

SYNTACTIC COMPLEXITY IN THE READING MATERIALS OF ENGLISH FOR ACADEMIC PURPOSES LEVELS 1 – 3

Widdy Wijanti

Sampoerna University, Jakarta, Indonesia

wijanti.widdy@gmail.com

DOI: doi.org/10.24071/llt.2017.200204

received 7 July 2017; revised 12 August 2017; accepted 5 September 2017

Abstract

In Indonesia, English is still considered as a foreign language and has become a crucial subject of study especially in the university level. For this reason, English for Academic Purposes has been conducted in the first year of college level for many years. Unfortunately, although many Asian countries including Indonesia have run the EAP course, the output is that there are still many Indonesian students who do not meet the vocabulary size and syntactic complexity that are expected while their learning process in the university. This results lower grades that they have in their assignments. Therefore, the recent study is aimed at evaluating the reading materials of EAP, especially in measuring the syntactic complexity containing in the texts as it is strongly believed in English learning that a good language output comes from a good language input. The data is taken from the collections of reading materials taken from EAP course Level 1, Level 2, and Level 3 which are compulsory subjects for students at Sampoerna University in their first 2 years of study. The data then is processed using the Syntactic Complexity Analyzer (Ai & Lu, 2013). The findings showed that the reading materials of EAP course Level 3 text is mainly suggested to be reviewed and revised in order to fulfill the five categories of syntactic complexity, i.e. the length of production unit, the sentence complexity, the subordination, the coordination, and the particular structure.

Keywords: vocabulary size, syntactic complexity, syntactic complexity analyzer, reading text

Introduction

Over the past decades, English has become a mandatory subject that should be taken in the first year of university life in Indonesia because students have to read imported textbooks and write assignments in English which is not their first language. In addition, college students who have good English competence have wider opportunities for student exchange program that will also enhance their learning experience in other countries. Several years later, they will use their knowledge and experience in the work life, and English is one of the important factors that support them in their career and further education. This is due to the fact that English is a language that is used internationally, which means it is understood and used by a lot of people in the world, especially in the

education and work life (Crystal, 2012, pp. 3-4). For that purpose, a number of lecturers use English (as a second language) as a medium of instructions in delivering their university subjects (Dudley-Evans, 1998, p. 38).

English for Academic Purposes has been taught in the universities for a long time, but especially in Indonesia, the students are still unable to reach the minimum target of English competence, that is to acquire the minimum size of Academic Word List as suggested by Nation (2001) i.e. as much as 10% or 570 word families or 2,570 words (p. 18). Based on the research conducted by Schmitt (2010) about the vocabulary size produced by university students between some Asian countries, i.e. Japan, China, Indonesia, and Oman, Indonesian students have the lowest number of vocabulary size after taking the similar number of instruction hours in English. This may be the result of poor quality of the materials given for learning English. One of several problems related to the materials is the lack of exposure in syntactic complexity in the reading materials given in the texts at university level. Therefore, it is necessary to conduct research on reviewing the materials given to the students at university level in Indonesia.

Besides knowing enough vocabulary needed to be used in the receptive skills (listening and reading skills), language learners need to be able to use the words they know in the productive skills (speaking and writing skills) to any forms of phrases, clauses, sentences, and even paragraphs by combining different words into one unity. Cutler (1983, p. 45) mentions that "Semantic complexity covers a fairly wide range of variations between words". Lu (2009, p. 4) also adds that language users have high syntactic complexity if they are able to use the vocabulary to produce simple sentences, compound sentences, and complex sentences. Simple sentences are produced when the words are combined together containing one subject (S), one predicate (P), one object (O) (optional), and some complements (C) (optional). For example: John writes a letter every month. John (S), writes (P), a letter (O), every month (C). Compound sentences are produced when there are two or more simple sentences combined into one sentence using one or more connectors, i.e. for, and, nor, but, or, yet, and so. For example: *John writes a letter every month* are combined with *Jane writes a letter every month*. Because there are several parts of speech that are the same, i.e. *writes* (P), *a letter* (O), *every month* (C), there are only the different subjects that are combined together using and that is to become *John and Jane write a letter every month*. The next is complex sentences, this kind of sentence can be produced when there are two or more simple sentences combined using one or more subordinators, i.e. *who*, *which*, *that*, *whom*, *because*, *therefore*, *while*, etc. For example: the first simple sentence is *John writes a letter every month*. The second simple sentence is *John likes a blue T-shirt*. From the second sentence, it can be seen that it is a sentence that explains more about John, therefore, the subordinator that is appropriate is using relative clause *who*. The sentence combination becomes *John **who likes a blue T-shirt** writes a letter every month*. The italic-bold words are called dependent clause because it is attached with the subordinator; while the italic words are called independent clause. In other words, complex sentences are the combination between independent clause and dependent clause. Therefore, if

the students are able to produce more complex sentences, their English proficiency is higher than the ones who can only produce simple sentences.

Syntactic Complexity means “the range and degree of syntactic structures that surface in language production and has been recognized as a very important construct in second language writing teaching and research” (Ortega, 2003) as quoted from Ai and Lu (2013, p. 249). In other words, if the language learners have a good use of syntactic complexity, then they will produce good variation of sentence structure, such as the use of compound and complex sentences, and also the varied use of transition signals which make the writing seem more coherent and grammatically correct. It is then shown from how many dependent clauses that the language learners can produce in their written works. It does not matter how long the sentences are as long as they produce many word/sentence elements in their works, i.e. the use of complex sentences, subordinators, coordinators, and good sentence structure. The present study uses *Syntactic Complexity Analyzer* (SCA) suggested by Ai and Lu (2013) in order to know whether the syntactic complexity performance in reading materials selected in the university, that the researcher is teaching, is suitable to the level of EAP course.

The researcher works at Sampoerna University which conducts EAP courses for 3 levels in the students’ first years of study. The materials are chosen and compiled from several sources from books or websites that are suitable with the topics in the syllabi which are designed by Institute of Languages and Communication (in which the researcher takes part in the material preparation process). However, this materials compilation technique has some drawbacks. Subjectivity is possible to happen as mentioned by Jordan (1997, p. 127) that one of several factors that materials provider choose the materials is “attractive”. In other words, the materials compiler tend to choose the materials which seem highly interesting to him and probably seem less interesting to the students. This gap will cause the materials delivery less effective because students are not in to what are being discussed. For this reason, the present study will evaluate the compilation of the reading materials whether they are chosen objectively and appropriate with the students’ levels.

This is significant to evaluate the reading materials of EAP course levels 1, 2, and 3 whether the selected compilation has already been in order according to the learners’ leveling or not because the compilation process of the reading materials was done only based on the chosen topics listed in the syllabus design, the source books available in the library at the SU, and the online materials provided to be downloaded. In addition, the result will be some indicators for ILC lecturers whether the syntactic complexity of the materials is suitable for the Sampoerna University students to improve their English competence and whether there is compulsory to augment the EAP materials for the future use.

Although there have been several studies of Syntactic Complexity in second language oral and written production conducted by Laufer & Nation (1995), and Lu (2012) (for oral production), not many studies have been done to analyze the quality of reading texts in terms of Syntactic Complexity. Therefore, the present study aims to know the quality of the reading materials selected for EAP course in each level (levels 1, 2, and 3) whether the Syntactic Complexity is suitable to

generate the English competence that the learners have, “What are the syntactic complexity differences between the reading materials in EAP course level 1, level 2, and level 3?”. In other words, suitable means whether the difficulty level of reading materials are gradually increasing or not, or whether the syntactic complexity is generating them to improve their language competence. By doing such research, there is a possibility whether there is a need to augment the materials in the future. Moreover, the study in the SC can give another practical contribution in the selection of sentence types to be successfully learned by language learners in the form of various contexts. In addition, the study can also give some suggestions on any types of reading materials that are fruitful for learners to use in the language learning process in order to develop their grammatical correctness in sentence level, coherence in paragraph level, and unity in a bigger frame, for example in an essay. If they are able to improve their language competence from the reading materials provided, they will be successful in their academic life.

After knowing enough frequently-used vocabulary, language learners are expected to be able to combine those words into good sentences using good sentence structure, which can be seen as syntactic complexity. Ortega (2003) as quoted from Ai and Lu (2013) defines that syntactic complexity is “the range and degree of syntactic structures that surface in language production recognized as a very important construct in second language writing teaching and research” (p. 3). Therefore, if someone is able to produce various kinds of sentences in their written products and grammatically correct, he or she is considered having good syntactic complexity. Various kinds of sentences are including the sentences using various sentence connectors, i.e. *for*, *and*, *nor*, *but*, *or*, *yet*, *so*, *who*, *which*, *that*, *whom*, *whose*, *because*, *because of*, *since*, *although*, etc. This means that a good syntactic complexity can be shown by the frequent use of the correct combination between independent clauses and dependent clauses, as known as compound sentences and complex sentences.

There are several computer programs to measure the syntactic complexity level as cited in Lu (2010). For example, computerized profiling (Long et al., 2008), Coh-Metrix (Graesser et al., 2004), and D-Level Analyzer (Lu, 2009) (p. 4). Those advance programs are usually used to measure people’s language development for productive skills, especially in written forms to see the second language learners’ progress after several weeks of language learning process, or in spoken forms to find out the range of sentences that can be produced, for example by people who have limitations with their language production in the brain system (people who are suffering from Alzheimer disease, or people in early ages—toddlers in their language acquisition period).

The present study measures the language quality of language input taken from reading materials that are used in the learning process of EAP course in the university level. The result will be used to determine whether language exposure provided in the text is suitable with the leveling of the course, so there can be some adjustments done for betterment. The measure that the researcher will use is the one that was developed by Lu (2009). She used this analyzer to measure the language development of toddlers: the syntactic complexity analyzer which is also

called as D-Level Analyzer (Ai & Lu, 2013, p. 4). In details, the present study will use the fourteen syntactic complexity measures, provided in the D-Level Analyzer, as reviewed by Wolfe-Quintero et al. (1998) and Ortega (2003) in students' written products as cited in Ai and Lu (2013, p. 5).

Method

This research uses quantitative approach according to Dörnyei (2007) because it "involves data collection procedures that result primarily in numerical data which is then analysed primarily by statistical methods," (p. 24). This research will use an online-computerized programs suggested by Lu (2013) that is Syntactic Complexity Analyzer to measure the Syntactic Complexity of the reading materials compiled for the three levels of EAP course at Sampoerna University.

In addition, this advance program is displaying numeric results; therefore the analysis will be based on quantitative approach where the list of vocabulary ratios of each data will be analyzed according to the numeric results from the computerized programs that measure the syntactic complexity which is parts of statistical methods. Moreover, the results will be analyzed to see the difference of Syntactic Complexity between the levels when compared.

4.2 Source of Data

- a. Printed Sources: (1) Brook-Hart, G. (2004). Academic Reading Passage 1. In *Cambridge instant IELTS: Ready-to-use tasks and activities* (pp. 33-36). Cambridge: Cambridge University Press, (2) Cambridge. (2002). *Cambridge IELTS 3: Examination papers from the University of Cambridge ESOL Examinations*. Cambridge: Cambridge University Press (pp. 8-11), (3) Cambridge. (2006). *Cambridge IELTS 5: Authentic examination papers from University of Cambridge ESOL examinations*. United Kingdom: Cambridge University Press, (4) Cambridge. (2009). *Cambridge IELTS 7: Authentic examination papers from University of Cambridge ESOL examinations*. Cambridge: Cambridge University Press, (5) Cambridge. (2013). *Cambridge IELTS 9: Authentic examination papers from Cambridge ESOL*. Cambridge: Cambridge University Press, (6) Hallows, R., Lisboa, M., & Unwin, M. (2013). *IELTS express: Intermediate course book*. Andover, Hampshire [England: Heinle Cengage Learning], (7) Jakeman, V., & McDowell, C. (2012). *Step up to IELTS*. UK: Cambridge University Press, (8) Lougheed, L., & Barron's Educational Series, Inc (2010). *IELTS practice exams: With audio cds*. Hauppauge, NY: Barron's Educational Series, (9) May, P. (2004). *IELTS practice tests: With explanatory key*. Oxford: Oxford University Press, (10) Phillips, D. (2003). *Preparation course for the TOEFL test*. USA: Pearson: Longman, and (11) Douglas, N. (2010). *Reading explorer 3: [student book]*. Boston: Heinle Cengage Learning.
- b. Online Sources: (1) Clark, D. (2011, July 12). Visual, auditory, and kinaesthetic learning styles (VAK). Retrieved July 3, 2014, from <http://www.nwlink.com/~donclark/hrd/styles/vakt.html>, (2) Gravity. (2013, June 8). Men vs. women: Who's saving and who's spending? |visualnews.com. Retrieved April 14, 2014, from

http://www.huffingtonpost.com/visualnewscom/men-vs-women-whos-saving-and-whos-spending_b_3714079.html, (3) Identify Irrelevant Sentences in Paragraphs 7. (2015). Retrieved from <http://www.grammarbank.com/identify-irrelevant-sentences.html>, (4) Paris offers free public transport to reduce severe smog - BBC News. (2014, March 14). Retrieved March 14, 2014, from <http://www.bbc.com/news/world-europe-26574623>, (5) Pravas, V. S. (2013). Different types of pollution: Causes of water, air, soil, thermal, light and noise pollution. Retrieved July 3, 2014, from <http://readanddigest.com/what-are-different-types-of-pollution/>.

In this present study, there were three steps taken in collecting and analyzing the data. The first step of data collection was re-typing the modules using Microsoft Word. This was done in order to make the texts easier to process in the software because there were some pdf files which could not be copy-pasted directly to the txt file. After that, the researcher converted the data from Word file into txt file with encoding code: UTF-8 to be able to run with the program by copy-pasting the whole texts compiled for each level of EAP course and saved the txt file in a different file, for example: Level1.txt was for EAP course level 1 containing 11 different texts, Level2.txt was for EAP course level 2 containing 8 various texts, and Level3.txt was for EAP course level 3 containing 10 lengthy texts.

The next step was to process the syntactic complexity of the reading texts. The researcher focused on measuring the syntactic complexity using the Syntactic Complexity Analyzer (SCA) as suggested by Lu (2013) and the program was used online via <http://www.personal.psu.edu/xxl13/downloads/l2sca.html>. The researcher used the online demo which has the same function—determining the syntactic complexity of each compilation of reading texts in Level 1, Level 2, and Level 3. Because the program cannot process the data which contain less than 50 words and more than 1500 words, the data in txt.file encoding UTF-8 needed to be separated into several txt files. Each Level 1, Level 2, and Level 3 was separated into 5 separated txt files, and then they were saved in zip files separately. With the zip files then the data could be uploaded to be processed using the SCA, that is Web-based L2 Syntactic Complexity Analyzer (Batch Mode) covering “(1) length of production units, (2) amounts of coordination, (3) amounts of subordination, (4) degree of phrasal sophistication and overall sentence complexity”. Then, the results of all Level 1, Level 2, and Level 3 were combined and analyzed to determine the syntactic complexity of the reading texts.

Finally, the last step was to find and analyze the differences of syntactic complexity between levels 1, 2, and 3.

Findings and Discussion

The Comparison of Syntactic Complexity of the Reading Texts

Table 1. The Comparison of Syntactic Complexity between Levels 1, 2, and 3

Syntactic Complexity Analyzer Result				
MEASURE	CODE	LEVEL 1	LEVEL 2	LEVEL 3
<i>Type 1: Length of production unit</i>				
Mean length of sentence	MLS	1,041.991	1,079.367	1,102.705
Mean length of T-unit	MLT	940.922	1,003.059	967.992
Mean length of clause	MLC	606.339	691.334	670.166
<i>Type 2: Sentence complexity</i>				
Sentence complexity ratio	C/S	86.029	78.813	83.384
<i>Type 3: Subordination</i>				
T-unit complexity ratio	C/T	77.736	72.983	73.165
Complex T-unit ratio	CT/T	2.3479	1.9821	2.1829
Dependent clause ratio	DC/C	2.0239	1.8964	1.7716
Dependent clauses per T-unit	DC/T	3.1665	2.7688	2.6483
<i>Type 4: Coordination</i>				
Coordinate phrases per clause	CP/C	1.4958	2.0245	1.7108
Coordinate phrases per T-unit	CP/T	2.2989	2.8579	2.4623
Sentence coordination ratio	T/S	55.341	53.839	57.024
<i>Type 5: Particular structures</i>				
Complex <u>nominals</u> per clause	CN/C	78.100	92.831	84.310
Complex <u>nominals</u> per T-unit	CN/T	120.970	135.783	122.239
Verb phrases per T-unit	VP/T	110.839	107.896	104.218

Table 1 displays the result comparison of syntactic complexity of the reading texts used in the Levels 1, 2, and 3 of EAP course. As mentioned earlier, the fourteen items analyzed in the text can be grouped into five types, i.e. *the length of the production unit*, *the sentence complexity*, *the subordination*, *the coordination*, and *the particular structure*. For the first type, i.e. The Length of the Production Unit is shown from the scores of the MLS, MLT, and MLC. Table 1 shows that the total MLS (Mean length of sentence) of Level 1 (1,041.991) is lower than that of Level 2 (1,079.367), and the MLS number is lower than that of Level 3 (1,102.705). This means that there is an increase in the total length of the sentences provided in the text as the level gets higher. This increase is good because when the English learners are in the higher level, they need to read more

sentences in order to expand their vocabulary and sentence structure knowledge. In addition, the calculation using the Chi-square shows that the chi-square value between MLS Level 1 and Level 2 is 0.42, which means that there is no significant difference in the value. While the chi-square value between MLS Level 2 and MLS Level 3 is 0.62, which means that there is also no significant difference in the value.

Next, Table 1 shows that the total MLT (Mean Length of T-Unit) of Level 1 (940.922) is lower than that of Level 2 (1,003.059). The total MLT (Mean Length of T-Unit) of Level 2 (1,003.059) is higher than that of Level 3 (967.992). This means that in Level 1 besides having more sentence length, the students are also reading more main clauses and dependent clauses. In Level 3, the students read less main clauses and dependent clauses. It is not good because as the level gets higher, the MLT number has to get higher in order to provide more exposure in main clauses and dependent clauses. The calculation using the Chi-square shows that the chi-square value between Level 1 and Level 2 is 0.16, which means that there is no significant difference in the value. The chi-square value between Level 2 and Level 3 is 0.43, which means that there is also no significant difference in the value.

Next, Table 1 also shows that the total MLC (Mean Length of clause) of Level 1 (606.339) is lower than that of Level 2 (631.334), and the total MLC (Mean Length of clause) of Level 2 (631.334) is higher than that of Level 3 (670.166). This means that between Level 1 and Level 2, there is a good increase in the MLC, while between Level 2 and Level 3, the decrease in number is not good because it is not suitable with the ideal concept of syntactic complexity where the higher level the students are, the more clauses that they need to learn from the reading materials which means the number of MLC should get higher as the level increases. The calculation using the Chi-square shows that the chi-square value between Level 1 and Level 2 is 0.02 which means that there is a significant difference in the value. The chi-square value between Level 2 and Level 3 is 0.57, which means that there is no significant difference in the value.

The numbers give a recommendation to have the materials of Level 3 to be augmented especially in the subtype of Mean Length of T-Unit (MLT) and Mean Length of clause (MLC) where the numbers should get higher because the revision can give more benefits for students in their learning process, that is for example as a good exposure of written texts.

The second type, i.e. the Sentence Complexity Ratio, is shown from the scores of the C/S. Table 1 shows that the total C/S (Sentence Complexity Ratio) of Level 1 (86.029) is higher than that of Level 2 (78.813), and the total C/S (Sentence Complexity Ratio) of Level 2 (78.813) is lower than that of Level 3 (83.384). This means that there is a decrease in the C/S in the Level 2 and the increase in the C/S in the Level 3. This fluctuating number in C/S is not good because the sentence of complexity ratio needs to increase as the level increases in order to raise the challenge for students to familiarize with the sentence structure. Moreover, the calculation using the Chi-square shows that the chi-square value between Level 1 and Level 2 is 0.57, which means that there is no significant

difference in the value. The chi-square value between Level 2 and Level 3 is 0.72, which means that there is no significant difference in the value.

The third category, i.e. the Subordination, is shown from the scores of the C/T, CT/T, DC/C and DC/T. Table 1 shows that the total C/T (T-unit complexity ratio) of Level 1 (77.736) is higher than that of Level 2 (72.983). While the total C/T (T-unit complexity ratio) of Level 2 (72.983) is lower than that of Level 3 (73.165). The fluctuating number of the T-unit complexity ratio (C/T) is not good because the C/T number should increase as the level gets higher. Therefore, the C/T of Level 1 should be lower than that of the Level 2, and the C/T of Level 2 should be lower than that of the Level 3. Students need to have more exposure in the complexity of main clause and dependent clauses as their English level gets higher. In addition, the calculation using the Chi-square shows that the chi-square value between Level 1 and Level 2 is 0.70, which means that there is no significant difference in the value. While the chi-square value between Level 2 and Level 3 is 0.99, which means that there is no significant difference in the value.

Next, from the table, it also shows that the total CT/T (Complex T-Unit ratio) of Level 1 (2.3479) is higher than that of Level 2 (1.9821). While the total CT/T of Level 2 (1.9821) is lower than that of Level 3 (2.1829). This fluctuating number means no good because the number of CT/T is supposed to be increasing as the level gets higher. The calculation using the Chi-square shows that the chi-square value between Level 1 and Level 2 is 0.86, which means that there is no significant difference in the value. While the chi-square value between Level 2 and Level 3 is 0.92, which means that there is also no significant difference in the value.

Lastly, the table also displays the total DC/C (Dependent clause ratio) of Level 1 (2.0239) which is higher than that of Level 2 (1.8964), and the total DC/C of Level 2 (1.8964) is higher than that of Level 3 (1.7716). This means that the difference in the use of dependent clause is not good because it is decreasing as the level gets higher. The calculation using the Chi-square shows that the chi-square value between Level 1 and Level 2 is 0.95, which means that there is no significant difference in the value. While the chi-square value between Level 2 and Level 3 is 0.95, which means that there is also no significant difference in the value.

The table also displays the total DC/T (Dependent clauses per T-unit) of Level 1 (3.1665) which is higher than that of Level 2 (2.7688). The total DC/T of Level 2 (2.7688) is higher than that of Level 3 (2.6483). This means that there is a decrease in the production of dependent clauses from Level 1 to Level 2 and from Level 2 to Level 3 where it is supposed to be increasing as the level gets higher. The calculation using the Chi-square shows that the chi-square value between Level 1 and Level 2 is 0.87, which means that there is no significant difference in the value. The chi-square value between Level 2 and Level 3 is 0.96, which means that there is also no significant difference in the value.

From the data given, it can be concluded that in the subordination category, the Level 2 text contains less beneficial materials for students compared to the Level 1 text because the complexity in the subtypes (C/T—T-unit complexity

ratio, CT/T—complex T-Unit ratio, DC/C—dependent clause ratio, and DC/T—dependent clauses per T-unit) tend to decrease as the level gets higher. Moreover, when the Level 2 texts are revised, the Level 3 texts should also have some adjustments to suit the criteria of a good text which provide good syntactic complexity where the number of C/T (T-Unit complexity ratio), CT/T (Complex T-Unit ratio), DC/C (Dependent clause ratio), and DC/T (Dependent clauses per T-unit) should get higher as the level increases in order to give more benefits for students in their learning process, that is for example as a good exposure of written texts.

The fourth category, i.e. the Coordination, is shown from the scores of the CP/C, CP/T, and T/S. Table 1 shows that the total CP/C (Coordinate phrases per clause) of Level 1 (1.4958) is lower than that of Level 2 (2.0245). While the total CP/C of Level 2 (2.0245) is higher than that of Level 3 (1.7108). This fluctuate number means that the Level 3 texts should be revised in order to have higher CP/C number than that of the Level 2 texts. The Level 2 text has already provided good exposure in the use of coordinate phrases per clause for students in their learning process because the ratio gets higher as the level of competence gets higher. The calculation using the Chi-square shows that the chi-square value between Level 1 and Level 2 is 0.78, which means that there is no significant difference in the value. While the chi-square value between Level 2 and Level 3 is 0.87, which means that there is no significant difference in the value.

The table also shows that the total CP/T (Coordinate phrases per T-unit) of Level 1 (2.2989) is lower than that of Level 2 (2.8579), and the total CP/T of Level 2 (2.8579) is higher than that of Level 3 (2.4623). This fluctuating number of CP/T means that the Level 3 texts need to be augmented because the number has to be higher than that of the Level 2 texts. The Level 2 text has provided good ratio in the use of coordinate phrases per T-unit because it gives more use of it as an exposure to students. The calculation using the Chi-square shows that the chi-square value between Level 1 and Level 2 is 0.81, which means that there is no significant difference in the value. The chi-square value between Level 2 and Level 3 is 0.86, which means that there is also no significant difference in the value.

Lastly, the table also displays the total T/S (Sentence coordination ratio) of Level 1 (55.341) which is higher than that of Level 2 (53.839), and the total T/S of Level 2 (53.839) which is lower than that of Level 3 (57.024). This fluctuating number of T/S means that the Level 2 texts and the Level 3 texts need to be augmented because the number has to be higher as the level gets higher in order to give more frequent sentence coordination to give more challenges to students in learning English. The calculation using the Chi-square shows that the chi-square value between Level 1 and Level 2 is 0.81, which means that there is no significant difference in the value. The chi-square value between Level 2 and Level 3 is 0.76, which means that there is also no significant difference in the value.

Based on the result in the subtypes of the Coordination category, the materials of Level 2 and Level 3 need to be augmented especially in the subtype of T/S (Sentence coordination), Coordinate phrases per clause (CP/C) and

Coordinate phrases per T-unit (CP/T), in order to give more samples for students in their learning process, i.e. a good exposure on how to provide good sentence combination.

The last category, i.e. the Particular Structure, is shown from the scores of the CN/C, CN/T and VP/T. Table 1 shows that the total CN/C (Complex nominals per clause) of Level 1 (78.100) is lower than that of Level 2 (92.831). The total CN/C of Level 2 (92.831) is higher than that of Level 3 (84.310). This means that the Level 2 text provides good use of complex nominals per clause (CN/C) and it also provides more challenge for students to understand the written texts as the level gets higher compared with that of the Level 1 text. The Level 3 text does not provide good use of complex nominals per clause (CN/C) and it also provides less challenge for students to understand the written texts as the level gets higher compared with that of the Level 2 text. The calculation using the Chi-square shows that the chi-square value between Level 1 and Level 2 is 0.26, which means that there is no significant difference in the value. The chi-square value between Level 2 and Level 3 is 0.52, which means that there is also no significant difference in the value.

Next, from the table, it also shows that the total CN/T (Complex nominals per T-unit) of Level 1 (120.970) is lower than that of Level 2 (135.783). The total CN/T of Level 2 (135.783) is higher than that of Level 3 (122.239). This means that the Level 2 text provides good use of complex nominals per T-unit (CN/T) and it also provides more challenge for students to understand the written texts as the level gets higher compared with that of the Level 1 text. The Level 3 text does not provide good use of complex nominals per T-unit (CN/T) and it also provides less challenge for students to understand the written texts as the level gets higher compared with that of the Level 2 text. The calculation using the Chi-square shows that the chi-square value between Level 1 and Level 2 is 0.36, which means that there is no significant difference in the value. The chi-square value between Level 2 and Level 3 is 0.40, which means that there is no significant difference in the value.

The last, the table also displays the total VP/T (Verb phrases per T-unit) of Level 1 (110.839) which is higher than that of Level 2 (107.896). This means that the Level 2 text does not provide a lot of good uses of verb phrases per T-unit (VP/T) and therefore, it does not provide more challenge for students to understand the written texts as the level gets higher compared with that of the Level 1 text. While the total VP/T (Verb phrases per T-unit) of Level 2 (107.896) is higher than that of Level 3 (104.218). This means that the Level 3 text does not provide good use of verb phrases per T-unit (VP/T) and therefore, it does not provide more challenge for students to understand the written texts as the level gets higher compared with that of the Level 2 text. The calculation using the Chi-square shows that the chi-square value between Level 1 and Level 2 is 0.84, which means that there is no significant difference in the value. The chi-square value is 0.80, which means that there is also no significant difference in the value.

The result in subtypes of the Particular Structure category gives a recommendation to have the materials of Level 2 and Level 3 augmented especially in the subtype of the use of verb phrases per T-unit (VP/T), the use of

complex nominals per clause (CN/C), the use of complex nominals per T-unit (CN/T), and the use of verb phrases per T-unit (VP/T) in order to give more samples for students in their learning process, i.e. a good exposure on how to provide good variety of the verb phrases use.

Based on the comparison data of syntactic complexity between Level 1, Level 2, and Level 3, the first category, the length of production unit between Level 1, Level 2, and Level 3, especially the sentence length (MLS) is already suitable with the ideal percentage, i.e. Level 1 is shorter than Level 2, and Level 2 is shorter than Level 3. Therefore, the Level 3 texts need to be augmented especially to increase the number of sentence length (MLS), T-unit length (MLT), and clauses length (MLC) in order to increase the challenge as the level gets higher. For the second category, the sentence complexity (C/S), the Level 1 has more sentence complexity than the Level 2. This result is not suitable with the ideal concept. As the level gets higher, students need to have higher sentence complexity. Therefore, the Level 2 text needs to be reviewed. However, the Level 2 has less sentence complexity than the Level 3. This result is suitable with the ideal concept. As the level gets higher, students need to have higher sentence complexity. Therefore, if the Level 2 is augmented, there needs to be some adjustments in the Level 3 that is to have higher C/S than that of the Level 2 texts.

The next category is the subordination. For all subtypes, the Level 1 has more numbers than the Level 2. This is not suitable with the ideal concept where Level 2 should have higher number of subordination in the T-unit complexity ratio, complex T-unit ratio, dependent clause ratio, and dependent clauses per T-unit in order to expose more variation of sentence types. The comparison between Level 2 and Level 3, for two out of four subtypes, i.e. the T-unit complexity ratio (C/T) and the complex T-unit ratio (CT/T), the Level 2 has smaller numbers than that of the Level 3. This is in line with the ideal concept where Level 3 should have higher number of subordination in all four subtypes. However, the numbers in the subtypes of dependent clause ratio and dependent clauses per T-unit of Level 2 text are bigger than that of Level 3 text. Therefore, the text in the Level 3 needs to be reviewed in order to expose more variation of sentence types.

The fourth category is coordination. The subtypes of Level 1, i.e. coordinate phrases per clause and coordinate phrases per T-unit, have less number compared with Level 2. This is already in line with the ideal concept. However, the number of the last subtype is the sentence coordination ratio of Level 1 is bigger than that of Level 2, where it is supposed to lower. Therefore, there is a need for Level 2 text to be reviewed, especially in this subtype. The last category is particular structure. Two of the three subtypes of the Level 1 text in this category are lower than that of the Level 2 text. This result is in line for the ideal concept. However, in the last subtype, i.e. the number of verb phrases per T-unit of Level 1 text is bigger than that of Level 2 where the higher the level, the bigger the number of verb phrases per T-unit. Therefore, there is a need for Level 2 to be reviewed in the future. The comparison between Level 2 and Level 3, the subtypes of Level 2, i.e. coordinate phrases per clause and coordinate phrases per T-unit, have bigger numbers compared with Level 3. This is not in line with the ideal concept. Therefore, the Level 3 text needs to be reviewed. However, the number of the last

subtype is the sentence coordination ratio of Level 2 is smaller than that of Level 3. This result is ideal.

The last category is particular structure. All three subtypes of the Level 2 text in this category are bigger than that of the Level 3 text. This result is not in line with the ideal concept where the higher the level, the bigger the number of particular sentence structure needs to be exposed to English learners. Therefore, there is a need for Level 3 to be reviewed in the future. After we see the comparison of lexical richness between Level 1, Level 2, and Level 3, let us now see the comparison of syntactic complexity between the three levels. The first category, the length of production unit between Level 1, Level 2 and Level 3, ideally there is a gradual increase as the level gets higher. However, from the data, it is shown that the Level 3 texts really need to be reviewed, especially in the sentence length (MLS), T-unit length (MLT), and clauses length (MLC) because the numbers of MLT and MLC need to be higher than that of Level 2 in order to increase the portion of the sentence length. The second category, the sentence complexity, the Level 3 needs to be reviewed because although it has more sentence complexity than the Level 2, the Level 2 has lower sentence complexity than the Level 1. These numbers show that there is no gradual increase in the sentence complexity where the ideal concept is as the level gets higher, students need to have higher sentence complexity.

Conclusion

From the data comparison, the numbers show that there is no gradual increase in the subordination, especially in all four subtypes. Therefore, the text in the Level 3 needs to be reviewed in order to expose more variation of sentence types. The fourth category is coordination. Similar to the previous category, the numbers show that there is no gradual increase in the coordination, especially in all three subtypes. Therefore, the text in the Level 3 needs to be reviewed in order to expose more variation of sentence coordination. The last category is particular structure. From the data comparison, the numbers show that there is also no gradual increase in this category, especially in all three subtypes. Therefore, the text in the Level 3 needs to be reviewed in order to expose more variation of sentence structure.

References

- Ai, H., & Lu, X. (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, the Netherlands: John Benjamins.
- Crystal, D. (2012). *English as a global language*. Cambridge, UK: Cambridge University Press.
- Cutler, A. (1983). Lexical complexity and sentence processing. In G. B. Flores d'Arcais, & R. J. Jarvella (Eds.), *The process of language understanding* (pp. 43-79). Chichester, Sussex: Wiley.
- Dörnyei, Z. (2007). *Research methods in applied linguistics: Quantitative, qualitative, and mixed methodologies*. Oxford: Oxford University Press.

- Dudley-Evans, T., & St John, M. J. (1998). *Developments in English for specific purposes: A multi-disciplinary approach*. Cambridge, UK: Cambridge university press.
- Jordan, R. R. (1997). *English for academic purposes: A guide and resource book for teachers*. Cambridge, UK: Cambridge University Press.
- Laufer, B., & Nation, P. (1995). Vocabulary size and use: Lexical richness in L2 written production. *Applied linguistics*, 16(3), 307-322.
- Lu, X. (2009). Automatic measurement of syntactic complexity in child language acquisition. *International Journal of Corpus Linguistics*, 14(1), 3-28.
- 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.
- Nation, I. S. (2001). *Learning vocabulary in another language*. Rowley, MA: Newbury House.
- Schmitt, N. (2010). *Researching vocabulary: A vocabulary research manual*. London: Palgrave Macmillan.