Bridging the Gap: Developing Lexical and Syntactic Complexity Through CLIL After Study Abroad Programmes

Jason Pipe, Kanto Gakuin University, Japan

The Asian Conference on Education 2025
Official Conference Proceedings

Abstract

This study investigates the development of lexical and syntactic complexity in English as a Foreign Language (EFL) learners after returning from a five-month study abroad programme. While previous research suggests that immersion in English-speaking environments enhances fluency and general communicative competence, questions remain regarding its long-term impact on academic writing proficiency once returned to their home countries. Drawing on Content and Language Integrated Learning (CLIL) as a post-study abroad teaching strategy, this research explores whether CLIL instruction can foster continued growth in advanced academic lexis and syntactic structures. Analysing weekly written samples from six Japanese university students collected over one academic year, this paper aimed at observing changes in linguistic complexity. Through quantitative measures, findings revealed no statistically significant or consistent improvements in lexical nor syntactic complexity despite maintaining overall motivation and engagement of students since returning to their home country. These results suggest that while CLIL can provide a meaningful and content-rich context for EFL language use, it may not in itself be sufficient to stretch learning to observe measurable development in academic writing proficiency. The study highlights the need for more explicit instruction, targeted scaffolding and feedback on complex language use within CLIL frameworks, especially for learners transitioning from immersive environments back to domestic academic contexts. It also raises important pedagogical considerations for supporting sustained linguistic development in post-study abroad programmes.

Keywords: CLIL (Content and Language Integrated Learning), lexical complexity, syntactic complexity, academic writing, study abroad

iafor
The International Academic Forum
www.iafor.org

Introduction

Study abroad programmes have always been seen as a feasible option for enhancing second language (L2) proficiency among university students. While these immersive experiences often lead to notable gains in confidence as well as spoken fluency, it remains unclear whether they also equip learners with the linguistic tools required for academic writing in English. Upon returning from study abroad, students may still struggle to bridge the gap between conversational proficiency and the demands of academic discourse.

A key concern in EFL language education is the absence of authentic learning, i.e. activities that mirror real-world scenarios and equip learners with meaningful, transferable knowledge that is applicable in practical contexts (Gilmore, 2007). The attractiveness of foreign exchange programme is that communicative activities can be integrated with authentic learning tasks. This has been shown to enhance learner motivation, provide culturally rich L2 (second language) input, and align second language instruction with real-life communicative contexts (Garcia-Pinar, 2019; Jacobs et al., 2022; Ozverir et al., 2017; Purnawarman & Dajarati, 2020; Vellanki & Bandu, 2021). Such authenticity would seem to be effective as it takes into account the learners' identities, the nature of the learning tasks, and the discourse in which the language is used to effectively communicate in a foreign coiuntry (Ahmadi, 2019). When studying abroad, such authenticity is likely to be embedded in instructional design, leading students to increase thei engagement, improved spoken L2 production, and enhanced knowledge transfer (Gilmore, 2007; Larsen et al., 2017; Perkins & Salomon, 2012; Wang et al., 2012). However, when returning back to their home countries, it is questionable that students can continue to maintain this momentum and develop further their academic skills.

To address this possible stagnation after returning from abroad post-study abroad English programme, a *Content and Language Integrated Learning* (CLIL) approach was implemented into the classroom as a pedagogical approach to integrate EAP subject content while also offering a potential pathway for continued language growth through authentic learning.

Literature Review

Study Abroad

Recent research continues to affirm the benefits of study abroad experiences for L2 development, particularly in enhancing oral fluency, pragmatic awareness, and vocabulary breadth (Friginal & Polat, 2022; Reynolds, 2020). Immersion in an L2 environment exposes learners to authentic communicative contexts and varied input, which can significantly improve their confidence and spontaneous language production. These gains are especially notable in functional spoken language and sociolinguistic competence (Bown & Dewey, 2020) as well as gains in fluency, pragmatic competence, and vocabulary breadth (Freed, 1995; Kinginger, 2009).

However, these benefits are not uniformly observed across all linguistic domains. While spoken interaction often improves, gains in academic language proficiency, particularly in academic writing and research, are less consistent (Pérez-Vidal & Barquin, 2014). This is due to study abroad programs catering in developing confidence, emphasising informal communication, and encouraging cultural engagement. These programmes, however, offer limited opportunities for students to engage with academic genres, advanced syntactic structures, or discipline-specific lexis (Lee & Curdt, 2021) as these students are rarely included

in truly academic and authentic enviuronments with native English speakers. As a consequence, students may return with strong conversational fluency but remain underprepared for the true linguistic demands of academic writing in the L2 (Taguchi & Greer, 2019).

Recently, there has been a small but growing body of research that would seem to suggest the necessity in providing post-return support. This should be an essential component for consolidating and hopefully extending the language gains made during study abroad. It would be beneficial for Japanese university students who have returned from their short-term programs abroad, showing improvements in fluency and lexical diversity in speech, but only marginal gains in written complexity (Imai & Kaneko, 2021). A study by Lasagabaster also observed a similar situation with Taiwanese students who studied abroad in English-speaking countries. While they significantly improved in interactive oral skills, they still required academic writing courses to develop their syntactic range and academic vocabulary further (Lasagabaster, 2021). It is evident that without explicit instruction in academic literacy, students may struggle to transfer their communicative gains to contexts that require higher-order cognitive and linguistic skills (Strobl & Baten, 2021). For instance, students' writing complexity, organisation, and vocabulary depth could continue to remain underdeveloped after studying abroad unless structured writing support is provided (Marijuan & Sanz, 2020; Strobl & Baten, 2021). Post-return programs that incorporate academic writing and critical thinking could, therefore, play a crucial role in helping learners internalise and transfer their linguistic gains into formal educational or professional contexts (Crossley & Kyle, 2021). As a result, while study abroad can offer rich, authentic input within immersive experiences to naturally enable L2 development, its impact on academic literacy remains limited without supplementary pedagogical support.

Content and Language Integrated Learning

Although often referred to as methodology, CLIL is not a method in the traditional pedagogical sense. It lacks a fixed set of teaching procedures or a universally agreed-upon implementation model. Rather, CLIL is a flexible educational approach that emerged in Europe in the 1990s, referring to programs in which a second or foreign language is used to teach subject content to learners who are still developing proficiency in that language (Coyle et al., 2010). CLIL is an increasingly influential approach in multilingual education that integrates the learning of academic subject matter with the simultaneous development of a second or foreign language (Coyle et al., 2010; Mehisto et l., 2019). In CLIL contexts, students learn curricular content, such as history, science, or the arts, through a language they are still acquiring.

This dual focus enables learners to use language for meaningful communication while also developing conceptual knowledge. CLIL operates as a dual-focused approach that integrates both content and language learning. While the emphasis may shift between content and language, the two are always interwoven. However, it is important to add that while content-based instruction teaches language through content, CLIL is subtly nuanced in that it focuses on teaching content through language (Echevarria et al., 2017; Lopriore, 2021). In other words, CLIL address the gap by shifting away from teaching language as an isolated subject to a language tool for actual learning (Morton, 2020). Its central aim is to promote meaningful language use in cognitively engaging contexts, thereby fostering deeper learning and broader communicative competence. By integrating language and content, CLIL encourages students to acquire disciplinary knowledge while developing the linguistic tools needed to discuss, analyse, and apply that knowledge in the target language. Such an approach could also have

the potential benefit of mirroring some of the positive aspects of studying abroad and thereby naturally continuing on from where the study abroad programme left off.

Empirical studies confirm that CLIL positively influences aspects of learning felt on foreign language exchanes. These EFL linguistic skills include: receptive skills, vocabulary acquisition, morphological development, creativity, fluency, and learner motivation (Dalton-Puffer, 2008). Students can also continue develop a broader range of academic and technical vocabulary than their peers in other traditional EFL courses. However, despite these obvious benefits, CLIL also appears to have minimal impact on other important aspects of second language acquistion, namely syntax, informal writing, pronunciation, and pragmatics (Table 1). In this research, the course wil adopt CLIL to teach aspects of social studies and other mumanity subjects in English within an EAP contect. Students will hopefully continue to engage in tasks such as discussions and mini-lectures to develop critical thinking and cross-cultural understanding from both British and Japanese perspectives in this EAP setting.

Table 1
Impact of CLIL on Specific EFL Linguistic Skills

Unaffected or Indefinite	Favorably Affected
Syntax	Receptive skills
Informal/non-technical writing	Vocabulary
Pronunciation	Morphology, creativity, risk-taking, fluency
Pragmatics	Emotive/affective outcomes

Lexical and Syntactic Complexity

Lexical and syntactic complexity are widely recognised as key indicators of advanced L2 writing proficiency. These constructs have, in fact, been necessary in assessing and supporting L2 writing development (Yang et al., 2023). Lexical complexity plays a particularly important role in academic writing and is commonly considered an indicator of L2 proficiency and development (Bulté & Housen, 2021). It refers to the richness, variety, and sophistication of vocabulary use, encompassing dimensions such as lexical diversity, lexical density, and word frequency (Lu & Ai, 2015). However, despite its importance, vocabulary acquisition and productive lexical use often receive less explicit attention in Japanese EFL education, particularly in terms of supporting students' academic English development (Paquot & Granger, 2020).

Syntactic complexity, in contrast, reflects the structural variety and grammatical elaboration in writing, including the use of subordinate clauses, compound sentences, passive constructions, and nominalisations (Kyle & Crossley,2019). Syntactic complexity has long been recognised as a key indicator of L2 proficiency, particularly in the context of academic writing. It refers to the range and sophistication of grammatical structures used by a learner and is widely employed in SLA research as a robust metric for evaluating linguistic development over time (Bulté & Housen, 2021; Vyatkina, 2012). When assessing syntactic development in written discourse, especially among university-level English learners, syntactic complexity provides critical insight into learners' ability to produce structurally elaborate and hierarchically organised sentences (Tavakoli & Wright, 2020).

Both types of complexity are essential elements in academic writing competence as they contribute to coherence, precision, and clarity. For instance, more sophisticated academic

writing is characterized by a high degree of lexical and syntactic complexity. It has also become evident that improvements in lexical sophistication, such as the use of low-frequency or academic words, can lead to improvements in reading fluency and writing quality in general (Paquot & Granger, 2020). Similarly, increases in syntactic complexity are associated with higher scores in holistic writing assessments, particularly in academic or argumentative tasks (Staples & Biber, 2021a).

Research Questions

Q1. Will the level of lexical complexity improve as a direct result of the introduction of CLIL?
Q2. Will the level of syntactic complexity improve as a direct result of the introduction of CLIL?

Methodology

Participants

Participants were six third-year students from a private university in Tokyo participated in this study. Each had recently returned from a five-month study abroad programme in Australia. TOEIC scores ranged from 660 to 800 before departure (CEFR: B1-B2) and from 750 to 835 after returning (CEFR: upper B2 to lower C1). Informal interviews indicated increased confidence and motivation to improve academic English.

CLIL Lessons

Over the academic year, students attended a weekly 90-minute CLIL class focusing on topics in cultural and social studies. CLIL classrooms often prioritise content delivery over language scaffolding, so learners may find few opportunities to develop grammatical accuracy and syntactic complexity (Ibarrola & Lasagabaster, 2021). Although CLIL's success depends on careful planning and responsiveness to learners' needs, time constraints meant that, in this course, lessons mainly focused on teaching content through English with minimal language-focused intervention. As a result, there was only limited explicit attention to the mechanics of language form.

Data Collection

Writing samples were collected bi-weekly over one academic year during a CLIL-based EAP course. Students were asked to write spontaniously for a 10-minute period at the end of class their opinions on the discussions introduced. Topics related to social studies and humanities, and were selected to promote critical thinking and expression.

Measuring Lexical Complexity

To assess lexical complexity in this study, we adopted Lu's (2012) Lexical Complexity Analyzer (LCA), a widely validated tool that automatically calculates 26 lexical indices across all three dimensions (Lu, 2012). This tool was chosen for its transparency, accessibility, and alignment with prior research in L2 lexical development. Due to expediency, lexical complexity commonly followed categorised the three distinct but related dimensions: lexical diversity, lexical sophistication, and lexical density (Yang et al., 2023) (Table 2). First, lexical

diversity refers to the range of different words used in a given text, often measured using indices such as the type-token ratio (TTR), and corrected TTR (CTTR) (Malvern et al., 2020). These metrics provide insight into the breadth of vocabulary employed by the learner and are widely used in studies of L2 academic writing. Second, lexical sophistication involves the use of low-frequency, academic, or otherwise advanced words, indicating the depth of a learner's vocabulary (Crossley & Kyle, 2021). Measures such as the proportion of words from the Academic Word List (AWL) and low-frequency word ratio are commonly employed to assess lexical sophistication. These measures are particularly relevant in evaluating learners' ability to engage with academic discourse, as they capture the extent to which students employ domain-specific or context-appropriate lexis (Staples et al., 2020). This paper used lexical sophistication (LS1) and verv sophistication (VS1) to determine this aspect of lexical complexity. Finally, lexical density, in contrast, assesses the informational load of a text by calculating the proportion of content words (nouns, verbs, adjectives, adverbs) to total words (Schmitt et al., 2019). Higher lexical density is generally associated with more formal or academic registers, distinguishing it from conversational or narrative writing. As a consequence, the measure lexical density (LD) was also included in this research.

Table 2

Dimensions of Lexical Complexity, Their Common Measures and Their Academic Relvance

Category	Definition	Typical Measures	Academic Relevance
Lexical Diversity	Range of different words used in a text. Reflects vocabulary breadth.	- Type-Token Ratio (TTR) - Corrected TTR (CTTR)	Indicates vocabulary range and learner variability; commonly used to assess productive vocabulary size. (Yang et al., 2023)
Lexical Sophist-ication	Use of low-frequency, academic, or advanced vocabulary items.	- Proportion of Academic Word List (AWL) words - Low-frequency word ratio such as Lexical Sophistication I (LS1); Verb Sophistication I (VS1)	Reflects lexical depth and learner ability to use domain-specific or contextually appropriate vocabulary. (Crossley & Kyle, 2021)
Lexical Density	Ratio of content words (nouns, verbs, adjectives, adverbs) to all words in the text.	- Lexical Density (LD) (Content Words / Total Words)	Indicates information load and formality; distinguishes academic from conversational writing. (Schmitt et al., 2019)

Measuring Syntactic Complexity

This study adopted Lu's L2 Syntactic Complexity Analyzer (L2SCA) to evaluate syntactic complexity across 14 established indices (Lu, 2010). Its reliability, accessibility, and consistency with corpus-based norms have made it one of the most widely used tools in academic writing research (Lu & Ai, 2015). This study focuses on nine syntactic measures that recent research has consistently found to be both developmentally sensitive and pedagogically relevant in academic contexts (Table 3). These measures are grouped into three primary categories: First are the length-based measures: Mean Length of Sentence (MLS); Mean Length of T-unit (MLT); Mean Length of Clause (MLC). These indices represent overall

grammatical elaboration and have shown significant correlation with writing proficiency in L2 learners, especially in argumentative and expository tasks (Knoch & Elder, 2017; Kyle & Crossley, 2019). The second group of measures are the subordination-based measures: Clauses per Sentence (C/S); Dependent Clauses per Clause (DC/C); Complex T-units per T-unit (CT/T). Subordination reflects learners' ability to build complex syntactic hierarchies, a characteristic of advanced academic writing. These measures are particularly effective in distinguishing between intermediate and advanced levels of syntactic development (Crossley & Kyle, 2020). The final measures concentrates on phrasal complexity: Complex Nominals per Clause (CN/C); Complex Nominals per T-unit (CN/T); Verb Phrases per T-unit (VP/T). These indices measure nominal and verbal elaboration, which are increasingly seen as indicators of mature academic style, especially in scientific and technical writing. Phrasal complexity has recently been recognised as a key developmental feature in academic writing that complements traditional clause-based complexity (Parkinson & Musgrave, 2014). Overall, this selection aligns with validated frameworks for assessing syntactic complexity in academic L2 writing (Staples & Biber, 2021b).

Table 3
Dimensions of Syntactic Complexity, Their Common Measures and Their Academic Relevance

Category	Measure	Definition	Academic Relevance
Length-Based	Mean Length of Sentence (MLS)	Average number of words per sentence.	Indicates overall grammatical elaboration; associated with advanced writing proficiency. (Kyle & Crossley, 2019)
	Mean Length of T-unit (MLT)	Average number of words per T-unit (an independent clause with any subordinate clauses).	Strongly correlated with writing fluency and syntactic maturity in academic genres. (Kyle & Crossley, 2019)
	Mean Length of Clause (MLC)	Average number of words per clause.	Reflects how learners elaborate on ideas within sentence-level grammar.(Kyle & Crossley, 2019)
Subordination-Based	Clauses per Sentence(C/S)	Average number of clauses per Sentence.	Captures sentence complexity through subordinate clause use. (Crossley & Kyle, 2020)
	Dependent Clauses per Clause (DC/C)	Ratio of dependent to total clauses.	Measures hierarchical sentence structure, signalling higher writing proficiency. (Crossley & Kyle, 2020)
	Complex T-units per T-unit (CT/T)	Proportion of T-units containing more than one clause.	Useful in identifying syntactic sophistication in expository and argumentative writing.(Crossley & Kyle, 2020)
Phrasal Complexity	Complex Nominals per Clause (CN/C)	Number of noun phrases with modifiers per clause.	Reflects nominal elaboration typical of academic writing style. (Parkinson & Musgrave, 2014)

	Complex Nominals per T-unit (CN/T)	Number of complex noun phrases per T-unit.	Associated with increased density and formality in academic prose. (Parkinson & Musgrave, 2020)
	Verb Phrases per T-unit (VP/T)	Number of verb phrases per T-unit.	Captures verbal complexity and helps identify developmental differences in syntactic use. (Parkinson & Musgrave, 2020)

Results

Lexical

Lexical Diversity

The mean TTR across the term was 0.63 (SD = 0.06), indicating a moderate range of lexical variety in student writing. A repeated-measures ANOVA revealed a significant difference across the 14 time points, $F(13, 182) = 230.43$, $p < .001$, $\eta^2 = 0.943$, suggesting that nearly 94% of the variation in TTR was linked to changes over time. However, a correlation analysis showed a non-significant downward trend ($r = –0.396$, $p = 0.161$), implying that while variation occurred, lexical diversity did not steadily increase.

Similarly, CTTR, which adjusts for text length, averaged 4.56 (SD = 0.39) over the 14 weeks. This measure also demonstrated significant differences over time, $F(13, 182) = 132.81$, $p < .001$, $\eta^2 = 0.905$, with a very large effect size. Yet again, correlation analysis revealed a non-significant decline ($r = ç0.282$, $p = 0.329$), suggesting that although CTTR values fluctuated, no consistent upward trajectory was observed.

Table 4

Lexical Diversity: Type-Token Ratio (TTR) and Corrected Type-Token Ratio (CTTR)

TTR	1	2	3	4	5	6	7	8	9	10	11	12	13	14
Mean	**0.63**	**0.66**	**0.60**	**0.62**	**0.61**	**0.69**	**0.69**	**0.67**	**0.63**	**0.59**	**0.60**	**0.63**	**0.63**	**0.55**
SD	0.10	0.06	0.04	0.06	0.06	0.06	0.05	0.05	0.06	0.07	0.09	0.07	0.06	0.07
Max	0.77	0.71	0.66	0.71	0.70	0.76	0.72	0.74	0.71	0.70	0.72	0.73	0.68	0.64
Min	0.52	0.57	0.56	0.55	0.54	0.60	0.61	0.60	0.55	0.52	0.49	0.54	0.56	0.46
CTTR														
Mean	**4.63**	**4.59**	**4.56**	**4.58**	**4.41**	**4.83**	**4.54**	**4.47**	**4.73**	**4.56**	**4.49**	**4.74**	**4.50**	**4.27**
SD	0.34	0.62	0.31	0.16	0.07	0.19	0.54	0.38	0.55	0.65	0.45	0.31	0.53	0.31
Max	5.01	5.22	4.97	4.79	4.51	5.07	5.03	5.10	5.62	5.54	4.99	5.10	4.92	4.61
Min	4.10	3.83	4.24	4.36	4.35	4.53	3.79	3.97	4.25	4.03	3.97	4.27	3.67	3.84

Taken together, the results indicate that while both TTR and CTTR varied significantly across time, neither measure showed sustained improvement. This may reflect the cognitive and linguistic demands of CLIL content, where students prioritised content comprehension and grammatical accuracy over increasing lexical variety.

Lexical Sophistication

Table 5

Lexical Sophistication: Lexical Sophistication I (LS1) and Verb Sophistication I (VS1)

LS1	1	2	3	4	5	6	7	8	9	10	11	12	13	14
Mean	**0.172**	**0.303**	**0.188**	**0.205**	**0.162**	**0.365**	**0.243**	**0.215**	**0.318**	**0.354**	**0.134**	**0.175**	**0.330**	**0.318**
SD	0.066	0.031	0.038	0.062	0.041	0.114	0.117	0.048	0.053	0.061	0.025	0.038	0.048	0.059
Max	0.244	0.343	0.222	0.258	0.225	0.473	0.354	0.278	0.357	0.438	0.176	0.217	0.404	0.386
Min	0.081	0.265	0.145	0.098	0.118	0.196	0.091	0.147	0.226	0.293	0.115	0.118	0.288	0.235
VS1														
Mean	**0.047**	**0.036**	**0.025**	**0.092**	**0.071**	**0.067**	**0.066**	**0.080**	**0.047**	**0.052**	**0.092**	**0.075**	**0.108**	**0.029**
SD	0.044	0.065	0.034	0.094	0.041	0.042	0.052	0.074	0.044	0.050	0.041	0.060	0.050	0.030
Max	0.087	0.150	0.062	0.250	0.100	0.105	0.111	0.174	0.087	0.105	0.133	0.148	0.174	0.071
Min	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.036	0.000	0.053	0.000

The average LS1 score was 0.25 (SD = 0.06), representing the proportion of low-frequency or academic vocabulary used in student writing. A repeated-measures ANOVA showed a highly significant difference across time, $F(13, 182) = 962.69$, $p < .001$, $\eta^2 = 0.986$, indicating nearly all variance in LS1 was accounted for by time. A correlation analysis revealed a non-significant positive trend ($r = 0.276$, $p = 0.340$), suggesting some upward movement in lexical sophistication, but without statistical confirmation of consistent growth.

VS1, representing verb-related lexical sophistication, had a mean of 0.06 (SD = 0.05). ANOVA results again indicated significant variation across time, $F(13, 182) = 109.76$, $p < .001$, $\eta^2 = 0.887$. Correlation analysis showed a non-significant positive trend ($r = 0.315$, $p = 0.272$), implying that students gradually increased their use of advanced or less frequent verbs, though not at a statistically significant rate.

In summary, both LS1 and VS1 demonstrated significant fluctuations across time, with very large effect sizes, yet neither showed a steady linear improvement. These results indicate that while CLIL instruction may encourage more sophisticated lexical choices, especially in verb use, gains may be topic-dependent or uneven in progression.

Lexical Density

Table 6

Lexical Density

LD	1	2	3	4	5	6	7	8	9	10	11	12	13	14
Mean	**0.571**	**0.600**	**0.600**	**0.585**	**0.589**	**0.597**	**0.599**	**0.550**	**0.581**	**0.618**	**0.544**	**0.578**	**0.579**	**0.597**
SD	0.049	0.030	0.024	0.064	0.047	0.043	0.040	0.086	0.017	0.017	0.037	0.060	0.045	0.041
Max	0.625	0.642	0.638	0.673	0.639	0.667	0.629	0.653	0.600	0.644	0.593	0.659	0.626	0.663
Min	0.516	0.571	0.579	0.500	0.514	0.560	0.541	0.447	0.559	0.595	0.500	0.510	0.509	0.566

The data showed no significant consistent improvements in lexical or syntactic complexity at the group level. However, individual differences were observed, suggesting that learners responded differently to the instructional context. The most notable finding was a strong negative correlation between lexical sophistication (LS1) and verb phrases per T-unit (VPT), $r = –0.628$, $p < .001$. Moderate negative correlations were also observed between LS1 and DCT ($r = ç0.450$) and LS1 and MLT ($r = –0.295$).

These patterns further support the idea that increased lexical complexity may coincide with reduced syntactic elaboration, likely due to cognitive constraints in managing both dimensions simultaneously. These findings align with the Limited Attentional Capacity Model (Skehan, 2009), which posits that learners may prioritise either lexical sophistication or syntactic complexity depending on task demands and available cognitive resources.

Syntax

Length-Based

Table 7

Length-Based Measures: Mean Length of Sentence (MLS), Mean Length of T-unit (MLT), Mean Length of Clause (MLC)

MLS	1	2	3	4	5	6	7	8	9	10	11	12	13	14
Mean	**15.89**	**15.45**	**17.55**	**14.88**	**15.12**	**15.06**	**18.14**	**16.31**	**15.45**	**16.79**	**18.80**	**16.99**	**15.60**	**12.48**
SD	4.44	3.38	4.21	3.34	1.83	2.20	7.58	1.67	1.70	0.82	6.93	2.58	3.17	1.48
Max	21.00	21.00	21.25	19.86	17.25	18.00	28.33	19.33	18.14	17.50	29.00	20.00	18.83	15.10
Min	9.91	12.14	11.00	11.89	13.14	12.13	11.25	14.80	14.00	15.63	12.14	13.67	11.11	11.50
MLT														
Mean	**14.04**	**13.33**	**15.73**	**13.31**	**13.42**	**12.23**	**13.23**	**13.97**	**13.63**	**13.76**	**16.76**	**15.17**	**11.85**	**10.93**
SD	4.27	1.53	2.97	2.60	1.94	2.00	5.36	2.04	0.60	0.77	7.21	2.23	2.40	1.32
Max	21.00	15.75	18.14	16.57	15.43	15.17	21.25	16.57	14.20	15.00	29.00	18.75	15.29	12.58
Min	9.91	12.00	11.00	10.38	11.50	9.80	10.00	10.67	12.73	13.00	10.83	13.00	9.09	9.00
MLC														
Mean	**8.57**	**8.65**	**8.51**	**8.48**	**7.34**	**8.83**	**6.96**	**8.06**	**11.38**	**10.50**	**9.70**	**7.90**	**8.19**	**8.47**
SD	1.12	2.17	2.19	0.30	0.92	1.27	0.88	1.60	1.76	1.70	3.09	0.72	0.61	1.08
Max	9.69	11.14	12.14	8.91	8.36	10.18	7.73	10.57	13.71	13.00	14.50	9.11	9.11	10.22
Min	6.81	6.30	6.50	8.23	6.57	7.00	6.00	6.17	9.57	8.75	6.55	7.32	7.53	7.31

This study investigated syntactic development in L2 writing following a study abroad experience, using three measures of complexity: Mean Length of Sentence (MLS), Mean Length of T-unit (MLT), and Mean Length of Clause (MLC). Data were collected over 14 CLIL-based instructional sessions, and changes across time were analysed using repeated-measures ANOVA, linear regression, and correlation analysis.

The analysis of MLS revealed an average sentence length of 16.04 words (SD = 3.24). A repeated-measures ANOVA showed that the variation in sentence length across time was statistically significant, $F(13, 182) = 36.09$, $p < .001$, with a large effect size ($\eta^2 = 0.720$). This indicates that sentence complexity fluctuated meaningfully over the course. However, a linear regression showed only a slight upward trend, and the correlation between MLS and time was not statistically significant ($r = –0.073$, $p = 0.805$). While sentence length varied across the 14 sessions, there was no consistent linear increase in complexity.

For MLT, the average T-unit length was 13.67 words (SD = 2.66). ANOVA results indicated significant changes across the time points, $F(13, 182) = 2.37 \times 10^{29}$, $p < .001$, with an extremely large effect size ($\eta^2 = 1.000$). Despite this, the linear regression suggested a slight downward trend, and correlation analysis confirmed no significant relationship with time ($r = –0.189$, $p = 0.519$). These findings suggest that while structural changes occurred in T-unit length, they did

not reflect a steady increase in syntactic elaboration; rather, students may have opted for more concise structuring over time.

In contrast, MLC showed more promising indicators of development. The average clause length was 8.68 words (SD = 1.39), and this measure exhibited the lowest variability of the three, suggesting relative stability in clause construction. Repeated-measures ANOVA confirmed a significant difference across time, $F(13, 182) = 98.53$, $p < .001$, with a very large effect size ($\eta^2 = 0.876$). A linear regression showed a gentle upward trend in clause length, and while the correlation with time was not statistically significant ($r = 0.166$, $p = 0.571$), the direction of the trend points to a gradual increase in clause-level syntactic complexity.

In summary, while all three measures (MLS, MLT, and MLC) demonstrated statistically significant variation over the course of the programme, only MLC showed a clear and sustained trend toward greater complexity. This suggests that CLIL-based instruction may be particularly effective in supporting internal clause development, even if broader sentence- or T-unit-level complexity does not progress in a strictly linear fashion.

Subordination-Based

Table 8

Subordination-Based Measures: Clauses per Sentence (C/S), Dependent Clauses per Clause (DC/C), Complex T-units per T-unit (CT/T)

C/S	1	2	3	4	5	6	7	8	9	10	11	12	13	14
Mean	**1.83**	**1.94**	**2.10**	**1.75**	**2.10**	**1.74**	**2.55**	**2.08**	**1.37**	**1.64**	**2.01**	**2.18**	**1.93**	**1.49**
SD	0.33	0.87	0.48	0.38	0.43	0.37	0.81	0.39	0.15	0.32	0.77	0.43	0.49	0.25
Max	2.17	3.33	2.83	2.29	2.63	2.20	3.67	2.40	1.56	2.00	3.33	2.50	2.50	1.80
Min	1.46	1.14	1.69	1.38	1.57	1.25	1.75	1.40	1.17	1.25	1.43	1.50	1.33	1.13
DC/C														
Mean	**0.37**	**0.32**	**0.45**	**0.34**	**0.44**	**0.27**	**0.39**	**0.42**	**0.17**	**0.23**	**0.39**	**0.43**	**0.24**	**0.24**
SD	0.12	0.18	0.13	0.12	0.11	0.13	0.07	0.11	0.12	0.16	0.17	0.10	0.14	0.09
Max	0.54	0.50	0.59	0.50	0.57	0.40	0.46	0.58	0.29	0.42	0.60	0.60	0.40	0.33
Min	0.23	0.13	0.29	0.20	0.27	0.09	0.30	0.29	0.00	0.00	0.20	0.33	0.08	0.11
CT/T														
Mean	**0.61**	**0.35**	**0.63**	**0.46**	**0.62**	**0.38**	**0.62**	**0.58**	**0.16**	**0.24**	**0.55**	**0.57**	**0.38**	**0.24**
SD	0.28	0.20	0.28	0.20	0.28	0.25	0.30	0.22	0.11	0.18	0.28	0.13	0.25	0.11
Max	1.00	0.58	1.00	0.78	0.86	0.60	1.00	1.00	0.30	0.43	0.88	0.75	0.71	0.33
Min	0.30	0.14	0.40	0.25	0.25	0.10	0.33	0.40	0.00	0.00	0.25	0.43	0.09	0.06

To examine development in subordinate clause usage, three measures were analysed over the 14-week CLIL-based course: Clauses per Sentence (C/S), Dependent Clauses per Clause (DC/C), and Complex T-units per T-unit (CT/T).

The mean number of clauses per sentence (C/S) was 1.91 (SD = 0.30), indicating moderately complex sentence constructions. A repeated-measures ANOVA revealed a statistically significant difference across the 14 time points, $F(13, 182) = 167.98$, $p < .001$, $\eta^2 = 0.923$, suggesting that over 92% of the variance in clause use per sentence was attributable to instructional time. Linear regression analysis showed a noticeable upward trend, indicating that students increasingly produced multi-clause sentences as the term progressed.

The proportion of dependent clauses per clause (DC/C) averaged 0.34 (SD = 0.13). ANOVA results confirmed significant variation across time, $F(13, 182) = 55.14$, $p < .001$, $\eta^2 = 0.798$. Although the regression trend was subtler than that of C/S, it was still upward, suggesting a gradual increase in learners' use of subordination strategies over time.

The ratio of complex T-units per T-unit (CT/T) was 0.46 on average (SD = 0.22), reflecting learners' integration of subordinate structures into complete thought units. ANOVA again revealed significant changes over time, $F(13, 182) = 38.76$, $p < .001$, $\eta^2 = 0.735$, with a moderate upward regression trend. These results indicate increasing syntactic maturity in how students structured and extended their ideas.

Taken together, all three subordination-based measures demonstrated statistically significant variation across time and positive developmental trends. These findings suggest that CLIL instruction following study abroad may support learners in advancing their use of embedded and hierarchical sentence structures—an essential aspect of academic writing fluency.

Phrasal Complexity

Table 9

Phrasal Complexity: Complex Nominals per Clause (CN/C); Complex Nominals per T-unit (CN/T); Verb Phrases per T-unit (VP/T)

CN/C	1	2	3	4	5	6	7	8	9	10	11	12	13	14
Mean	**0.75**	**1.03**	**1.09**	**0.72**	**0.62**	**0.69**	**0.81**	**0.68**	**1.13**	**1.23**	**0.70**	**0.80**	**1.05**	**0.93**
SD	0.13	0.34	0.43	0.17	0.24	0.13	0.22	0.27	0.36	0.14	0.28	0.49	0.28	0.36
Max	0.89	1.57	1.71	0.94	1.00	0.90	1.09	1.14	1.57	1.36	1.13	1.67	1.56	1.33
Min	0.62	0.75	0.68	0.55	0.42	0.57	0.63	0.42	0.79	1.08	0.50	0.50	0.87	0.62
CN/T														
Mean	**1.24**	**1.62**	**2.01**	**1.13**	**1.13**	**0.98**	**1.60**	**1.15**	**1.37**	**1.65**	**1.29**	**1.46**	**1.52**	**1.18**
SD	0.39	0.44	0.71	0.34	0.38	0.34	0.94	0.32	0.43	0.38	0.82	0.63	0.43	0.40
Max	1.67	2.00	3.00	1.67	1.67	1.50	3.00	1.60	2.00	2.29	2.25	2.50	2.00	1.69
Min	0.80	0.86	1.15	0.75	0.83	0.70	1.00	0.71	0.89	1.30	0.67	0.86	1.00	0.75
VP/T														
Mean	**2.43**	**1.91**	**2.87**	**2.35**	**2.57**	**1.81**	**2.03**	**2.21**	**1.78**	**1.49**	**2.73**	**2.49**	**1.73**	**1.51**
SD	0.47	0.48	0.59	0.67	0.30	0.21	0.66	0.41	0.33	0.14	0.59	0.46	0.52	0.25
Max	3.17	2.50	3.86	3.11	2.86	2.00	3.00	2.86	2.10	1.71	3.50	3.13	2.43	1.75
Min	2.00	1.29	2.44	1.63	2.13	1.50	1.57	1.67	1.29	1.33	2.00	1.83	1.27	1.13

This section reports the development of phrasal complexity over 14 CLIL-based sessions, using three key measures: Complex Nominals per Clause (CN/C), Complex Nominals per T-unit (CN/T), and Verb Phrases per T-unit (VP/T).

The mean value of complex nominals per clause (CN/C) was 0.70 (SD = 0.17), reflecting a moderate density of noun phrase elaboration at the clause level. A repeated-measures ANOVA revealed a statistically significant difference across time points, $F(13, 182) = 23.02$, $p < .001$, $\eta^2 = 0.622$, indicating that nearly two-thirds of the variance was attributable to instructional changes. However, correlation analysis showed no significant trend across time ($r = 0.089$, $p = 0.763$), suggesting that while variability occurred, it did not consistently increase or decrease.

The complexity of noun phrases at the T-unit level (CN/T) averaged 1.25 (SD = 0.28), indicating frequent use of extended noun constructions. The ANOVA results showed significant differences across the term, $F(13, 182) = 15.94$, $p < .001$, $\eta^2 = 0.533$, with a medium-to-large effect size. The Pearson correlation showed a weak negative trend ($r = –0.271$, $p = 0.345$), indicating that although T-units were often complex, their elaboration did not increase over time and may have slightly decreased.

Verb phrase density, measured by verb phrases per T-unit (VP/T), yielded an average of 1.79 (SD = 0.25). The ANOVA showed significant differences across the 14 time points, $F(13, 182) = 8.97$, $p < .001$, $\eta^2 = 0.391$, suggesting moderate variability. However, the trend analysis showed a weak, non-significant positive trend ($r = 0.142$, $p = 0.631$), implying that learners may have incrementally increased their use of verb phrases, though not in a statistically meaningful way.

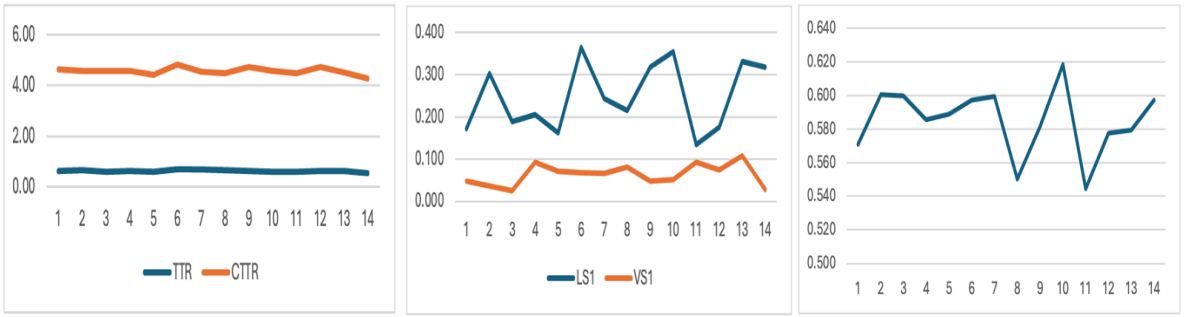
In summary, while all three phrasal complexity measures showed statistically significant changes across time, none displayed a consistent linear trend. These findings indicate that although CLIL instruction may foster structural variation and phrasal elaboration, such development does not necessarily progress in a steady or upward direction.

Discussion

Q1. Will the level of lexical complexity improve as a direct result of the introduction of CLIL?

Graphs 1–3

Lexical: 1. Diversity (left), 2. Sophistication (centre), 3. Density (right)



The results revealed statistically significant variation over time in all measured aspects of lexical complexity. Lexical diversity (TTR, CTTR) (Graph 1), lexical sophistication (LS1, VS1) (Graph 2), and lexical density (LD) (Graph 3) show no consistent upward trend. For example, while LS1 and VS1 showed positive correlations with time, these trends were not statistically significant. Similarly, TTR and CTTR showed slight downward trends, and LD remained relatively flat. These findings suggest that CLIL alone does not lead to sustained improvements in lexical complexity.

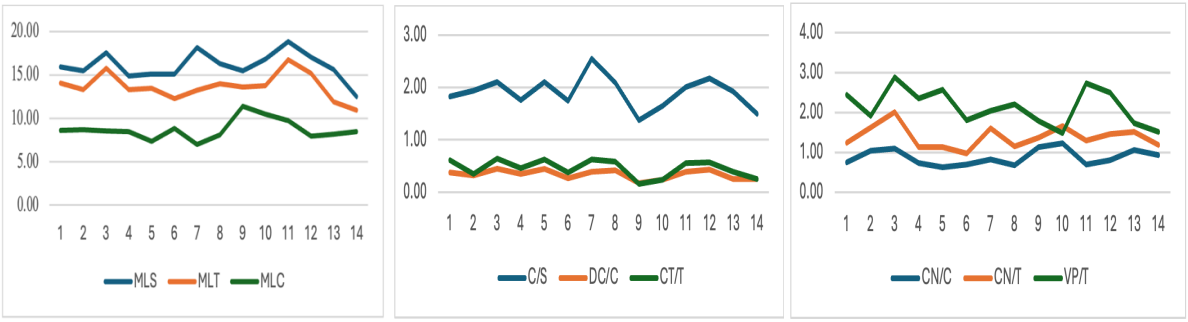
These outcomes suggest that while CLIL may foster variability and topic-related fluctuations in lexical usage, it does not in itself ensure steady gains in academic lexical development. This may be due to the cognitively demanding nature of CLIL tasks, which place pressure on learners' processing capacity. Learners may simplify vocabulary in favour of maintaining content coherence, illustrating the often-cited lexico-syntactic trade-off (Révész et al., 2020). Furthermore, although CLIL encourages exposure to domain-relevant texts, it often lacks

explicit focus on academic vocabulary development (Paquot & Granger, 2020), and students may struggle to internalise and actively produce advanced lexis without scaffolded vocabulary tasks. These findings align with previous studies suggesting that post-return academic language instruction must be explicitly targeted, particularly if lexical development is to be fostered (Crossley & Kyle, 2021; Lasagabaster, 2021).

Q2. Will the level of syntactic complexity improve as a direct result of the introduction of CLIL?

Graphs 4–6

Syntactic: 4. Length-Based (Left), 5. Subordination-Based (Centre), 6. Phrasal Complexity (Right)



The introduction of CLIL did not result in overall improvement across all aspects of syntactic complexity. Focusing on Length-Based Measures (Graph 4) , Mean Length of Sentence (MLS), T-unit (MLT), and Clause (MLC) indicated no consistent upward trend. In fact, by the end of the academic year, MLS and MLT values became slightly lower than at the beginning of the course. Subordination-Based Measures (Graph 5) also indicated no progress. Measures like Clauses per Sentence (C/S) and Dependent Clauses per Clause (DC/C) showed inconsistent results, with fluctuations but no overall imporvement. This suggests that learners did not systematically improve in producing subordinated structures. Phrasal Complexity (Graph 6) was the only aspect of syntactic complexity where an upward trend in the metrics, CN/C, CN/T, VP/T, was observed although not linearly and not significantly. It could be argued that phrasal complexity measures improved across most learners, which could indicate a possible shift toward denser, noun-based academic style, even if sentence or clause complexity remained static. These outcomes support the idea that CLIL may not systematically shift students toward more complex global syntactic constructions unless explicit instruction is included (Ibarrola & Lasagabaster, 2021).

Implications

Table 10

Summary of Significant Improvement in Lexical and Syntactic Complexity Measures and Instructional Suggestions

Domain	Measure	Significant Change	Correlation Trend	Effect Size (η^2)	Instructional Suggestion
Lexical Diversity	TTR	Yes	Non-significant ↓	0.943	Introduce targeted vocabulary tasks with range-building goals.
Lexical Diversity	CTTR	Yes	Non-significant ↓	0.905	Use longer writing prompts to naturally increase CTTR.
Lexical Sophistication	LS1	Yes	Non-significant ↑	0.986	Focus on low-frequency and academic vocabulary introduction.
Lexical Sophistication	VS1	Yes	Non-significant ↑	0.887	Support verb-focused lexical expansion using academic verbs.
Lexical Density	LD	Yes	Non-significant ↓	0.829	Encourage summarisation & explanation tasks to increase lexical density.
Sentence Complexity	MLS	Yes	Non-significant ↓	0.72	Encourage varied sentence structure but pair with cohesion training.
Sentence Complexity	MLT	Yes	Non-significant ↓	1	Support with sentence planning to prevent overly short T-units.
Sentence Complexity	MLC	Yes	Non-significant ↑	0.876	Reinforce internal clause building with practice in relative clauses.
Subordination	C/S	Yes	Significant ↑	0.923	Introduce scaffolded exercises on subordination and conjunction use.
Subordination	DC/C	Yes	Non-significant ↑	0.798	Encourage awareness of clause roles in complex ideas.
Subordination	CT/T	Yes	Non-significant ↑	0.735	Use structured sentence combination tasks to strengthen T-unit complexity.
Phrasal Complexity	CN/C	Yes	Non-significant ↑	0.622	Expose students to academic noun phrase models (e.g., nominalisation).
Phrasal Complexity	CN/T	Yes	Non-significant ↓	0.533	Provide sentence frames with embedded nominal structures.
Phrasal Complexity	VP/T	Yes	Non-significant ↑	0.391	Practice expanding verb phrases through tense, aspect, and modality.

As there was no significant progress evident in any of the lexical and synactic complexity over the academic year (Table 10), it would appear that there is some kind of cognitive trade-off in L2 writing (Skehan, 2009). Learners may prioritise either vocabulary or syntax depending on

task demands and cognitive resources. This finding has important pedagogical implications: without structured support, students may be unable to optimise both dimensions of complexity, limiting their academic writing development.Taken together, the findings confirm that CLIL serves as a valuable platform for extending fluency and discourse-level writing, but its impact on advanced academic language development, especially lexical sophistication and broader syntactic growth, requires additional pedagogical scaffolding. The data support calls in recent literature for integrating genre-based writing support and explicit linguistic objectives within CLIL instruction (Gierlinger, 2020; Matsuda, 2023).

As students return from immersive English programmes abroad, their fluency and confidence often outpace their academic writing capabilities. While it is evident that CLIL provides a valuable opportunity to reinforce language learning through meaningful content, additional support would seem to be required to ensure that students can meet the demands of academic discourse. CLIL does promote meaningful language use but it does not automatically result in increased complexity unless it is combined with explicit instruction in academic language. If participants have not been directed to develop specific lexical or syntactic features, it may explain the lack of measurable progress. Moreover, the absence of feedback on writing quality would also have further limited the opportunities for targeted improvement on specific aspects of academic writing. CLIL, by itself, may serve more as a platform to maintain and arguably build skills in L2 communication than a mechanism for developing refined academic writing skills. To develop complexity at a lexical and syntactic level, there may be the need for finely-tuned interventions and carefully scaffolded tasks to develop the mechanics in academic Language. Given the challenging transition from spoken to written academic English, returning students may benefit from scaffolding activities that target sentence structure, academic collocations, and cohesive devices. Combining CLIL with explicit writing instruction could better support learners' needs.

However, it remains unclear at what level that would not interfere with fluency and motivation. One method would be to include focused instruction on aspects such as advanced lexis, and syntactic structures. To optimise gains, instructors may consider: introducing corpus-informed vocabulary tasks that target academic collocations; providing sentence combining and expansion exercises to foster syntactic fluency; or incorporating reflective self-assessment and peer review, helping learners track development over time. Such support may involve tailored writing workshops, peer-review sessions, and exposure to model texts, all of which can guide students in transforming conversational competence into academic proficiency. Hopefully, these measures can support students in transforming the gains made during immersive language use abroad into academic literacy gains that are sustainable and transferrable to university and professional settings.

Conclusion

While the data indicated slight but not significant improvement in syntactic complexitity, except at subordination, overall there was no evidence in clear progress. However, ultimately, CLIL alone is not sufficiently effective in developing higher-level academic writing proficiency. This research would argue for the inclusion of supplementing CLIL-based instruction with focused academic writing support to assist students returning from their EFL programmes abroad. While immersion builds communicative confidence and general proficiency, academic writing development requires targeted strategies to target issues and enhance lexical and syntactic complexity. The integration of CLIL with explicit instruction would, therefore, seem to offer a promising direction for supporting learners’ transition from

general spoken English to more advanced EAP courses. Future implementations of CLIL should consider incorporating academic literacy goals into course design, including formative assessment, metacognitive reflection, and task-based writing development to assist in and stretch further each student's L2 academic endeavours.

References

Ahmadi, H. (2019). Principles of authenticity in second Language instructional materials development: A social-constructivist perspective. *Journal of Applied Linguistics and Language Research*, *6*(3), 132–138.

Ball, P., Kelly, K., & Clegg, J. (2015). *Putting CLIL into Practice.* Oxford University Press.

Biber, D., Gray, B., & Poonpon, K. (2011). Should we use characteristics of conversation to measure grammatical complexity in L2 writing development? *TESOL Quarterly*, *45*(1), 5–35.

Bown, J., & Dewey, D. (2020). The impact of study abroad on L2 fluency and interactional competence. *Study Abroad Research in Second Language Acquisition and International Education*, *5*(1), 36–59.

Bulté, B., & Housen, A. (2021). Conceptualising and measuring L2 writing complexity: New approaches and classroom applications. *Language Teaching Research*, *25*(5), 723–745.

Coyle, D., Hood, P., & Marsh, D. (2010). *CLIL: Content and Language Integrated Learning*. Cambridge University Press.

Crossley, S. A., & Kyle, K. (2020). Measuring syntactic complexity in L2 writing using fine-grained indices. *Applied Linguistics*, *41*(2), 146–168.

Crossley, S. A., & Kyle, K. (2021). Modeling lexical sophistication: Understanding its relationship to writing quality. *Journal of Writing Research*, *13*(2), 331–356.

Crossley, S. A., Roscoe, R., & McNamara, D. S. (2020). Predicting writing proficiency using linguistic features: Tools for academic writing development. *Journal of Writing Research, 2020;12*(1), 171–200.

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* (pp. 139–158). Carl Winter.

Dalton-Puffer, C., & Smit, U. (2021). Introduction to the special issue: Research on CLIL. *International Journal of Bilingual Education and Bilingualism*, *24*(2), 153–161.

de Graaff, R., & Koopman, G. (2020). Teacher education for CLIL: Developing content and language integration. *International Journal of Applied Linguistics*, *30*(1), 98–116.

de Graaff, R., Koopman, G. J., Anikina, Y., & Westhoff, G. (2007). An observation tool for effective L2 pedagogy in content and language integrated learning (CLIL). *International Journal of Bilingual Education and Bilingualism*, *10*(5), 603–624.

Echevarria, J., Vogt, M. E., & Short, D. (2017). *Making content comprehensible for English learners: The SIOP Model*. Pearson.

Fernández-Fontecha, A. (2015). Vocabulary learning in CLIL and EFL contexts: A comparative study. *Complutense Journal of English Studies*, *23*, 79–96.

Fontecha, A. F. (2015). Motivation and Vocabulary Breadth in CLIL and EFL Contexts. Different age, Same Time of Exposure. *Complutense Journal of English Studies*, *23*, 79–96.

Freed, B. F. (1995). What makes us think that students who study abroad become fluent? In B. F. Freed (Ed.), *Second Language Acquisition in a Study Abroad Context* (pp. 123–148). John Benjamins.

Friginal, E., & Polat, B. (2022). *The Routledge Handbook of Second Language Acquisition and Technology*. Routledge.

García-Pinar, A. (2019). Getting closer to authenticity in the course of technical English: Task-based instruction and TED talks. *English Language Teaching*, *12*(11), 10.

Gierlinger, E. (2020). Teaching academic language through CLIL: Challenges and opportunities. *TESOL Quarterly*, *54*(2), 488–515.

Gilmore, A. (2007). Authentic materials and authenticity in foreign language learning. *Language Teaching*, *40*(2), 97–118.

Honnibal, M., & Montani, I. (2017). spaCy 2: Natural language understanding with Bloom embeddings. *ACL Proceedings*.

Ibarrola, A., & Lasagabaster, D. (2021). Syntax in CLIL: Still lagging behind? *Language Awareness*, *30*(4), 329–347.

Imai, H., & Kaneko, T. (2021). Short-term study abroad and the development of academic language in Japanese EFL learners. *Asia TEFL Journal*, *18*(2), 502–520.

Jacobs, G. M., Asmawi, A., & Renandya, W. A. (2022). Teacher authenticity in language education. *International Journal of Education, 15*(2), 69–78.

Johnson, M., & Myles, F. (2020). Cognitive resources and L2 academic writing: Implications for lexical-syntactic integration. *Applied Linguistics Review*, *11*(1), 143–164.

Kinginger, C. (2009). *Language Learning and Study Abroad: A Critical Reading of Research*. Palgrave Macmillan.

Knoch, U., & Elder, C. (2017). Short- and long-term effects of teacher feedback on grammar in second language writing. *CALICO Journal, 34*(3), 1–21.

Kyle, K., & Crossley, S. (2019). Assessing syntactic development in L2 writing: A multi-dimensional approach. *Language Testing*, *36*(2), 249–270.

Larsen, C. C., Walsh, N., Almond, X., & Myers, C. (2017). The “Real Value” of field trips in the early weeks of higher education: The student perspective. *Educational Studies*, *43*(1), 110–121.

Lasagabaster, D. (2021). CLIL in higher education: A ten-year review of research (2010–2020). *Language Teaching*, *54*(1), 1–19.

Lee, J., & Curdt, L. (2021). The role of academic writing instruction in post-study abroad contexts. *TESOL Quarterly*, *55*(3), 870–899.

Llanes, À., & Muñoz, C. (2009). A short stay abroad: Does it make a difference? *System*, *37*(3), 353–365.

López-Medina, E. F., & Casado Casado, J. M. (2024). CLIL, a competence-based coping strategy against foreign language anxiety in higher education. *Discover Psychology*, *4*, 68.

Lopriore, L. (2021). Content and Language Integrated Learning in an ELF-aware perspective. In A. Graziano, B. Turchetta, F. Benedetti & L. Cinganotto (Eds.), *Pedagogical and Technological Innovations in (and through) Content and Language Integrated Learning* (pp. 80–98). Cambridge Scholars.

Lorenzo, F., Casal, S., & Moore, P. (2010). The effects of content and language integrated learning in European education: Key findings from the Andalusian bilingual sections evaluation project. *Applied Linguistics*, *31*(3), 418–442.

Lu, X. (2010). Automatic analysis of syntactic complexity in second language writing. *International Journal of Corpus Linguistics*, *15*(4), 474–496.

Lu, X. (2012). The relationship of lexical richness to the quality of ESL learners' oral narratives. *The Modern Language Journal, 96*(2), 190–208.

Lu, X. (2024). The L2 Syntactic Complexity Analyzer (L2SCA). https://ai-lab.gsu.edu/l2sca. Accessed 12th August, 2024.

Lu, X. (2024). *The Lexical Complexity Analyzer (LCA)*. https://ai-lab.gsu.edu/lca. Accessed 12th August 2024.

Lu, X., & Ai, H. (2015). Syntactic Complexity in College-Level English Writing: Differences among Writers with Diverse L1 Backgrounds. *Journal of Second Language Writing, 29*, 16–27.

Malvern, D., Richards, B., Chipere, N., & Durán, P. (2020). *Lexical Diversity and Language Development: Quantification and Assessment*. Palgrave Macmillan.

Marijuan, S., & Sanz, C. (2020). Beyond conversation: Examining L2 writing outcomes from study abroad. *Foreign Language Annals*, *53*(1), 100–118.

Matsuda, A. (2023). Supporting academic language use in Japanese CLIL classrooms. *Journal of Language, Identity & Education*, *22*(1), 46–61.

Mehisto, P., Marsh, D., & Frigols, M. J. (2019). *Uncovering CLIL: Content and Language Integrated Learning in Bilingual and Multilingual Education*. Macmillan Education.

Morton T. (2020). Integrating content and language in CLIL: A discourse perspective. *Language Teaching Research*, *24*(3), 388–404.

Ozverir, I., Osam, U. V., & Herrington, J. (2017). Investigating the Effects of Authentic Activities on Foreign Language Learning: A Design-based research approach. *Educational Technology & Society*, *20*(4), 261–274.

Paquot, M., & Granger, S. (2020). Academic vocabulary in learner writing: From extraction to pedagogy. *ReCALL*, *32*(3), 365–383.

Parkinson, J., & Musgrave, J. (2014). Development of noun phrase complexity in the writing of English for academic purposes students. *Journal of English for Academic Purposes*, *14*, 48–59.

Pavón-Vázquez, V., & Ellison, M. (2013). Examining teacher roles and competences in Content and Language Integrated Learning (CLIL). *Linguarum Arena*, *4*, 65–78.

Pérez-Vidal, C., & Barquin, E. (2014). Comparing progress in academic writing after formal instruction and study abroad. In C. Pérez-Vidal (Ed.), *Language acquisition in study abroad and formal instruction contexts* (pp. 217–234). John Benjamins.

Perkins, D. N., & Salomon, G. (2012). Knowledge to go: A motivational and dispositional view of transfer. *Educational Psychologist*, *47*(3), 248–258.

Pfenninger, S. E., & Meyer, L. (2021). Assessment in CLIL: Current practices and future directions. *Language Assessment Quarterly*, *18*(4), 311–330.

Purnawarman, P., & Dajarati, U. (2020). Authentic Speaking Assessment Applied by English Teachers During Online Learning. *Advances in Social Science, Education and Humanities Research*, *509*, 560–566.

Révész, A., Michel, M., & Gilabert, R. (2020). Measuring and teaching complexity in L2 writing: What we know and what we need. *Language Teaching Research*, *24*(3), 329–345.

Reynolds, D. (2020). Study abroad and second language learning: A critical overview. *Language Teaching*, *53*(4), 443–456.

Schmitt, N., Jiang, X., & Grabe, W. (2019). The frequency of lexical words in academic texts. *TESOL Quarterly*, *53*(4), 1010–1030.

Skehan, P. (2009). Modelling second language performance: Integrating complexity, accuracy, fluency, and lexis. *Applied Linguistics, 30*(4), 510–532.

Spring, R., & Johnson, E. R. (2022). Comparing three lexical analysis tools: LCA, SpaCy, and NLTK. *System*, *108*, 102770.

Staples, S., & Biber, D. (2021a). Syntactic complexity in argumentative writing: L1 and L2 development over time. *TESOL Quarterly*, *55*(1), 230–256.

Staples, S., & Biber, D. (2021b). Exploring the use of phrasal structures in L2 academic writing. *English for Specific Purposes*, *61*, 29–45.

Staples, S., Reppen, R., & Biber, D. (2020). Using multidimensional analysis to explore lexical density in academic writing. *Applied Linguistics*, *41*(2), 169–196.

Strobl, C., & Baten, K. (2021). Writing development during study abroad: The role of language contact and social networks. *Study Abroad Research in Second Language Acquisition and International Education*, *6*(1), 124–162.

Taguchi, N., & Greer, T. (2019). Pragmatics in study abroad: Gains, limitations, and future directions. *Annual Review of Applied Linguistics*, *39*, 236–253.

Tavakoli, P., & Wright, C. (2020). *Second Language Speech Fluency: From Research to Practice*. Cambridge University Press.

Urmeneta, E. C. (2019). An Introduction to Content and Language Integrated Learning (CLIL) for Teachers and Teacher Educators. *CLIL Journal of Innovation and Research in Plurilingual and Pluricultural Education, 2*(1), 7–19.

Vellanki, S., & Bandu, S. (2021). Engaging students online with technology-mediated task-based Language teaching. *Arab World English Journal*, (1), 107–126.

Vyatkina, N. (2012). Developing syntactic complexity in L2 German writing: A longitudinal study. *Modern Language Journal*, *96*(4), 576–598.

Wadden, P., Hale, C. C, Shiroza, S., Otsuru, S., & Martin, H. M. (2024). Transforming university English with CLIL curricula: An in-depth case study. In B. Lacy, R. P. Lege, & P. Ferguson (Eds.), *Growth Mindset in Language Education*. JALT. https://doi.org/10.37546/JALTPCP2023-26

Wang, J., Dyehouse, M., Weber, N., & Strobel, J. (2012). Conceptualizing Authenticity in Engineering Education: A Systematic Literature Review. Conference at *American Society for Engineering Education.* https://doi.org/10.18260/1-2–21098

Yang, Y., Yap, N. T., & Mohamad Ali, A. (2023). Predicting EFL expository writing quality with measures of lexical richness. *Assessing Writing*, *57*, 100762.

Contact email: jpipe@kanto-gakuin.ac.jp