

Development in Lexical and Syntactic Complexity through EAP after Five-month Study Abroad Programme

Jason Pipe and Teruaki Tsushima¹⁾

ABSTRACT

The purpose of this paper is to examine whether a course on English for Academic Purposes (EAP) at a Japanese university could foster greater critical thinking skills and develop the learners' academic skills. After recently completing a five-month study abroad programme, a group of third-year university students (n=14) was introduced to an EAP course that would hopefully result in improvement in lexical and syntactic complexity. With very little research into developing L2 skills post study-abroad, this paper aimed at improving their English proficiency through EAP classes rather than general English classes. Through quantitative analysis, the data observed alterations in syntactic complexity and lexical complexity. The findings showed that students were able to improve syntactic complexity and, to a lesser degree, lexical complexity. In the case of syntactic complexity, there was a significant improvement in the mean length of clauses and mean length of T-units and, to a marginal degree, with complex T-units per T-unit and T-unit per sentence. Lexical complexity also observed significant improvement in the Uber Index, the D measure, and lexical variation, especially with regards to adjective variation and modifier variation. While other measures of complexity in both syntax and lexis did not establish significant progress, there was no negative performance observed in the research. This paper, therefore, stresses the importance of providing EAP courses that continue to stretch student learning beyond general EFL classes in order to improve their academic abilities and further develop their proficiency in English.

key words; English for Academic Purposes, critical thinking, syntactic complexity, lexical complexity, L2 writing

1. Introduction

University students are often provided courses to develop their skills at an academic level. These courses can prove useful in preparing students before they embark on studying overseas. However, there would appear to be a limitation on the quality of courses available when these students return to their Japanese universities (Iida & Herder, 2019). As a consequence, little research has been conducted as to how students can further improve their academic skills post-international study. While it is agreed that students develop greater fluency than those in the intensive program at home (e.g., Lafford, 2004; Serrano et al., 2011), there is even less research available on how this proficiency can be maintained (Tanaka, 2010) except that which focuses on qualitative studies related to student motivation (e.g., Iida, 2013; Kimura, 2011; Lehner, 2011; Sasaki, 2007). Although one quantitative study observed that a group of students who participated in the overseas programme made greater linguistic gains in English than students who only studied in their own country, this was evident in terms of fluency. Changes in complexity and accuracy remained unnoticeable (Tanaka & Ellis, 2003). Another study noted that university students who studied overseas developed general English skills but no significant improvement in their academic English skills (Iida & Herder, 2019). As a result, this paper is concerned with how to maintain and develop academic skills in terms of lexical and syntactic complexity.

Obviously, studying abroad has benefits as it can build self-confidence and self-efficacy. Through their efforts, students can observe generally incremental successes that lead to further motivation. Furthermore, such an opportunity to participate in a study-abroad programme debunks the ubiquitous Grammar Translation Method in Japanese high school English classes which focuses almost completely on complexity and accuracy (Nishino, 2008; Sakui, 2004). Instead, students experience English as a communicational tool and learn the valuable lesson that expressing one's thoughts is more important than being accurate (Iida & Herder, 2019). Finally, increased fluency prepares students to be able to interact much more successfully in an international environment during their study abroad.

However, a possibility must also be considered that students may lose their momentum in their English studies which could lead to stilted improvement and possibly even a drop in their L2 proficiency. As a consequence, this pilot paper aims at developing a

course to continue student learning that stretches their abilities and maintains a relevant impetus in their English studies. Students were, therefore, introduced to a course that focused on English for Academic Purposes (EAP). By considering how to design and develop an EAP programme to match the needs of a group of students (n=14) who recently studied in Australia for five months, it is believed that this cohort could significantly improve their academic abilities in terms of lexical and syntactic complexity.

2. Literature Review

Courses that aim to develop skills in EAP are different from those that teach general English. While general courses focus on improving vocabulary, grammatical structures, and general language skills for socializing (Campion, 2012; Gao & Bartlett, 2014), EAP aims at stretching students' abilities in reading and writing (Hamp-Lyons, 2011) and requires students to have the necessary language and study skills to embark on academic purposes (Hyland & Shaw, 2016). This would require EAP courses to consider aspects such as research aims, student needs, and pedagogical goals (Ding & Bruce, 2017) as the ultimate goal; and that academic outcome is driven by guiding students to enter an international university (Alexander et al., 2008). In order to benefit from this level of study, students also need to sensitise themselves to 'the salient rhetorical and linguistic features of key disciplinary genres' (Evans & Morrison, 2011: 389).

Previous research on needs analysis of the EAP students would suggest the importance of closing gaps between the level of L2 learners and the high demands of EAP tasks (e.g., Berman & Chang, 2001; Hosogoshi & Takahashi, 2015; Nishikawa et al., 2023). However, with general English teaching approaches mainly adopted in university classes in Japan, teachers may impede the transition necessary for the university to provide courses that orient towards EAP (Alexander, 2012; Farrell & Yang, 2017; Han et al., 2021). EAP requires additional abilities, skills, and approaches that regular teachers may not be aware of (Alexander, 2012; Han et al., 2021; Martin, 2014). As a result, students from this research who returned from post-international study were orientated towards typical EAP teaching approaches taught at other international universities in the UK. This would infer that students would be taught so that they can communicate effectively in academic environments in the UK (Pandey, 2019).

While EAP pedagogy has developed over the past 40 years, there would appear to be no clear definition. In fact, EAP has become a larger field, but also more patchwork and

Table 1: Multiple Literacies of EAP Concepts.

| EAP concepts | Definitions |
|------------------------------|--|
| EAP as academic literacy | Instruction to students of the necessary English language and study skills to complete their academic studies and research. |
| EAP as disciplinary cultural | Teaching students common knowledge and epistemology confirmed in a discursive culture, to help them learn and socialize in academic environments. |
| EAP as critical literacy | Developing a broader conceptual understanding of discipline and the academic world so that students may have a deeper understanding of what they are learning, rather than “blindly” following; and thereby understand power relationships and ideology in the academia. |
| EAP as digital literacy | Teaching students the essential ICT skills for completing academic research and helping them to adjust to the new academic media they may encounter in their future. |

fragmented, at least from the point of view of programme delivery (Hamp-Lyons, 2011). Current research systematically organised theories of EAP into a combination of multiple literacies such as academic literacy, disciplinary cultural literacy, critical literacy and digital literacy (refer to Table 1). As a result, EAP practitioners are challenged to incorporate “a large body of work that has both expanded and deepened the intellectual, theoretical and empirical foundations available to inform and direct praxis” (Ding & Campion, 2016: 1019). However, it is generally agreed that this term refers to the concern with language that is embedded in the practices, discourses, and texts of the academic world in which L2 students aspire to learn (Li, 2017; Bruce, 2017; Ding & Bruce, 2017).

To simplify the theoretical underpinnings, especially in the context of a Japanese university environment, this paper endeavoured to raise EAP teaching standards that suited the needs of the learners. As students were not studying in an international environment and were not focusing on their main field of study, little attention was placed on disciplinary cultural and digital literacy. Instead, this EAP course concentrated on academic and critical literacy. Materials, therefore, incorporated content-based activities that developed learners’ overall language competencies within an academic framework (Sayed, 2018). Central to the course would be the necessity to foster learner critical thinking skills (De Chzal, 2013), an ability to read, analyse, summarise and synthesise complex texts (Tanaka, 2010), and skills to produce academic essays and research papers (Intersegmental Committee of the Academic Senate, 2002). Materials also entailed a focus on teaching

critical thinking and argumentative writing skills (Tanaka, 2009). Through these aspects of EAP, it is believed that students could develop their L2 competencies.

3. Research focus

The main purpose of the present study is to observe any improvement in the level of L2 proficiency in terms of language complexity over a single academic term. This paper is concerned with how to maintain and develop academic writing skills in terms of lexical and syntactic complexity. As a result, attention will be placed on the following two research questions:

1. Is there a difference in the level of syntactic complexity as a result of the introduction of EAP classes?
2. Is there a difference in the level of lexical complexity as a result of the introduction of EAP classes?

4. Methodology

4.1 Participants

The participants were 14 third-year students from a private university in Tokyo. Their English ability was categorized as intermediate to advanced because of improvement in their L2 fluency whilst studying abroad in Australia for five months. This can be verified by their rapid progress in their TOEIC scores which ranged from 400 to 755 before embarking on their studies abroad (CEFR levels of between lower B1 and upper A1). After returning, their TOEIC scores progressed to a higher range of 570 to 835 (a CEFR level of the upper B2 and lower C1). Most noticeably, students appeared to have grown in their L2 spoken fluency and confidence due to their daily experience in conversational English and general English skills courses taught at a university in Sydney. Through impromptu discussions with students, it became evident that students were keen to continue their progress in English by focusing more on EAP skills. They seemed motivated to invest themselves in this EAP project to improve their English abilities at a more academic level. However, it remained unclear whether they were motivated to continue studying abroad at a post-graduate level.

4.2 Lesson procedure

Students were taught once a week for 90 minutes. The classes were split between two types of lessons (refer to Table 2). The first type of lesson concentrated on developing more critical thinking when reading through a consistent format. This was achieved by asking students to first work on paired warm-up tasks on a current event and then read a short article on the more controversial aspect of this event provided by the website: *Breaking English*. This enabled students to acquire higher-level vocabulary, stretch their reading comprehension, and later debate on contentious issues while developing critical thinking (Halpern, 2000; Lipman, 2003). Not only would students naturally be encouraged to develop greater awareness of fundamental academic skills such as hedging, conceding, and countering arguments, but cultivate critical thinking attitudes that are curious and committed to finding answers (Forrester, 2008; Thomas, 2011).

The other type of lesson focused on developing their academic skills. This was based on the coursebook: *Academic Writing Skills 2* (Chin et al., 2012) but had to be supple-

Table 2: Syllabus of EAP course.

| | Content | Aim |
|---------|-----------------------|--|
| Week 1 | Academic Discussion 1 | Whaling should be banned. |
| Week 2 | Academic Discussion 2 | Online shopping is better than shopping in malls. |
| Week 3 | Academic Skills 1 | Essay structure and the introductory paragraph. |
| Week 4 | Academic Skills 2 | Body paragraphs, concluding paragraphs, and outlining. |
| Week 5 | Academic Discussion 3 | What is the real age for society to accept people as adults? |
| Week 6 | Academic Discussion 4 | All countries should drastically curb their reliance on plastics. |
| Week 7 | Academic Skills 3 | Introduce common mistakes in essay writing, and the necessity to research and plan beforehand. |
| Week 8 | Academic Skills 4 | Practising autonomous seminar skills to discuss essay prior to writing. |
| Week 9 | Academic Skills 5 | Improving academic writing and peer review. |
| Week 10 | Academic Skills 6 | Introduction to research and citation. |
| Week 11 | Academic Discussion 5 | No one should have to work more than four days a week. |
| Week 12 | Academic Skills 7 | Developing arguments. |
| Week 13 | Role Play Conference | Defending current investment on SDG's of a country. |
| Week 14 | Academic Discussion 6 | Wealth inequality is immoral and should be made illegal. |

mented with more tailored materials to meet the needs and interests of the class. Students were first introduced to the general structural and grammatical aspects of writing at an academic level with tasks that focused on developing particular writing and research skills. However, through ‘process’ learning (Chunling & Guoping, 2009) which included activities that demand higher-level thinking such as analysis, synthesis, and evaluation (Abdullah et al., 2022); and scaffolding as a means to assist learners in achieving independence (Kurt, 2020), students had the opportunity to also develop more autonomous learning (Abdullah et al., 2022). This would encourage students to become better aware of their own abilities in their academic writing skills while improving their attitudes, self-efficacy, and ‘can do’ skills (Wrigley, 1998). The focus on developing these academic skills was to better prepare them to effectively participate in a seminar in week 8, produce a research-based, argumentative essay by week 9, and independently research for a role-play conference in week 13.

4.3 Complexity measures

4.3.1 Syntactic Complexity Measures

When focusing on providing valid and reliable measures of complexity, research seems to focus prominently on syntactic complexity as an impartial gauge of the overall proficiency of non-native learners (e.g., Larsen-Freeman, 2009; Lu, 2011; Norris & Ortega, 2009; Ortega, 2003; Wolfe-Quintero et al., 1998). Syntactic complexity is, therefore, an important construct in second language evaluation in order to describe the level of proficiency in a second language of a learner. For example, syntactic complexity is reliable in describing grammatical competence in L2 assessment (Park, 2017) or describing changes in learners’ proficiency over time.

It is agreed that the quality of language output is affected by a variety of reasons. This would include aspects such as cognitive load and input processing (Polat, 2016); task design (Frear & Bitchener, 2015; Kormos & Trebits, 2012; Spring & Johnson, 2022), types of writing tasks or registers (Kim et al., 2018), teaching approach (Kim et al., 2018), age (Llanes & Munoz, 2013), first language interference (Lu & Ai, 2015; Crossley & McNamara, 2014); and other stimuli such as studying abroad (Serrano et al., 2012) or pedagogic intervention (Benevento & Storch, 2011; Norris & Ortega, 2009; Storch, 2009). However, despite these concerns, syntactic complexity measures would appear to be the standard of large-scale studies (e.g., Ai & Lu, 2013; Asención-Delaney et al., 2011; Bulté & Housen, 2014; Crossley & McNamara, 2014; Spring & Johnson, 2022; Vyatkina et al., 2015;

Development in Lexical and Syntactic Complexity through EAP after Five-month Study Abroad ...
Yoon & Polio, 2016).

To expediently determine the syntactic complexity of assignments produced, this paper incorporated the reliability of the L2 Syntactic Complexity Analyser (L2SCA), a computational system for automatic measurement of syntactic complexity. Following the recommendations of Lu (2010), the syntactic complexity measures selected for this paper consisted of the six measures covered in both Wolfe-Quintero et al. (1998) and Ortega (2003), another five measures that were shown by at least one previous study to have at least a weak correlation with or effect for proficiency, and three other measures that have not been explored in previous studies but were recommended by Wolfe-Quintero et al. (1998) to pursue further (Lu, 2010: 478). Unfortunately, recent research on the automatic measurement of syntactic complexity would appear limited to analysis on cross-sectional studies (Lorenzo & Rodríguez, 2014; Mancilla et al., 2015) and to comparing groups of writers (Lu & Ai, 2015; Mancilla et al., 2015).

To use L2SCA, the text was simply converted to a plain text format and uploaded to L2SCA, after which the fourteen measures were analysed (refer to Table 3). To achieve this, the text had to first follow the preprocessing stage, in which the system L2SCA used a syntactic parser to analyse the syntactic structures of the sample text. The output was a parsed sample that consisted of a sequence of parse trees, with each parse tree representing the analysis of the syntactic structure of a sentence in the sample. Next, the text was put through the syntactic complexity analysis stage, which entailed retrieving and counting the occurrences of all relevant production units and syntactic structures necessary for calculating one or more of the syntactic complexity indices measures.

Furthermore, due to the failure of previous research to clearly define and consistently apply explicit definitions of each measure of unit (Wolfe-Quintero et al., 1998), this paper was researched in accordance with the exemplar of Lu (2010). By maintaining the explicit and consistent definitions provided in Lu's work (2010, 2011) of the six production units and syntactic structures (refer to Table 4) when calculating syntactic complexity measures, one aimed to enable fairer and more accurate comparisons with previous research.

Finally, due to the reliable statistical analysis of research conducted by Lu (2010), which has been followed up in further research (for example, Ai & Lu; 2013; Lu & Ai, 2015), this paper compared the full range of syntactic complexity measures. In other words, the measurements found in the analysis of the research group (RG) were compared with the original findings of Lu's research group (LG) (2010), to determine the accuracy and proficiency of results. In this seminal piece of work, Lu randomly selected 40

Table 3: Summary of the fourteen measures of syntactic complexity (Mancilla, 2017).

| Measure | Code | Definition |
|---|------|--|
| Type 1: Length of production unit | | These three measures gauge length of production at the clausal, sentential and T-unit level |
| Mean length of clause | MLC | Mean length of clause: Number of words / # of clauses |
| Mean length of sentence | MLS | Mean length of sentence: Number of words / # of sentences |
| Mean length of T-unit | MLT | Mean length of T-unit: Number of words / # of T-units |
| Type 2: Overall Sentence complexity | | This measure concentrates on the sentence complexity ratio |
| Sentence complexity ratio | C/S | Number of clauses per sentence (Not to be used) |
| Type 3: Degree of Subordination | | Contains four ratios that reflect the amount of subordination |
| T-unit complexity ratio | C/T | Number of clauses / # of T-units (Not to be used) |
| Complex T-unit ratio | CT/T | Number of complex T-units / # of T-units (Not to be used) |
| Dependent clause ratio | DC/C | Number of dependent clauses / # of clauses |
| Dependent clauses per T-unit | DC/T | Number of dependent clauses / # of T-units |
| Type 4: Amount of Coordination | | Includes three ratios that measure the amount of coordination |
| Coordinate phrases per clause | CP/C | Number of coordinate phrases / # of clauses |
| Coordinate phrases per T-unit | CP/T | Number of coordinate phrases / # of T-units |
| Sentence coordination ratio | T/S | Number of T-units / # of sentences |
| Type 5: Degree of phrasal sophistication | | Comprises of three ratios that consider particular structures in relation to larger production units |
| Complex nominals per clause | CN/C | Number of complex nominals / # of clauses |
| Complex nominals per T-unit | CN/T | Number of complex nominals / # of T-units |
| Verb phrases per T-unit | VP/T | Number of verb phrases / # of T-units (Not to be used) |

essays from a corpus of 3,554 essays written by English majors from nine colleges in China. While it was expected for there to be some disparity between LG and RG data sets due to differences in first language, culture, and education, it is believed that the reliability of Lu's data would provide some enlightenment as regards the level of academic ability by RG in terms of identifying relevant production units and syntactic structures from the essays produced.

Table 4: Summary of six production units and syntactic structures when calculating syntactic complexity measures (Lu, 2010, 2011).

| Production Unit | Syntactic Structures |
|-------------------|--|
| Sentence | A sentence is least problematic and is defined by punctuation. If the punctuation results in a fragment or run-on, it is still to be counted as a sentence (Hunt, 1965). If there is a period, but no capitalization, count it as a sentence. Thus a sentence is a group of words (including sentence fragments) punctuated with a sentence-final punctuation mark, including a period, exclamation mark, question mark, and occasionally elliptical marks or closing quotation marks. |
| Clause | A structure with a subject and a finite verb, including independent, adjective, adverbial, and nominal clauses, but not non-finite verb phrases, which are included in the definition of verb phrases instead (Hunt 1965; Polio 1997). |
| Dependent clause | A finite adverbial, adjective, or nominal clause - usually analysis of dependent clauses (Hunt 1965; Cooper 1976; Kameen 1979). |
| T-unit | A T-unit consists of a main clause plus any dependent clause or non-clausal structure attached or embedded in it (Hunt 1970: 4). A complex T-unit is defined as an independent clause and its dependent clauses, even if they span more than one "sentence", as defined by punctuation. |
| Coordinate phrase | Refers to coordinate adjective, adverb, noun, and verb phrases. (e.g., They should be responsible and do not make any trouble. // They can have their own bedroom and bathroom.). Also count phrases that may have the coordinator missing (e.g., For dinner, I ate chicken, salad, soup.) |
| Complex nominal | (1) noun phrases with one or more of the following pre- or post-modifiers: adjective, possessive, prepositional phrase, adjective clause, participle, or appositive; Noun phrases with certain premodifiers (i.e., a noun modified by a participle, an adjective with a determiner, an adjective) (e.g., the last spring // my small problems // the process for getting a new visa // a lot of money) (2) nominal clauses: a Double count embedded complex nominal. (e.g. There is a big man that I saw yesterday. (count as 2 CNs)) (3) gerunds and infinitives in subject position (Cooper 1976). Noun clauses (in any position) and infinitives and gerunds in the subject position (e.g., Studying English is my goal for a long time. // I found that many information help to open my mind. // What students should do is that -) |

4.3.2 Lexical Complexity Measures

Researchers have provided various definitions and interpretations of lexical complexity, but it is commonly defined as "a wide variety of basic and sophisticated words that are available and can be accessed quickly, whereas a lack of complexity means that only a narrow range of basic words are available or can be accessed" (Wolfe-Quintero et al.,

1998: 102). Despite the important role of lexical competence in language learning and teaching (Alqahtani, 2015; Caro & Mendinueta, 2017; Lu, 2012), there still seems to be relatively little attention placed on the successful acquisition of lexical items in L2 English classes at schools in Japan. To be lexically more competent, the learner has to be encouraged to apply the lexical resource in an effort to maintain their utterances and deepen their understanding of vocabulary from the perspectives of form, meaning, and use (Nation, 2013). Most research in this field differentiates lexical diversity (i.e., the size of the lexicon measured by means of, for example, type-token ratio measures), lexical sophistication (i.e., the depth of lexis measured by means of, for example, frequency of rare or academic words), and lexical density (i.e., the amount of information in a text, typically measured by the ratio of lexical words per function words) (Michel, 2017). This paper also concentrates on all these aspects to gain a clearer picture as regards the level of lexical diversity.

In order to analyse the lexical complexity or lexical richness of a text, therefore, it is manifested at the observational level in L2 performance to encompass a broad range of aspects of vocabulary usage, specifically the percentage of meaningful words versus filler or grammatical words (lexical density), the ability to use non-standard words (lexical sophistication) and the range of vocabulary used by a speaker or writer (lexical variation or diversity) (Bulté & Housen, 2012; Lu, 2012). While there is an increasing number of software to assist in evaluating lexical complexity, it remains difficult for analysts to clearly and objectively identify what software performs the best (Schmitt et al., 2019). Analysis of lexical complexity was achieved by incorporating Lu's (2012): Lexical Complexity Analyzer (LCA) to automatically analyse pre-process text files run to determine 26 different lexical measures of richness (Ai & Lu, 2010) as this would appear to be the general standard (refer to Table 5). This analyser compares words from the data analysis with those from either the American National Corpus (ANC) or the British National Corpus (BNC) to determine the general level of lexical richness. While there have been recent developments in language processing systems available in Python with Natural Language ToolKit (NLTK) and SpaCy tools to ensure relative accuracy of calculations of variables to determine lexical EFL proficiency, it would appear that despite improved technical compatibility of running Python and expediency of not requiring pre-processing to parse, tag, and lemmatize input data, (Honnibal & Montani, 2017; Schmitt et al., 2019; Spring & Johnson, 2022), there would seem to be very little variation in these three automated systems (Spring & Johnson, 2022). As a result, to maintain clarity and objectivity (Schmitt

Table 5: Summary of 26 automated measures of lexical complexity (Lu, 2012).

| Measure | Code | Formula | Examples |
|--|----------|---|------------------------------------|
| Type 1: Lexical Density Lexical Density | LD | N_{lex}/N | Engber (1995) |
| Type 2: Lexical Sophistication: Lexical Sophistication-I | LS1 | $N_{\text{slex}}/N_{\text{lex}}$ | Linnarud (1986), Hyltenstam (1988) |
| Lexical Sophistication-II | LS2 | T_{s}/T | Laufer (1994) |
| Verb Sophistication-I | VS1 | $T_{\text{sverb}}/N_{\text{verb}}$ | Harley & King (1989) |
| Corrected VS1 | CVS1 | $T_{\text{sverb}}/\sqrt{2N_{\text{verb}}}$ | Wolfe-Quintero et al. (1998) |
| Verb Sophistication-II | VS2 | $T_{\text{sverb}}^2/N_{\text{verb}}$ | Chaudron & Parker (1990) |
| Measure | Code | Formula | |
| Type 3: Lexical Variation | | | |
| Number of Different Words | NDW | T | |
| NDW (first 50 words) | NDW-50 | T in the first 50 words of sample | |
| NDW (expected random 50) | NDW-ER50 | Mean T of 10 random 50-word samples | |
| NDW (expected sequence 50) | NDW-ES50 | Mean T of 10 random 50-word sequences | |
| Type-Token Ratio | TTR | T/N | |
| Mean Segmental TTR (50) | MSTTR-50 | Mean TTR of all 50-word segments | |
| Corrected TTR | CTTR | $T/\sqrt{2N}$ | |
| Root TTR | RTTR | T/\sqrt{N} | |
| Bilobarithmic TTR | LogTTR | $\text{Log}T/\text{Log}N$ | |
| Uber Index | Uber | $\text{Log}_2N/\text{Log}(N/T)$ | |
| D Measure | D | $TTR = \frac{D}{N} \left[\left(1 + 2 \frac{N}{D} \right)^{\frac{1}{2}} - 1 \right]$ | |
| Lexical Word Variation | LV | $T_{\text{lex}}/N_{\text{lex}}$ | |
| Verb Variation-I | VV1 | $T_{\text{verb}}/N_{\text{verb}}$ | |
| Squared VV1 | SVV1 | $T_{\text{verb}}^2/N_{\text{verb}}$ | |
| Corrected VV1 | CVV1 | $T_{\text{verb}}/\sqrt{2N_{\text{verb}}}$ | |
| Verb Variation-II | VV2 | $T_{\text{verb}}/N_{\text{lex}}$ | |
| Noun Variation | NV | $T_{\text{noun}}/N_{\text{lex}}$ | |
| Adjective Variation | AdjV | $T_{\text{adj}}/N_{\text{lex}}$ | |
| Adverb Variation | AdvV | $T_{\text{adv}}/N_{\text{lex}}$ | |
| Modifier Variation | ModV | $(T_{\text{adj}} + T_{\text{adv}})/N_{\text{lex}}$ | |

et al., 2019), the mini-essay texts were simply converted to a plain text format and uploaded to a modified version of LCA (see Spring & Johnson, 2022, for details).

Finally, due to the full availability of findings of all 26 different lexical measures, the datasets from RG were compared with the original findings of Lu's group (LG) (2012) to determine the accuracy and proficiency of results. As LG data was taken from transcripts

of 408 test takers' oral productions from second-year majors at Chinese colleges, it is expected that RG would outperform LG in lexical complexity. RG had ten minutes to produce their written essays which put them at an advantage to carefully craft their sentences. However, it is also assumed that the test takers from LG would have possibly memorized more sophisticated lexical phrases to maintain their utterances to convey the persona of being academically capable. Furthermore, to the best of this writer's knowledge, there has been no other research that provides complete datasets to compare the results of RG. Consequently, while it is expected for there to be better results observed from RG, due to the reliability of Lu's data which provides a large and complete set of lexical measures from a large dataset, it is still conceivable that analysis of the results will possibly illuminate the reader as to the level of academic ability by RG in terms of lexical complexity from the essays produced.

4.4 Data Elicitation

In total, data from 6 points were obtained during a single semester. The writing was supervised by the teacher in class so that students could only rely on their own L2 skills. Due to a limitation in time, these pieces of written work were conducted at the end of the academic discussion on weeks 1, 2, 5, 6, 11, 14. Furthermore, due to a limit in resources, students were provided 10 minutes to complete the tasks at the end of these classes. While these mini-essays could have been written at home so that students could spend as much time as they felt suitable, it was necessary to conduct the data elicitation under specific conditions so that the results could be quantifiably analysed and compared. However, it became apparent that these pieces of writing were quite short due to the time restrictions. To increase the number of word tokens in one file and reduce the influence of a particular topic, two passages were combined. In other words, from the 6 mini-essays, 1&2, 3&4, 5&6 were analysed together, resulting in a total of three files: P1/2, P3/4, and P5/6. These three files were then analysed through the automatic syntactic analyser, L2SCA, and the automatic lexical analyser, LCA.

5. Results

5.1 Syntactic Complexity

The first type of syntactic complexity measure consists of three measures that gauge length of production at the clausal, sentential, or T-unit level, namely, mean length of

Table 6: Comparison of the research group with Lu's group within fourteen syntactic complexity measures (Lu, 2010).

| | Research Group | | | | | | Proficiency Level (Lu, 2010) | | | | | |
|-----------------------------------|----------------|------|-------|------|-------|------|------------------------------|------|---------|------|---------|------|
| | P1/2 | | P3/4 | | P5/6 | | Level 1 | | Level 2 | | Level 3 | |
| | Mean | SD | Mean | SD | Mean | SD | Mean | SD | Mean | SD | Mean | SD |
| Mean Length of Clause (MLC) | 12.30 | 2.91 | 13.52 | 2.62 | 14.33 | 3.05 | 8.80 | 1.44 | 9.10 | 1.50 | 9.63 | 1.74 |
| Mean Length of Sentence (MLS) | 7.15 | 1.26 | 8.13 | 1.36 | 8.07 | 1.18 | 14.68 | 3.67 | 14.67 | 3.34 | 15.75 | 3.51 |
| Mean Length of T-Unit (MLT) | 14.09 | 3.17 | 15.63 | 3.38 | 16.35 | 3.55 | 13.18 | 2.56 | 13.32 | 2.63 | 14.43 | 2.95 |
| Clauses/Sentence (C/S) | 1.98 | 0.38 | 1.94 | 0.39 | 2.03 | 0.36 | 1.69 | 0.43 | 1.63 | 0.34 | 1.66 | 0.35 |
| T-unit Complexity Ratio (C/T) | 0.52 | 0.12 | 0.53 | 0.11 | 0.56 | 0.15 | 1.51 | 0.27 | 1.48 | 0.24 | 1.51 | 0.27 |
| Complex T-unit Ratio (CT/T) | 0.18 | 0.10 | 0.25 | 0.21 | 0.32 | 0.20 | 0.40 | 0.14 | 0.38 | 0.14 | 0.39 | 0.13 |
| Dependent Clause Ratio (DC/C) | 0.39 | 0.07 | 0.40 | 0.06 | 0.40 | 0.10 | 0.33 | 0.09 | 0.32 | 0.10 | 0.33 | 0.09 |
| Dependent Clauses/T-Unit (DC/T) | 0.69 | 0.26 | 0.69 | 0.18 | 0.74 | 0.27 | 0.52 | 0.22 | 0.49 | 0.22 | 0.51 | 0.22 |
| Coordinate Phrases/Clause (CP/C) | 0.80 | 0.25 | 0.92 | 0.21 | 0.88 | 0.22 | 0.22 | 0.13 | 0.23 | 0.13 | 0.28 | 0.14 |
| Coordinate Phrases/T-unit (CP/T) | 1.15 | 0.13 | 1.17 | 0.20 | 1.14 | 0.11 | 0.33 | 0.18 | 0.34 | 0.18 | 0.42 | 0.22 |
| Sentence Coordination Ratio (T/S) | 0.10 | 0.06 | 0.15 | 0.11 | 0.18 | 0.10 | 1.11 | 0.15 | 1.10 | 0.14 | 1.09 | 0.10 |
| Complex Nominals/Clause (CN/C) | 2.28 | 0.44 | 2.23 | 0.25 | 2.25 | 0.42 | 0.94 | 0.28 | 0.99 | 0.28 | 1.10 | 0.33 |
| Complex Nominals/T-unit (CN/T) | 1.38 | 0.52 | 1.54 | 0.44 | 1.57 | 0.46 | 1.42 | 0.45 | 1.45 | 0.48 | 1.65 | 0.52 |
| Verb Phrases/T-unit (VP/T) | 1.73 | 0.30 | 1.67 | 0.20 | 1.78 | 0.29 | 2.04 | 0.40 | 1.97 | 0.38 | 2.06 | 0.38 |

clause (MLC), mean length of sentence (MLS), and mean length of T-unit (MLT). While it is noted that MLS performed at a lower level for RG (ranging from 7.15 to 8.13) compared to LG (ranging from 14.68 (level 1) to 15.75 (level 3)), one can observe marked improvement in the quality of MLC and MLT. MLC showed significant improvement from 12.30 (P1/2) to 14.33 (P5/6), $F(2, 24)=4.42, p=0.023, \eta^2=0.269^2$, and that the findings were higher than LG (ranging from 8.80 (level 1) to 9.63 (level 3)). MLT also showed similar performance by also resulting in significant progress from 14.09 (P1/2) to 16.35 (P5/6), $F(2, 24)=3.49, p=0.047, \eta^2=0.225$, while LG's analysis maintained at a lower level (ranging from 13.18 (level 1) to 14.43 (level 3)). One can infer that while students from the RG found it challenging to consistently maintain lengthy sentences, they were able to provide longer clauses and T-units.

The second type consists of a sentence complexity ratio (clauses per sentence, or C/S). While RG seemed to show no significant progress, the results indicate a non-significant change from 1.98 (P1/2) to 2.03 (P5/6), $p=.792, \eta^2=0.019$, but C/S were notably higher than LG (ranging from 1.69 (level 1) to 1.66 (level 3)). It can be deduced that LG made little progress in producing sentences with more complexity in terms of the number of clauses in a sentence.

The third type comprises four ratios that reflect the amount of subordination, including a T-unit complexity ratio (clauses per T-unit, or C/T), a complex T-unit ratio (complex T-units per T-unit, or CT/T), a dependent clause ratio (dependent clauses per

clause, or DC/C), and dependent clauses per T-unit (DC/T). Looking at C/T, while RG was lower than LG (hovering around 1.50), there was a modest, but non-significant improvement from 0.52 (P1/2) to 0.56 (P5/6). A similar pattern was observed in CT/T. Again, results from RG were lower than LG (averaging around 0.39) but had managed to show a marginally significant improvement from 0.18 (P1/2) to 0.32 (P5/6), $F(2, 24) = 2.89$, $p = 0.075$, $\eta^2 = 0.194$. On the other hand, DC/C and DC/T provide no significant progress for RG (DC/C: 0.39 (P1/2) to 0.40 (P5/6); DC/T: 0.69 (P1/2) to 0.74 (P5/6), but still managed to outperform LG (DC/C: 0.33; DC/T: 0.51). One can conclude that there was a general but not significant improvement in subordination.

The fourth type is made up of three ratios that measure the amount of coordination, namely, coordinate phrases per clause (CP/C), coordinate phrases per T-unit (CP/T), and a sentence coordination ratio (T-units per sentence, or T/S). When analysing CP/C and CP/T, one can note a similar pattern that there was no significant change but still performed better than LG. For RG, there was no obvious progress (CP/C: 0.80 (P1/2) to 0.88 (P5/6); CP/T: 1.15 (P1/2) to 1.14 (P5/6) while LG performed less well (CP/C: 0.22 (level 1) to 0.28 (level 3); CP/T: 0.33 (level 1) to 0.42 (level 3). Turning attention towards T/S, however, one can observe an opposite pattern. RG progressed from 0.10 (P1/2) to 0.18 (P5/6), showing a marginally significant improvement, $F(2, 24) = 2.93$, $p = 0.073$, $\eta^2 = 0.196$, but less impressively compared to LG which averaged 1.10. As a result, RG showed a greater ability in CP/C and CP/T than LG and a steady but modest improvement in T/S.

The final type consists of three ratios that consider the relationship between particular syntactic structures and larger production units, i.e., complex nominals per clause (CN/C), complex nominals per T-unit (CN/T), and verb phrases per T-unit (VP/T). Regarding CN/C, there would appear to be no significant change (from 2.28 (P1/2) to 2.25 (P5/6) but notably higher than LG (from 0.94 (level 1) to 1.10 (level 3)). Nor did CN/T show significant progress (from 1.38 (P1/2) to 1.57 (P5/6) and matched the abilities of the more proficient students from LG (1.45 (level 2) to 1.65 (level 3)). However, the most disappointing result from the fourteen syntactic complexity measures by RG was VP/T. There was little significant improvement in the data (from 1.73 (P1/2) to 1.78 (P5/6), and RG performed less well when compared to LG (from 2.0 (level 1) to 2.06 (level 3)).

5.2 Lexical Complexity

The first type of lexical complexity measure is lexical density. This measure refers to

Table 7: Comparison of the research group with Lu's group within 26 lexical complexity measures (Lu, 2012)

| | Research Group | | | | | | Proficiency Level (Lu, 2012) | |
|------------------------------------|----------------|-------|--------|-------|--------|-------|---------------------------------|-------|
| | 1 & 2 | | 3 & 4 | | 5 & 6 | | Level 1 | |
| | Mean | SD | Mean | SD | Mean | SD | Mean | SD |
| Lexical Density (LD) | 0.54 | 0.02 | 0.56 | 0.02 | 0.56 | 0.03 | 0.41 | 0.33 |
| Lexical Sophistication-I (LS1) | 0.25 | 0.04 | 0.16 | 0.05 | 0.19 | 0.05 | 0.23 | 0.07 |
| Lexical Sophistication-II (LS2) | 0.20 | 0.03 | 0.20 | 0.04 | 0.19 | 0.04 | 0.26 | 0.04 |
| Verb Sophistication-I (VS1) | 0.10 | 0.04 | 0.12 | 0.04 | 0.09 | 0.04 | 0.07 | 0.05 |
| Corrected VS1 (CVS1) | 0.47 | 0.18 | 0.58 | 0.18 | 0.35 | 0.18 | 0.33 | 0.22 |
| Verb Sophistication-II (VS2) | 0.51 | 0.35 | 0.74 | 0.40 | 0.31 | 0.23 | 0.31 | 0.39 |
| Number of Different Words (NDW) | 116.38 | 13.09 | 133.38 | 15.60 | 106.00 | 22.53 | 119.83 | 22.53 |
| NDW-50 (first 50 words) | 38.31 | 2.59 | 37.54 | 2.82 | 38.08 | 2.56 | 34.00 | 3.64 |
| NDW-ER50 (expected random 50) | 38.64 | 1.38 | 40.39 | 1.33 | 39.52 | 1.58 | 36.83 | 2.03 |
| NDW-ES50 (expected sequence 50) | 37.88 | 1.51 | 39.30 | 1.52 | 39.17 | 2.06 | 34.23 | 2.85 |
| Type-Token Ratio (TTR) | 0.48 | 0.05 | 0.51 | 0.05 | 0.56 | 0.08 | 0.41 | 0.06 |
| Mean Segmental TTR (50) (MSTTR-50) | 0.76 | 0.03 | 0.78 | 0.03 | 0.78 | 0.05 | 0.69 | 0.05 |
| Corrected TTR (CTTR) | 5.25 | 0.29 | 5.81 | 0.31 | 5.39 | 0.50 | 4.94 | 0.57 |
| Root TTR (RTTR) | 7.43 | 0.41 | 8.22 | 0.44 | 7.62 | 0.70 | 6.99 | 0.80 |
| Bilogarithmic TTR (LogTTR) | 0.87 | 0.01 | 0.88 | 0.02 | 0.89 | 0.02 | 0.84 | 0.02 |
| Uber Index (Uber) | 17.83 | 1.58 | 20.16 | 2.12 | 21.06 | 3.77 | 26.15 | 4.63 |
| D Measure (D) | 53.62 | 9.33 | 70.31 | 13.17 | 69.28 | 20.01 | 55.56 | 13.39 |
| Lexical Word Variation (LV) | 0.66 | 0.05 | 0.70 | 0.07 | 0.72 | 0.09 | 0.57 | 0.08 |
| Verb Variation-I (VV1) | 0.63 | 0.05 | 0.70 | 0.09 | 0.70 | 0.13 | 0.58 | 0.12 |
| Squared VV1 (SVV1) | 18.05 | 3.08 | 22.13 | 4.54 | 14.78 | 3.83 | 13.42 | 4.32 |
| Corrected VV1 (CVV1) | 2.99 | 0.26 | 3.31 | 0.34 | 2.70 | 0.36 | 2.56 | 0.42 |
| Verb Variation-II (VV2) | 0.22 | 0.03 | 0.22 | 0.03 | 0.20 | 0.05 | 0.19 | 0.04 |
| Noun Variation (NV) | 0.67 | 0.07 | 0.67 | 0.08 | 0.70 | 0.11 | 0.69 | 0.10 |
| Adjective Variation (AdjV) | 0.11 | 0.02 | 0.11 | 0.03 | 0.15 | 0.03 | 0.11 | 0.04 |
| Adverb Variation (AdvV) | 0.09 | 0.02 | 0.10 | 0.03 | 0.09 | 0.03 | 0.04 | 0.02 |
| Modifier Variation (ModV) | 0.20 | 0.03 | 0.21 | 0.04 | 0.24 | 0.04 | 0.15 | 0.04 |

the number of lexical words (nouns, adjectives, verbs (excluding modal and auxiliary verbs), and adverbs) (i.e., not grammatical words) to the number of words in the text (Ure, 1971). Unsurprisingly, there is a higher ratio than LG (0.41), but one can also note only slight and non-significant improvement (from 0.54 (P1/2) to 0.56 (P5/6)).

The second type of lexical complexity measure is lexical sophistication or lexical rareness as it aims to determine the proportion of relatively unusual words to advanced words in a learner's L2 repertoire (Read, 2000). Due to the considerable variability in how sophisticated words are defined across previous studies, this paper concentrates on five measures: Lexical Sophistication-I (LS1), Lexical Sophistication-II (LS2), Verb Sophistication-I (VS1), Corrected VS1 (CVS1), and Verb Sophistication-II (VS2). Surprisingly, LG outperformed RG in terms of LS1 and LS2, and RG seemed to deteriorate over the period. LS1 for LG was maintained at a mean of 0.23, while RG fell from 0.25 (P1/2) to 0.19

(P5/6). This would indicate that LG used a greater amount of sophisticated lexical words (beyond the 2,000 most frequent English words) than the number of lexical words while RG opted to use more common words. LS2 for LG averaged at 0.26 while RG dropped slightly from 0.20 (P1/2) to 0.19 (P5/6). Again, this would suggest that LG used a greater proportion of the second 1,000 most frequent words compared to the first 1,000 most frequent words in their oral narratives compared to RG. However, when focusing more specifically on verb sophistication, RG outperformed LG in VS1, CVS1, and VS2. While LG performed at 0.07, 0.33, and 0.31 respectively, RG maintained but not improved at a level of VS1 (0.09–0.12), CVS1 (0.35–0.58), and VS2 (0.31–0.74). One can, therefore, suggest that RG preferred to write with more frequently used lexis compared to LG but had the ability to apply less commonly known verbs.

The final type of lexical complexity measure concentrates on lexical variation. This aspect of lexical richness aims to discover the wide range of learners' vocabulary (Crystal, 1982) and hence this paper concentrates on 20 measures of lexical variation. From a glance, RG outperformed LG except in the number of different words (NDW), noun variation (NV), verb variation (VV), and Uber Index (Uber). In terms of progress, there would seem to be a mixed range of improvement.

Focusing on NDW, RG and LG have a similar range. However, RG outperformed LG within NDW-50 (RG average: 37.98 to LG: 34.88), NDW-ER50 (RG average: 39.52 to LG: 36.83), and NDW-ES50 (RG average: 38.78 to LG: 34.23). Statistical analyses showed that, in NDW-50, there was a significant improvement between 1&2 and 3&4, $p=0.004^3$, while the difference was not significant between 1&2 and 5&6 (from 38.64 (P1/2) to 39.52 (P5/6)), $p=0.275$. As for NDW-ES50, there was a significant improvement between 1&2 and 2&3, $p=0.046$, the difference between 1&2 and 5&6 (from 37.88 (P1/2) to 39.17 (P5/6)) was not significant, $p=0.346$. This would seem to indicate that despite the modest improvement in NDW throughout the semester, there was a limit on how many different words students could apply to their writing by the end of the semester.

Turning attention towards TTR, i.e., the ratio of the number of word types (T) to the number of word tokens (N) in the text, it would seem that RG again outperformed LG in all aspects of the earlier TTR transformation models. TTR (RG average: 0.52 to LG: 0.41), TTR (50) (RG average: 0.77 to LG: 0.69), CTTR (RG average: 5.48 to LG: 4.94), RTTR (RG average: 7.76 to LG: 6.99), LogTTR (RG average: 0.88 to LG: 0.84). TTR has been criticized due to unsatisfactory validity and reliability of data due to a sample size problem and unequal dividing of standard-sized segments (Malvern et al. 2004). Instead,

more “robust measures of lexical diversity which are not a function of sample size” would also be considered (Malvern et al. 2004: 60). Among these measures, MSTTR-50, CTTR, RTTR, and LogTTR did not show a clear pattern of improvement. However, significant improvements were found in both the Uber Index, $F(2, 24)=6.53$, $p=0.005$, $\eta^2=0.352$, and the D measure, $F(2, 24)=6.72$, $p=0.006$, $\eta^2=0.343$, suggesting that the type-token ratio increased through the term. RG outperformed LG (RG average: 64.40 to LG: 55.56) in the D measure, while the Uber Index would seem to contradict earlier trends as LG performed better than RG (LG: 26.16 to RG average: 19.68).

Concentrating on the other eight measures of lexical variation, there would appear to be mixed results. First of all, LV showed a significant improvement, $F(2, 24)=3.87$, $p=0.35$, $\eta^2=0.244$, suggesting that the variation of lexical words became larger through the course. The RG’s average (0.69) was higher than that of LG (0.57). Looking at verb variation, i.e. the ratio of the number of verb types to the total number verbs in the text, one can observe that again, RG outperformed LG in each category: VV1 (RG average: 0.68 to LG: 0.58); SVV1 (RG average: 18.32 to LG: 13.42); CVV1 (RG average: 3.00 to LG: 2.56); and VV2 (RG average: 0.21 to LG: 0.19). However, when analysing performance over the semester, there would appear to be no significant improvement: VV1 (from 0.63 (P1/2) to 0.70 (P5/6)); SVV1 (from 18.05 (P1/2) to 14.78 (P5/6) with a spike of 22.13 (P3/4)); CVV1 (from 2.99 (P1/2) to 2.70 (P5/6) with a spike of 3.31 (P3/4)); and VV2 (from 0.20 (P1/2) to 0.24 (P5/6)). Turning to noun, adjective, and adverb variation, RG seemed comparable to LG in NV (RG average: 0.68 to LG: 0.69); and AdjV (RG average: 0.12 to LG: 0.11) but stronger in AdvV (RG average: 0.09 to LG: 0.04); and modifier variation, a measure which concentrates on variation of both adverbs and adjectives (ModV) (RG average: 0.22 to LG: 0.15). Over the term, RG showed a significant improvement in AdjV (from 0.11 (P1/2) to 0.15 (P5/6)), $F(2, 24)=9.92$, $p<0.001$, $\eta^2=0.453$, and ModV (from 0.20 (P1/2) to 0.24 (P5/6)), $F(2, 24)=4.86$, $p=0.017$, $\eta^2=0.288$, while the improvement was not significant in NV (from 0.67 (P1/2) to 0.79 (P5/6)), $p=0.414$, $\eta^2=0.071$. Finally, AdvV remained constant at around 0.09 throughout the semester.

6. Discussion

1. Is there a difference in the level of syntactic complexity as a result of the introduction of EAP classes?

Overall, there is a marked improvement in syntactic complexity when compared to LG. RG outperformed in the length of the production unit in terms of clause (MLC) and T-unit (MLT), the overall sentence complexity (C/S), the degree of subordination of the dependent clause (DC/C and DC/T), the amount of coordinate phrasing (CP/C and CP/T), and the degree of phrasal sophistication (CN/C and CN/T). However, due to the shorter mean sentence length produced by RG, one can also note lower scores compared to LG in MLS and TS. Also, despite the greater length of clauses (MLC), one can also see lower scores in CT/T and CT/T due to less frequency of T-units compared to LG. Finally, there was less of a range of verb phrases which resulted in a lower VP/T score.

When considering progress made by RG from within the fourteen syntactic complexity measures, one can also observe significant progress. In the following four measures: MLC, MLT, CT/T, and T/S, there was at least a marginally significant improvement. While the other measures did not establish any significant progress, it is also true to say that there were no negative results in the group's performance.

2. Is there a difference in the level of lexical complexity as a result of the introduction of EAP classes?

Again, when compared to LG, there is a marked improvement in syntactic complexity within RG. This was evident in LD, VS1, CVS1, VS2, NWD-50, NDW-ER50, NDW-ES50, TTR, MSTTR-50, CTTR, RTTR, LogTTR, D, LV, VV1, SVV1, CVV1, VV2, AdvV and ModV. Only in NDW, Uber, NV and AdjV is there more comparable results to LG. However, this is not unexpected as the former group was evaluated on oral performance while the latter was evaluated on written performance.

However, when observing alterations in lexical complexity over the semester, there would seem to be modest but not significant gains. Over the 26 lexical complexity measures, one can note some progress. Although there were positive changes in overall lexical sophistication generally, within the TTR, and specifically in terms of noun, verb, adjective, and adverb variation, such alterations were not in a linear manner. This may be due to the challenges of acquiring additional lexical resources to complete the essays or/and the task provided not matching the needs or interests of the students. However, it is also noted that over the semester, the quality and range of vocabulary did not reduce after completing their five-month stay in another country.

7. Conclusion

Providing an academic course that matches student needs after studying abroad for five months would seem to pose a challenge for any educator as the honeymoon period is over and students go back to the familiar study routines at their home university. It is, therefore, important to continue the momentum gained from experience studying abroad in their home environment. While it has been acknowledged that students generally are able to improve their fluency while studying overseas (Serrano, et al. 2011), the aim of this research was to observe how students could continue to nurture a positive belief that their English skills can still improve after their stay in another country. This was achieved by developing their English abilities in terms of syntactic and lexical complexity through a course that facilitates each student's ability to acquire EAP skills.

While results from this study vary, the overall impression from a wide pool of measures at the syntactic and lexical level would suggest at the very least that the level of motivation had been maintained. After all, it can be argued that there was no noticeable decline in any of the measurements. In fact, there was a marked improvement in syntactic complexity in terms of MLC, MLT, CT/T, and T/S. In terms of lexical complexity, significant improvements were obtained in the Uber index, the D measure, NV, AdjV, and ModV, while the difference between the first two and the last two tests was not significant in the other measures. The overall results, therefore, suggest that some progress has been made in the proficiency of students after their stay in another English-speaking country.

This progress resulted from carefully catering to student needs. This was achieved by providing an EAP course with discussion tasks with higher leveled vocabulary, which stretched their reading comprehension, and with debate tasks on contentious issues, which developed their critical thinking (Halpern, 2000; Lipman, 2003) on one hand, and writing materials that provided general structural and grammatical aspects of writing at an academic level, on the other. Due to the careful preparation of both current, tailored materials on the one hand and more formal, academic materials on the other, students were not only able to access the tasks autonomously and independently (Abdullah et al., 2022) but also seemed to cultivate critical thinking attitudes that were curious and committed to finding answers (Forrester, 2008; Thomas, 2011). As a result, students seemed motivated to invest themselves in this EAP course to improve their English abilities at a

more academic level.

While this paper cannot directly claim that students from a Japanese university could foster greater critical thinking skills and develop their academic skills, one can recommend the implementation of tasks that facilitate EAP skills to maintain their motivation (Iida, 2013; Kimura, 2011), and to foster critical thinking and argumentative writing skills (Tanaka, 2009) in order to encourage a steady improvement in their proficiency in English.

References

- Abdullah H., Harun, H., Wahab, N.A., Yan Juo, S., & M.A. Ali, 2022. Teaching English for academic purposes through project-based learning method and process writing approach. *International Journal Languages and Education*, 1 (1), 49-62.
- Ai, H., & X. Lu, 2013. A corpus-based comparison of syntactic complexity in NNS and NS university students' writing. In A. Díaz-Negrillo, N. Ballier, & P. Thompson (Eds.), *Automatic treatment and analysis of learner corpus data* (pp. 249-264). Amsterdam/Philadelphia: John Benjamins.
- Alquahtani, M., 2015. The importance of vocabulary in language learning and how to be taught. *International Journal of Teaching and Education*, 3 (3), 21-34.
- Alexander, O., Argent, S., & J. Spencer, 2008. EAP essentials. A teacher's guide to principles and practice. Reading: Garnet Publishing.
- Freebody, P., & Luke, A. (1990). Literacies programs: debates and demands in cultural context. *Prospect: Australian Journal of TESOL*, 5, 7-16.
- Alexander, O., 2012. Exploring teacher beliefs in teaching EAP at low proficiency levels. *Journal of English for Academic Purposes*, 11 (2), 99-111.
- Asención-Delaney, Y., & J. Collentine, 2011. A multidimensional analysis of a written L2 Spanish corpus. *Applied Linguistics*, 32 (3), 299-322.
- Benevento, C., & N. Storch, 2011. Investigating writing development in secondary school learners of French. *Assessing Writing*, 16 (2), 97-110.
- Berman, R., & L. Cheng, 2001. English academic language skills: Perceived difficulties by undergraduate and graduate students, and their academic achievement. *Canadian Journal of Applied Linguistics*, 4 (1), 25-40.
- Bruce, I., 2017. What knowledge do practitioners need to master to inform and direct not only their teaching but also, more broadly, their professional activities including understandings of academia in both its epistemological and sociological dimensions? *BALEAP ResTes Knowledge and the EAP Practitioner: A Symposium*. The University of Leeds.
- Bulté, B., & A. Housen, 2014. Conceptualizing and measuring short-term changes in L2 writing complexity. *Journal of Second Language Writing*, 26, 42-65.
- Campion, G.C., 2012. The learning never ends' Investigating teachers'experiences of moving from English for General Purposes to English for Academic Purposes in the UK context; What

- are the main challenges associated with beginning to teach EAP, and how can these challenges be overcome? (*Master's Thesis*). Retrieved from BALEAP.
- Caro, K., & N.R. Mendinueta, 2017. Lexis, Lexical Competence and Lexical Knowledge: A Review. *Journal of Language Teaching and Research*, 8 (2), 205–13.
- Chunling S. & F. Guoping, 2009. Process approach to teaching writing applied in different teaching models. *English Language Teaching*, 1 (2), 150–155.
- Cohen, J. (1988). *Statistical Power Analysis for the Behavioral Sciences* (2nd ed.). Hillsdale, NJ: Erlbaum.
- Crossley, S.A., & D.S. McNamara, 2014. Does writing development equal writing quality? A computational investigation of syntactic complexity in L2 learners. *Journal of Second Language Writing*, 26, 66–79.
- Crystal, D., 1982. *Profiling linguistic disability*. London: Edward Arnold.
- Ding, A., & I. Bruce, 2017. *The English for academic purposes practitioner*. Palgrave Macmillan.
- Ding, A., & G. Campion, 2016. EAP teacher development. In K. Hyland & P. Shaw (Eds.), *The Routledge handbook of English for Academic Purposes* (pp.1014–1018). London and New York: Routledge.
- Farrell, T. S. C., & D. Yang, 2017. Exploring an EAP teacher's beliefs and practices in teaching L2 speaking: A case study. *RELC Journal*, 48, 1–14.
- Forrester, J., 2008. Thinking creatively; thinking critically. *Asian Social Science*, 4 (5), 100–105.
- Frear, M. W., & J. Bitchener, 2015. The effects of cognitive task complexity on writing complexity. *Journal of Second Language Writing*, 30, 45–57.
- Gao, Y., & B. Bartlett, 2014. Opportunities and challenges for negotiating appropriate EAP in China. In I. Liyanage & T. Walker (Eds.), *English for Academic Purposes (EAP) in Asia: Negotiating appropriate practices in a global context* (pp.13–32). Sense Publishers.
- Halpern, D., 2000. Teaching for critical thinking: helping college students develop the skills and dispositions of a critical thinker. In M. D. Svinicki (Ed.), *Teaching and learning on the edge of the millennium: building on what we have learned. New Directions for Teaching and Learning* (pp.69–74). San Francisco, CA: Jossey-Bass.
- Hamp-Lyons, L., 2011. English for academic purposes. In E. Hinkel (Ed.), *Handbook of research in second language teaching and learning* (Vol. 2, pp. 89–103). Routledge.
- Han, J., Zhao, Y, Liu, M., & J. Zhang, 2021. The development of college English teachers' pedagogical content knowledge (PCK): from General English to English for Academic Purposes. *Asia Pacific Education Review*, 22, 609–62. Retrieved from: <https://doi.org/10.1007/s12564-021-09689-7>
- Hosogoshi, K., & S. Takahashi, 2015. The use of integrated listening, reading, speaking and writing tasks on students' productive skills in a university EAP course. *Professional and Academic English*, 45, 22–30.
- Hyland, K., & P. Shaw, 2016. *The Routledge handbook of English for academic purposes*. Routledge.
- Iida, T., 2013. Study abroad and university students' English proficiency: to have a better rela-

- tionship between home university and ESL overseas. *JACET Kansai Journal*, 15, 106-124.
- Iida, T. & S. Herder, 2019. Second Language Development Before, During, and After Study Abroad: A Longitudinal Study at a Japanese Women's University. *JACET Journal*, 63, 27-45.
- Intersegmental Committee of the Academic Senate, 2002. Academic Literacy: A Statement of Competencies Expected of Participants Entering California's Public Colleges and Universities. ICAS. Sacramento CA.
- Kim, M., Crossley, S. A., & K. Kyle, 2018. Lexical sophistication as a multidimensional phenomenon: Relations to second language lexical proficiency, development, and writing quality. *Modern Language Journal*, 102 (1), 120-141. Retrieved from: <https://doi.org/10.1111/modl.12447>
- Kimura, K., 2011. Effects of short-term overseas programs on Japanese university students studying Australia. *JABAET Journal* 14/15, 59-74.
- Kormos, J., & A. Trebits, 2012. The role of task complexity, modality, and aptitude in narrative task performance. *Language Learning*, 62 (2), 439-472.
- Kurt, S., 2020. Vygotsky's Zone of Proximal Development and Scaffolding. *Educational Technology*. Retrieved from: <https://educationaltechnology.net/vygotskys-zone-of-proximal-development-and-scaffolding/>
- Lafford, B., 2004. The effect of the context of learning on the use of communication strategies by learners of Spanish as a second language. *Studies in Second Language Acquisition*, 26 (2), 201-225.
- Larsen-Freeman, D., 2009. Adjusting expectations: The study of complexity, accuracy, and fluency in second language acquisition. *Applied Linguistics*, 30 (4), 579-589. Retrieved from: <https://doi.org/10.1093/applin/amp043>
- Lehner, A., 2011. Globalization, the English language, and international universities: The case of Japan. In B.C. Swaffield, & I. Guske (Eds.), *Global Encounters, Pedagogical Paradigms and Educational Practices* (pp.80-92). Newcastle Upon Tyne: Cambridge Scholars Publishing.
- Li, Y., 2017. A phenomenological ethnographic study of Chinese college English teachers' transition from teaching English for General Purposes to teaching English for Academic Purposes. A Thesis Submitted to The Education University of Hong Kong In Partial Fulfillment of the Requirement for The Doctor of Philosophy.
- Lipman, M., 2003. *Thinking in education*. Cambridge, U.K.: Cambridge University Press.
- Llanes, A., & C. Munoz, 2013. Age effects in a study abroad context: Children and adults studying abroad and at home. *Language Learning*, 63 (1), 63-90.
- Lorenzo, F., & L. Rodriguez, 2014. Onset and Expansion of L2 Cognitive Academic Language Proficiency in Bilingual Settings: CALP in CLIL. *System* 47, 64-72.
- Lu, X., 2010. Automatic analysis of syntactic complexity in second language writing. *International Journal of Corpus Linguistics*, 15 (4), 474-496.
- Lu, X., 2011. A corpus-based evaluation of syntactic complexity measures as indices of college level EDL writers' language development. *TESOL Quarterly*, 45 (1), 36-62. Retrieved from:

<https://doi.org/10.5054/tq.2011.240859>

- Lu, X., & H. Ai, 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., & Dur'an, P. 2004. *Lexical diversity and language development: Quantification and assessment*. Houndmills, England: Palgrave MacMillan.
- Mancilla, R. L., N. Polat, & A. O. Akcay, 2015. An Investigation of Native and Nonnative English Speakers' Levels of Written Syntactic Complexity in Asynchronous Online Discussions. *Applied Linguistics*, 39, 112-134.
- Martin, P., 2014. Teachers in transition: The road to EAP. In P. Breen (Ed.), *Cases on teacher identity, diversity and cognition in higher education* (pp. 287-315). IGI Global.
- Michel, M., 2017. Complexity, Accuracy and Fluency (CAF). In S. Loewen & M. Sato (Eds.), *The Routledge Handbook of Instructed Second Language Acquisition*. London: Routledge.
- Nation, I.S.P., 2013. *Learning vocabulary in another Language*. Cambridge: Cambridge University Press.
- Nishikawa, M., Benom, C., Cupchak, D., Cutrone, P., Collins, W., Datzman, B., Hiroe, A., Kyaw-Oo, M, Lucas, S., Moji, K., Nishihara, T., Sullivan, S. A., Tada, A., Tanigawa S., & T. Zhang, 2023. Investigating Japanese EFL University Students' Needs, Motivations and Perceptions of the New EAP Program. *Nagasaki University Graduate School Journal of Multicultural Sociology*, 9, 336-361.
- Nishino, T., 2008. Japanese secondary school teachers' beliefs and practices regarding communicative language teaching: An exploratory survey. *JALT Journal*, 30 (1), 27.
- Norris, J. M., & L. Ortega, 2009. Towards an organic approach to investigating CAF in instructed SLA: *The case of complexity*. *Applied Linguistics*, 30 (4), 555-578. <https://doi.org/10.1093/applin/amp044>
- Ortega, L., 2003. Syntactic complexity measures and their relationship to L2 proficiency: A research synthesis of college-level L2 writing. *Applied Linguistics*, 24 (4), 492-518, <https://doi.org/10.1093/applin/24.4.492>
- Pandey, G.P. 2019. English for Academic Purposes: Theory, Trends and Practices. Nepal Journals Online. <https://doi.org/10.3126/ed.v29i0.32562>
- Park, J-H., 2017. Syntactic complexity as a predictor of second language writing proficiency and writing quality. *Doctorate dissertation* submitted to submitted to Michigan State University.
- Read, J., 2000. *Assessing vocabulary*. Oxford: Oxford University Press.
- Sakui, K., 2004. Wearing two pairs of shoes: Language teaching in Japan. *ELT Journal*, 58 (2), 155-163.
- Sasaki, M., 2007. Effects of study abroad experiences on EFL writers: a multiple-data analysis. *The Modern Language Journal*, 91 (4), 602-620.
- Sayed, S.T., 2018. English for Academic Purposes: English for general skills Writing Course. Bulletin of Advanced English Studies, 1 (1), 2018, 87-110.
- Serrano, R., Llanes, A., & E. Tragant, 2011. Analyzing the effect of context of second language learning: domestic intensive and semi-intensive courses vs. study abroad in Europe. *System*,

- 39, 133-143.
- Serrano, R., Tragant, E., & A. Llanes, 2012. A longitudinal analysis of the effects of one year abroad. *The Canadian Modern Language Review*, 68 (2), 183-163.
- Schmitt, X., Kubler, S., Robert, J., Papadakis M., & LeTraon, Y. 2019. A Replicable Comparison Study of NER Software: StanfordNLP, NLTK, OpenNLP, SpaCy, Gate. *2019 Sixth International Conference on Social Networks Analysis, Management and Security (SNAMS)*. Conference Paper, Granada, Spain. DOI: 10.1109/SNAMS.2019.8931850
- Spring, R., & M. Johnson, 2022. The possibility of improving automated calculation of measures of lexical richness for EFL writing: A comparison of the LCA, NLTK and SpaCy tools. *System 106*. ISSN 0346-251X, DOI: 10.1016/j.system.2022.102770.
- 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, 103-118.
- Tanaka, K., 2009. The learning outcomes of an academic writing course: A study of Japanese university students. *JACET Journal*, 49, 123-140.
- Tanaka, K., 2010. The philosophical, political, and the practical dimensions of English for academic purposes education: A focus on critical thinking. *Kokusai Gaku Kenkyu*, 37, 67-77.
- Tanaka, K., & R. Ellis, 2003. Study abroad, language proficiency, and learner beliefs about language learning. *JALT Journal*, 25 (1), 63-85.
- Thomas, T.A., 2011. Developing first year students' critical thinking skills. *Asian Social Science*, 7 (4), 26-35.
- Vyatkina, N., Hirschmann, H., & F. Golcher, 2015. Syntactic modification at early stages of L2 German writing development: A longitudinal learner corpus study. *Journal of Second Language Writing*, 29, 28-50.
- Wrigley, H.S., 1998. Knowledge in action: The promise of project-based learning. *Focus on Basics*, 2 (D), 13-18.
- Wolfe-Quintero, K., Inagaki, S., & H.-Y. Kim, 1998. *Second Language Development in Writing: Measures of Fluency, Accuracy, and Complexity*. University of Hawai'i Press.
- Yoon, H. J., & C. Polio, 2016. The linguistic development of students of English as a second language in two written genres. *TESOL Quarterly*, 275-301. Retrieved from: doi: 10.1002/tesq.296

Note

- 1) The first author was in charge of running the EAP course including the design and preparation of training materials and of writing the manuscript, while the second author focused on written data management and lexical and syntactic analyses.
- 2) According to Cohen (1988), the effect size is small if the value of η^2 varies around 0.01, medium if it varies around 0.06, and large if it varies around 0.14.
- 3) The p -values were adjusted for the post-hoc multiple comparisons using Bonferroni correction.