.

In formal education settings, research on written complexity, accuracy, and fluency demonstrates clear improvement developments. Accordingly, Verspoor et al. (2012) examined the written texts of a group of first- and third-year ESL students. They examined 64 distinct factors affecting sentence structures, clause constructions, verb phrase constructions, the lexical, and accuracy assessments crosswise five distinct skill levels, from beginner to intermediate. As proficiency levels increased, all variables examined improved, more multifaceted structures appeared at all stages, and fewer errors existed. Additionally, the findings indicated that the phrase length, complexity, total number of dependent clauses, and errors all differentiated among levels of writing competence.

Other research conducted by Lahuerta (2020) show that in the setting of CLIL instruction, accuracy improved dramatically with grade, whereas in the non-CLIL group, only lexicogrammatical errors shrank significantly. With articles creating issues in both contexts and prepositions, determiners, voice, subordination, and word order in the non-CLIL context, the examination of mistake subtypes indicated certain cases of regression and stabilization tendencies. In terms of gender, participants found that they performed similarly to their male

counterparts in the CLIL group as they did in the non-CLIL group for written accuracy, suggesting that CLIL may help balance gender inequalities.

Based on the results of their English proficiency tests, the participants were classified as having a high-intermediate level of English proficiency. The participants' written products were evaluated on the basis of their fluency (the total number of words and words per T-unit), accuracy (the number of errors per number of words), academic vocabulary use, and grammatical structures. The results indicate that there was no significant improvement in accuracy, fluency, or vocabulary use. However, there was an improvement in the text structure and content, even though the content was not statistically significant. In the research conducted by Aguilar and Muñoz (2019)shows that the difference in the mean scores in the pre- and post-listening exams was significant, but not for the pre- and post-grammar tests, according to paired- sampled t-tests. A repeated measure ANOVA revealed that less proficient students improved their hearing and grammatical abilities more than more competent students when students were divided into three groups based on their pre-test scores.

Additionally, progress in Thewissen (2013) provided insight into error behaviour (2013). Using an error-tagged version of EFL learner corpus, she examined the dynamics of second language accuracy. Learners' writings were corrected and rated based on one of frameworks of reference for language criteria of language competence. Lexis was the most significantly advanced skill from intermediate to advanced levels, according to the findings of this study. The findings indicate that tendencies toward development and stability dominate EFL error development patterns, and that progress is frequently distributed.

In other sides, some research is worth discussing regarding the differences between first and second language writing. For instance, Cabrera Solano et al. (2014) examined the effect of the Spanish language on second language writing skills in secondary education.

Data collection tools included learner and instructor questionnaires, as well as a writing test in which students were asked to write a narrative text. According to the findings, the most significant amount of first language influence occurred in English grammar and vocabulary. The most frequent Spanish intervention errors were verb misuse, a lack of personal and object pronouns, preposition misuse, incorrect punctuation, excessive use of articles, spelling errors, and an inappropriate word pattern.

Furthermore, the positive effects of writing development are still unclear in the CLIL research. While some studies indicate that writing growth is constrained in the CLIL classes, others demonstrate that the CLIL enhances written competence. Among previous studies, Llinares and Whittaker (2007) discovered significant deficits in the writing production of secondary CLIL students. Similar to their previous study, Llinares and Whittaker (2006) discovered that their Spanish CLIL participants at secondary school studying social science through English achieved some subject-specific features of their discipline (e.g., the allocation of the most frequently used words in compositions), but that other resources, such as modality or clause elaboration, were rarely used in their compositions. Among the numerous studies demonstrating the benefits of writing, Whittaker et al. (2011)demonstrated that the CLIL students advanced in the academic inventory requirement and produced more coherent and constructive writing. Lasagabaster (2008) examined language performance in English holistically, accounting for multiple language skills into account, using a one-time data collection design consists of two CLIL groups and one non-CLIL group. His study discovered that the CLIL students outperformed their peers in terms of overall written productivity and the majority of the measures examined.

The effects of Content and Language Integrated Learning (CLIL) on writing in other countries are generally evident in the improvement of foreign language skills and a deeper understanding of specific content. By integrating language learning with other subjects such as science or arts, students can develop their language

proficiency contextually and in a more profound manner. This expansion of insight enables them to convey information more clearly and precisely in a foreign language.

His findings paralleled those of Ruiz de Zarobe (2010), who used a similar holistic methodology with a different strategy, including limiting introduction to the target language. The results of this study indicated that 15–16-year-old secondary learners in CLIL groups outperformed their non-CLIL peers in vocabulary, the use of language, and mechanics. In another study, Jexenflicker & Dalton-Puffer(2010) compared the writing of CLIL and non-CLIL students in Austrian higher colleges of technology. They discovered that CLIL students had considerable advantages in lexicogramm as well as vocabulary range and syntax correctness. Differences in speech abilities and textual arrangement, on the other hand, were difficult to distinguish. A rating scale was used to assess task completion, organization and structure, grammar, and vocabulary. Similar with the study conducted by Lee (2020) in terms of writing quality, lexical sophistication, lexical diversity, and mean length of clause, the CLIL group beat the non-CLIL group; nevertheless, the non-CLIL group's writing displayed higher semantic cohesiveness. The findings contribute to CLIL research by extending it to examine CLIL effects on syntactic complexity, lexical complexity, and cohesiveness in addition to writing quality. Previously, CLIL research was restricted to a focus on morphosyntactic aspects or writing quality.

In Indonesia, the implementation of CLIL also has positive impacts on writing. Besides enhancing foreign language skills, CLIL can positively influence the mastery of specific subject matter. Writing in a foreign language can become more skillful and focused, allowing students to articulate their ideas more effectively. However, challenges in the implementation of CLIL in Indonesia include resource availability and training for teachers. Nevertheless, the adoption of CLIL in Indonesia has the potential to improve the quality of education and students' language abilities in a global context.

To summarize, complexity, fluency, and accuracy have been established as objective indicators of L2 writing proficiency in EFL contexts where English learning in higher education was a major focus. However, one distinct educational CLIL and Non-CLIL classes remains largely unexplored. In other words, little consideration is given to whether EFL writing will improve in a setting where content is typically the primary focus. As a result, research is needed to determine whether EFL college students' complexity, accuracy, and fluency in morphological, lexical, lexicogram, spelling, and punctuation skills can be improved through CLIL method.

To that end, this study examines the aspects of linguistic proficiency, which are the complexity, accuracy, and fluency of students enrolled in the CLIL and the Non-CLIL classes in their writing. It is important to address this question because a thorough understanding of CAF in this context can provide valuable insights into the effectiveness of the CLIL approach in language teaching. The choice of CAF as the focus of this research is made because language proficiency is a key aspect in evaluating students' language abilities and can offer a more holistic understanding of their proficiency in using a particular language. Additionally, this research can provide a useful comparison between students participating in CLIL programs and those who are not, allowing for the identification of potential benefits of this language learning approach. The comparison of the two contexts focuses on the complexity, accuracy, fluency, syntactic, morphological, lexical, lexicogram, spelling, and punctuation characteristics of both classes. The current study seeks to clarify this issue by examining language proficiency in writing when a foreign language is employed as a form of interaction in the CLIL and the Non-CLIL classes. The method employed in this paper is based on examining specific errors made by students in two educational settings. This method has not been used previously in the literature to our knowledge.

This study's novelty lies in its exploration of the linguistic proficiency aspects—complexity, accuracy, and fluency (CAF)—

specifically within the context of CLIL (Content and Language Integrated Learning) and Non-CLIL classes. By focusing on these variables, we aim to offer a comprehensive analysis of students' writing abilities within both educational settings. This approach allows for a nuanced understanding of the impact of CLIL on language proficiency development compared to traditional language instruction methods. Additionally, by examining these aspects, we contribute to the existing body of literature on CLIL effectiveness and provide valuable insights for educators and policymakers seeking to enhance language learning outcomes.

METHOD
Research Design

This study aimed to compare the complexity, accuracy, and fluency of EFL written text in CLIL and non-CLIL classes among undergraduate students at the State University of Malang, Indonesia, focusing on quantitative evaluation of various aspects of linguistic proficiency and classification of errors made in their essays. The research design employed in this study is experimental research. The main objective is to evaluate the impact of two different teaching methods (CLIL and Non-CLIL) on the complexity, accuracy, and fluency of EFL written texts. A T-test was performed to examine the difference in mean values of observations between the CLIL and Non-CLIL groups for each test variable.

Research Participant

This study enrolled 100 students, divided into two groups of undergraduate students from the State University of Malang, Indonesia: a CLIL as experimental class (N = 50; 22 males and 28 females) and a comparable non-CLIL as control class (N = 50; 19 males and 31 females). The average age of CLIL and non-CLIL students was 19 years, and they started studying English at five years old. Participants were enrolled in an English for Academic Purposes

(EAP) course at the undergraduate level. The participants' English proficiency ranged from intermediate to advance.

Research Treatment Procedure

This study was conducted at a university as part of an initiative to promote bilingual education in Indonesia. Bahasa is the primary language of instruction at universities, while English is taught as a foreign language. CLIL and Non-CLIL classes had been given the same durations of EFL learning by the time data collection began. While both CLIL classes took a communicative method, the CLIL method provide more time on language structures and grammar. Written text received no special treatment in the non-CLIL contexts studied, with the majority of EFL writing tasks being brief and generic, whereas CLIL compositions were more specific. Regularly, learners did not receive personalized responses to their writing assignments.

Data Collection

The data for this study were gathered through a written task that participants completed in their own classrooms. Weekly assignments related to course content were also required. They were informed that their essays would be graded on the following criteria: organization, content, complexity, fluency, and accuracy (morphology, lexical, punctuation, spelling, and glaring grammatical errors that impede communication). Students were required to write about a single topic for the written activity. Thirty minutes were allotted to each participant for the writing process. Thus, both time and subject matter constraints were maintained to ensure that the findings were equivalent (Wolfe-Quintero et al., 1998). The first, middle, and final weeks of the semester's assignments were chosen for comparison to determine whether the participants' complexity, accuracy, and fluency increased throughout the semester.

Following a functional linguistics approach (Rose, 2020), written competency was defined in this study as the capacity to

communicate meaningfully through the written text (Myles, 2002). It was evaluated quantitatively using three aspects of linguistic proficiency: complexity, accuracy, and fluency (CAF). Even though there are some concerns and disagreements about the CAF measures (e.g. Pallotti, 2009), they are widely used in studies to measure language learners' writing.

The term of complexity in this study has been used to refer to the total environment of specific linguistic elements, as well as the interaction between these constituent components, including syntax and lexis (Bulté & Housen, 2012). The coordination index (CI) was used to calculate syntactic complexity, which is determined by dividing the number of independent clause coordination by the total number of combined clauses, that is, the total number of clauses minus the total number of sentences (Wolfe-Quintero et al., 1998). This measure, with values ranging from 0 to 1, should drop as learners enhance their competency. Furthermore, this study added the subordination ratio (SR), which is clauses/T-unit. This ratio accepts values equal to or greater than one, with one signifying the least amount of subordination, and consequently, the least complexity. A T-unit or smallest terminable unit of language is defined by Hunt(1964) as a major independent clause plus all of its subordinate dependent clauses. Consequently, this study considered any verb phrase, finite or non-finite, to be a clause.

Syntactic complexity was calculated using the coordination index (CI), which is determined by taking the total number of independent clause coordination by the total number of merged clauses, or the total number of clauses minus the total number of sentences(Wolfe-Quintero et al., 1998). This metric, ranging from 0 to 1, should decrease as learners improve their competency. Additionally, we included the subordination ratio (SR), which is expressed as clauses/T-unit. This ratio accepts values of one or greater, with one indicating the least amount of subordination and thus the least complexity. Excerpt 1 demonstrates how we viewed a sentence as a single T-unit or as a collection of coordinated T-units

between punctuation marks. Hunt (1964) defines a T-unit, or smallest terminable unit of language, as a major independent clause plus all of its subordinate dependent clauses. Any verb phrase, whether finite or non-finite, was considered a clause.

A variety of quantitative criteria measuring various aspects of L2 accuracy were also used to evaluate the essays. The error-free sentence ratio (SR) and the errors per word ratio (the total number of errors divided by the total number of words) were used (Wolfe-Quintero et al., 1998), with values ranging from 0 to 1. In the first ratio, a higher value indicates that the text is further precise. Lower values for the second ratio, on the other hand, indicate greater accuracy.

This study identified and evaluated errors of syntactic, morphological, lexical, lexicogrammatical, spelling, and punctuation. Nominal morphology (plural, case), verbal morphology (tense, modality, subject-verb agreement, and passive construction), determiner and article errors, prepositional morphology errors, and derivational morphology errors were all included in the morphological errors (e.g. lack of suffixes, etc.). Syntactic errors included misspellings, omissions, and repetitions of constituents, as well as errors in subordination and coordination. As for lexical errors, we counted connector errors, lexical errors on single words, and lexical phrase errors. Lexicographic errors included errors with countable/uncountable nouns and verb complementation.

We classified punctuation errors as omission, misunderstanding, and redundancy. Spelling errors were included in the category of form errors. The percentage of errors was calculated by dividing the total number of instances of a specific error by the total number of words. Fluency is related to the length concerns in this study, as it is in Wolfe-Quintero et al. (1998). The following ratios were used: T-units per minute (T/M), words (i.e. tokens) per T-unit (W/T), and words (i.e. tokens) per minute (W/M). For that purpose, we examined the composition's 25-minute time limit. Although the number of words and T-units may decrease with the increasing

expertise as T-units lengthen (Wolfe-Quintero et al., 1998), higher values of these ratios may indicate enhanced competence, and thus higher scores in the three fluency measures are expected to increase over time.

Data Analysis

Data analysis in this study was carried out by first testing the normality of the data. The normality test was used to analyze whether the distribution of data within a data group or variable was normally distributed. The normality test was visually performed using a Q-Q Plot graph, followed by descriptive analysis to find out the mean, standard deviation, minimum and maximum data on each variable. Descriptive analysis was continued by creating a graph of the mean of each variable to examine the descriptive difference between the variables using the CLIL and Non-CLIL methods. Finally, T-test examined the difference in the mean value of observations between the samples taken randomly from a population and Non-CLIL for each test variable.

FINDINGS

This study conducted to examine the comparison of the two groups between the CLIL and the Non-CLIL classes in in writing. This study focuses on the complexity, accuracy, fluency, syntactic, morphological, lexical, lexicogram, spelling, and punctuation characteristics of both classes.

To clarify those issues, we need to take normality test of the data. The normality test of the data is carried out visually with the Q-Q Plot graph presented in Figure 1. It shows that the variable data in the CLIL and Non CLIL methods are normally distributed so that data analysis can be continued.
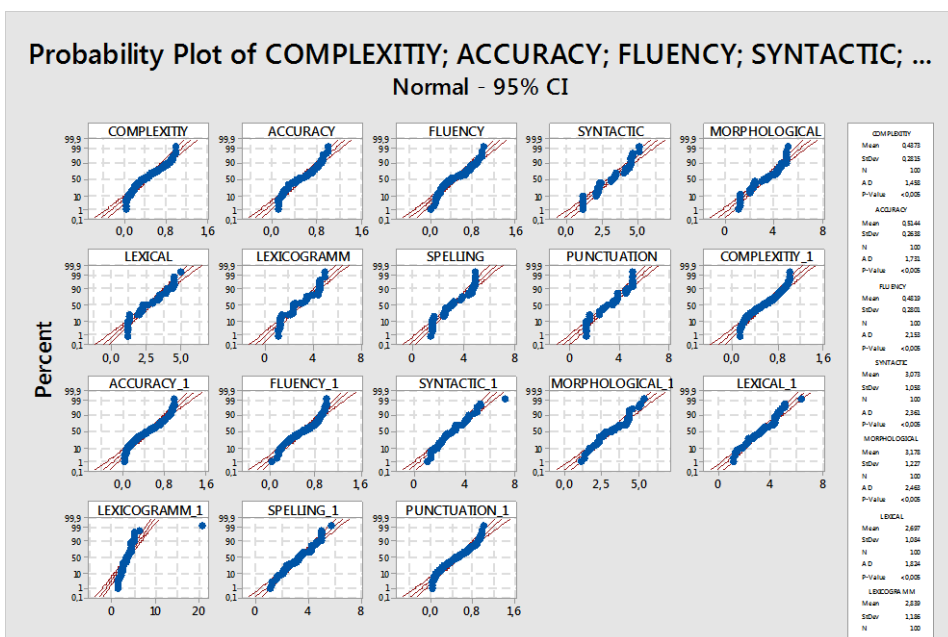
Figure 1. The normal distribution of each variable

Next, we proceed the descriptive data of students using the CLIL and Non-CLIL methods. Table 1 presents the detailed descriptive data for both classes.

Table 1.
Research descriptive statistics

Variable	N	CLIL				NON-CLIL			
		Mean	SD	Min	Max	Mean	SD	Min	Max
Comp-lexity	100	0.4373	0.2627	0.01	1	0.5444	0.2815	0,1	1
Accuracy	100	0.5144	0.2934	0.09	1	0.5017	0.2638	0,01	1
Fluency	100	0.4819	0.2712	0.01	1	0.5719	0.2801	0,01	1
Syntactic	100	3.073	1.142	1.01	5	3.109	1.058	1	7,25
Morpho-logical	100	3.178	1.034	1	5	3.111	1.227	1,05	5,31
Lexical	100	2.697	1.079	1.1	5	3.231	1.084	1,04	6,25
Lexico-gramm	100	2.839	2.098	1	4.99	3.414	1.186	1	21
Spelling	100	3.162	1.072	1.4	4.89	3.326	1.11	1,08	5,79
Punctu-ation	100	3.029	0.2884	1.1	4.99	0.5008	1.176	0,01	1

Table 1 shows an overall representation of the research variables, including the average of each variable, data variation, minimum and maximum values and the number of studies.
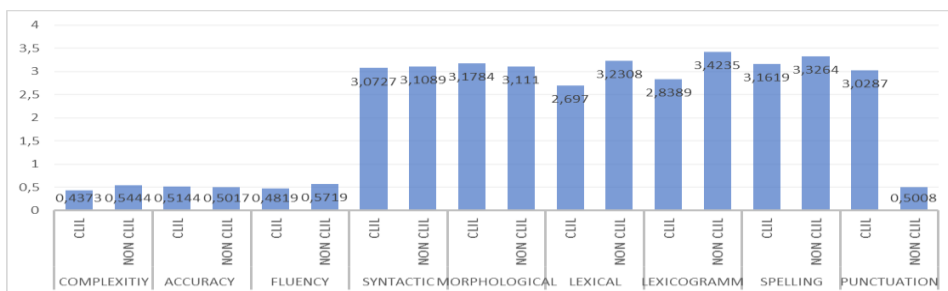

Figure 2. The mean of each research variable

Figure 2 presents the mean of each variable in the CLIL and Non-CLIL methods. It shows that the average of complexity, accuracy, fluency, syntactic, lexical and spelling variables for students using the CLIL and the Non-CLIL method are descriptively different, with the average scores of the Non-CLIL method being higher than the CLIL. On the other hands, the averages of morphological, lexicogram, and punctuation are descriptively different, with the average scores of the CLIL method being higher than Non-CLIL.

Furthermore, to know whether those differences are significant we proceed through T-test. Table 2 reveals the significant value of the score of Complexity, Accuracy, Fluency, Syntactic, Morphological, Lexical, Lexicogram, Spelling and Punctuation in the application of the CLIL and Non-CLIL methods

Table 2.

The comparison of the means of CLIL and Non-CLIL method

Variable	Mean 1 (CLIL)	Mean2 (Non-CLIL)	p
Complexity	0.4373	0.5444	0.006
Accuracy	0.5144	0.5017	0.748
Fluency	0.4819	0.5719	0.022
Syntactic	3.073	3.109	0.816
Morphological	3.178	3.111	0.674
Lexical	2.697	3.231	0.001

Lexicogram	2.839	3.414	0.019
Spelling	3.162	3.326	0.288
Punctuation	3.029	0.5008	0.000

The descriptive analysis and t-test indicate that the CLIL and Non-CLIL techniques yield distinct scores for all variables, but not all of these differences are statistically significant. This study found no significant differences in accuracy, syntactic, morphological, and spelling scores between the CLIL and Non-CLIL techniques. On the other hand, there are notable differences in terms of complexity, fluency, lexical, lexicogram, and punctuation between the CLIL and Non-CLIL approaches. These differences are more pronounced in the non-CLIL method.

DISCUSSION

CLIL and Non-CLIL classes emphasize communication, but the formal teaching program emphasizes linguistic structures and grammar considerably. Writing processes are incorporated into both EFL curriculum programs. In addition to their English classes, bilingual students are given some writing instruction and assignments.

The CLIL utilizes materials that are not typically utilized in countries with an English-speaking population (Lasagabaster & Sierra, 2009). This is because CLIL learning objectives are designed to enable students to achieve a level of proficiency in the instruction language that is less than native-like. As a result, instructional materials have to be customized to the students' proficiency level. Typically, a CLIL method begins when learners have acquired basic literacy skills in their first language during elementary education (Lasagabaster & Sierra, 2009). Therefore, instructional materials should be differentiated according to the learners' proficiency level. CLIL programs frequently begin in primary school, after students have demonstrated basic literacy abilities in their first language. EFL learners joined the CLIL class voluntarily and were free to choose the subject matters and stages involved, as long as their content teachers

possessed the required language qualifications and complied with all applicable legal requirements. Teachers who took part were encouraged to attend CLIL instruction.

In the last decade, the implementation of bilingual instruction programs in the context of Indonesian has undergone significant transforms. In this case, the government discontinued the International Standards School (ISS) program in 2015, which required students to participate in bilingual education programs in primary and secondary schools. Due to the ISS program's discontinuation, studies in Indonesia assessing the performance and progress of bilingual instruction systems are limited. As a result, CLIL is considered more applicable in Indonesian settings, especially in higher educational areas.

The findings of this study indicate that, on average, the CLIL method is effective at increasing students' written competence, implying that CLIL appears to have a beneficial effect on writing (Dalton-Puffer, 2008). To address the first research question, the findings corroborated Storch's (2009) finding that syntactic complexity did not improve after one semester. The lack of complexity improvement could be attributed to the study's brief duration. Syntactic complexity development, according to Ortega's (2003) meta-analysis, may take more than 12 months of higher education instruction. Additionally, one might argue that the participants already have a high level of proficiency, making enhancement more difficult or time-consuming for them (Green, 2004). Additionally, familiarity with the subject may contribute to the lack of complexity improvement (Tedick, 1990). Additionally, this CLIL method was not designed to increase syntactic complexity. As a result, it is indeed unsurprising that participants demonstrated no increase. The research conducted by Aguilar and Muñoz (2019) shows that the difference in the mean scores in the pre- and post-listening exams was significant, but not for the pre- and post-grammar tests, according to paired- sampled t-tests. A repeated measures ANOVA revealed that less proficient students improved their hearing and

grammatical abilities more than more competent students when students were divided into three groups based on their pre-test scores.

Consistent with Casanave's (1994) findings, participants improved their accuracy over the semester. Although this CLIL class's objective was not on the accuracy, the findings can be attributed to various factors. As a result, their accuracy significantly increased. Concerning the second research question, the findings indicated that participants produced more fluent texts after a semester of using the CLIL method. In other words, their T-units and words were significantly larger in number. A possible explanation for this increase is practice effects, in which repetition of tasks enables learners to produce more fluent output because a portion of conceptualization, formulation, and articulation is retained in memory and can be used the next time they encounter a similar task (Bygate, 1999). As discussed in the theoretical background section above, Whittaker et al. (2011) and Llinares et al. (2012) discovered development in textual resource control as well as an increase in nominal group complexity over four years. Subsequently, the researchers proposed that CLIL settings that place a premium on content learning provide an appropriate context for developing written discourse, as the students in the study appeared to have a solid knowledge base from which to generate their texts. This viewpoint could have been expanded to account for the current study's significant improvement in accuracy.

Similar with Lahuerta (2020) in the setting of CLIL instruction, accuracy improved dramatically with grade, whereas in the non-CLIL group, only lexicogram errors shrank significantly. When considering articless, issues arose in both contexts and prepositions, determiners, voice, subordination, and word order in the non-CLIL context, the examination of mistake subtypes indicated certain cases of regression and stabilization tendencies. In terms of gender, participants found that they performed similarly to their male counterparts in the CLIL group as they did in the non-CLIL group for written accuracy, suggesting that CLIL may help balance gender inequalities. Similar

with the study conducted by Lee (2020) in terms of writing quality, lexical sophistication, lexical diversity, and mean length of clause, the CLIL group beat the non-CLIL group; nevertheless, the non-CLIL group's writing displayed higher semantic cohesiveness. The findings contribute to CLIL research by extending it to examine CLIL effects on syntactic complexity, lexical complexity, and cohesiveness in addition to writing quality. Previously, CLIL research was restricted to a focus on morphosyntactic aspects or writing quality.

In this study participants might prefer the native English grammatical norm. He and Zhang (2010) conducted a large survey sample of Chinese students and discovered that most respondents preferred Standard English for grammar. Additionally, Kirkpatrick and Zhichang (2002) discovered significant attitude differences between English and other majors, with female English majors being more likely to pursue Standard English. Thus, more than half of the participants in this study were female English majors; their attitude may have contributed to an improvement in accuracy. Third, the assessment criteria used in this class may affect their improved accuracy. The assessment criteria stated that they would be graded on their organization and should avoid obvious grammatical errors such as misspelled words and improper punctuation that impede communication. On the one hand, this result was unexpected, given that prior research in the CLIL literature indicated that the greatest gains in oral production typically occur in lexis and pragmatic features. The findings, however, are consistent with prior research on CLIL writing, most notably in terms of accuracy (Jexenflicker & Dalton-Puffer, 2010).

CONCLUSION

This study explored students' writing ability in their second languages sequentially to determine which errors are distinct to the second language. In this study, the CLIL and Non-CLIL methods used in the experimental and control groups produced similar accuracy, fluency, syntactic, morphological, lexicogrammatical and

spelling scores. Meanwhile, for complexity and lexical, the CLIL and Non-CLIL methods resulted in significantly different average, with the Non-CLIL method being higher. This study appears to demonstrate that a CLIL method benefits education learners' longitudinal English writing development in terms of complexity, fluency, and accuracy over an academic year. The evidence accumulated thus far indicates that the CLIL method is effective at supplementing formal language instruction by increasing students' introduction to English as a foreign language in an integrative and naturalistic manner. In line with Kormos (2012), hopefully, this study will serve as a springboard for future research that incorporates a larger quantity and a multiple diversity of writing forms.

ACKNOWLEDGMENTS

REFERENCES

Aguilar, M., & Muñoz, C. (2019). The effect of proficiency on CLIL benefits in Engineering students in Spain. *International Journal of Applied Linguistics (United Kingdom)*, 24(1), 1–18. https://doi.org/10.1111/ijal.12006

Arnó-Macià, E., & Mancho-Barés, G. (2015). The role of content and language in content and language integrated learning (CLIL) at university: Challenges and implications for ESP. *English for Specific Purposes*, 37(1), 63–73. https://doi.org/10.1016/j.esp.2014.06.007

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

Bruton, A. (2013). CLIL: Some of the reasons why... and why not. *System*, 41(3), 587–597. https://doi.org/10.1016/j.system.2013.07.001

Bulté, B., & Housen, A. (2012). *Defining and operationalising L2 complexity*. 21–46. https://doi.org/10.1075/LLLT.32.02BUL

Bygate, M. (1999). Quality of language and purpose of task: Patterns of learners' language on two oral communication tasks. *Language Teaching Research*, 3(3), 185–214. https://doi.org/10.1177/136216889900300302

Cabrera Solano, P. A., Gonzalez Torres, P. F., Ochoa Cueva, C. A., Quinonez Beltran, A. L., Castillo Cuesta, L. M., Solano Jaramillo, L. M., Espinosa Jaramillo, F. O., & Arias Cordova, M. O. (2014). Spanish interference in EFL writing skills: A case of Ecuadorian senior high schools. *English Language Teaching*, 7(7), 40–48. https://doi.org/10.5539/elt.v7n7p40

Casanave, C. P. (1994). Language development in students' journals. *Journal of Second Language Writing*, 3(3), 179–201. https://doi.org/10.1016/1060-3743(94)90016-7

Cenoz, J., & Gorter, D. (2014). Focus on multilingualism as an approach in educational contexts. In *Educational Linguistics* (Vol. 20, pp. 239–254). Springer, Dordrecht. https://doi.org/10.1007/978-94-007-7856-6_13

Coyle, D. (2005). *CLIL planning tools for teachers 4cs curriculum guidance 3as lesson planning tool matrix audit tool for tasks & materials*.

Dalton-Puffer, C. (2008). Outcomes and processes in Content and Language Integrated Learning (CLIL): current research from Europe. In *Werner Delanoy and Laurenz Volkmann, (eds.), Future Perspectives for English Language Teaching*. (Issue May, pp. 139–157). Heidelberg: Carl Winter.

Green, T. (2004). Making the grade: Score gains on the IELTS writing test. *Research Notes*, 16(4), 9–3.

He, D., & Zhang, Q. (2010). Native speaker norms and China English: From the perspective of learners and teachers in China. *TESOL Quarterly*, 44(4), 769–789. https://doi.org/10.5054/tq.2010.235995

Heras, A., & Lasagabaster, D. (2015). The impact of CLIL on affective factors and vocabulary learning. *Language Teaching Research*, 19(1), 70–88. https://doi.org/10.1177/1362168814541736

Hunt, K. (1964). Differences in grammatical structures written at

three grade levels. In *NCTE Research Report No. 3*.

Jexenflicker, S., & Dalton-Puffer, C. (2010). The CLIL differential: Comparing the writing of CLIL and non-CLIL students in higher colleges of technology. *Language Use and Language Learning in CLIL Classrooms.*, 7(2010), 169.

Kirkpatrick, A., & Zhichang, X. (2002). Chinese pragmatic norms and 'China English.' *World Englishes*, 21(2), 269–279. https://doi.org/10.1111/1467-971X.00247

Kormos, J. (2012). The role of individual differences in L2 writing. *Journal of Second Language Writing*, 21(4), 390–403. https://doi.org/10.1016/j.jslw.2012.09.003

Krashen, S. D. (1981). Second language acquisition and second language learning. *Language Education, Oxford: Pergamon.* https://doi.org/10.4324/9781003054368-4

Lahuerta, A. (2020). Analysis of accuracy in the writing of EFL students enrolled on CLIL and non-CLIL programmes: the impact of grade and gender. *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., & Sierra, J. M. (2009). Immersion and CLIL in English: More differences than similarities. *ELT Journal*, 64(4), 367–375. https://doi.org/10.1093/elt/ccp082

Lee, J. (2020). Assessing the effects of CLIL on Korean high school students' writing. *Linguistic Research*, 37(Special Edition), 89–112. https://doi.org/10.17250/khisli.37..202009.004

Llinares, A., Morton, T., & whittaker, R. (2012). The Roles of Language in CLIL. *New York: Cambridge University Press.* https://doi.org/10.18823/asiatefl.2017.14.1.18.209

Llinares, Ana, & Whittaker, R. (2006). Linguistic analysis of secondary school students' oral and written production in CLIL contexts: studying social science in English. *VIEWS Vienna English Working Papers*, 15(3), 33–40. https://doi.org/10.16660/j.cnki.1674-098x.2012.23.052

Llinares, Ana, & Whittaker, R. (2007). Talking and Writing in a Foreign Language in CLIL contexts: a Linguistic Analysis of

secondary school learners of geography and history. *Revista Española de Lingüística Aplicada*, *1*, 83–94.

Lyster, R. (2011). Content-based second language teaching. In *E. Hinkel (Ed.), Handbook of Research in Second Language Teaching and Learning, Vol .2* (pp. 611–630). New York :Routledge.

Myles, J. (2002). Second language writing and research: The writing process and error analysis in student texts. *TESL-EJ*, *6*(2).

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.

Pallotti, G. (2009). CAF: Defining, refining and differentiating constructs. *Applied Linguistics*, *30*(4), 590–601. https://doi.org/10.1093/APPLIN/AMP045

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

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

Pérez-Vidal, Carmen, & 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

Rallo Fabra, L., & Jacob, K. (2015). Content-based Language Learning in Multilingual Educational Environments. In M. Juan-Garau & J. Salazar-Noguera (Eds.), *Content-based Language Learning in Multilingual Educational Environments* (Vol. 23, pp. 163–177). Springer. https://doi.org/10.1007/978-3-319-11496-5

Rose, D. (2020). Literacy Education and Systemic Functional Linguistics. In *The Cambridge Introduction to Applied Linguistics* (pp. 115–132). Cambridge University Press. https://doi.org/10.1017/9781108658089.012

Ruiz de Zarobe, Y. (2010). *Written production and CLIL* (pp. 191–210). John Benjamins. https://doi.org/10.1075/aals.7.10rui

Storch, N. (2009). The impact of studying in a second language (L2)

medium university on the development of L2 writing. *Journal of Second Language Writing*, *18*(2), 103–118. https://doi.org/10.1016/j.jslw.2009.02.003

Tedick, D. J. (1990). ESL writing assessment: Subject-matter knowledge and its impact on performance. *English for Specific Purposes*, *9*(2), 123–143. https://doi.org/10.1016/0889-4906(90)90003-U

Thewissen, J. (2013). Capturing L2 Accuracy Developmental Patterns: Insights From an Error-Tagged EFL Learner Corpus. *Modern Language Journal*, *97*(SUPPL.1), 77–101. https://doi.org/10.1111/j.1540-4781.2012.01422.x

Verspoor, M., Schmid, M. S., & Xu, X. (2012). A dynamic usage based perspective on L2 writing. *Journal of Second Language Writing*, *21*(3), 239–263. https://doi.org/10.1016/j.jslw.2012.03.007

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, K. S., & Kim, H. . (1998). Second language development in writing: Measures of fluency, accuracy, and complexity. *University of Hawaii Press*, *19*(2), 225–257. https://doi.org/10.1017/s0142716400010055

Xudong, D., Cheng, L., Varaprasad, C., & Leng, L. (2010). Academic Writing Development of ESL/EFL Graduate Students in NUS. *Nus.Sg*, *9*(2), 119–138