LEXICAL RICHNESS OF THE EXPOSITORY WRITING IN INDONESIAN BY SENIOR HIGH SCHOOL STUDENTS

Fara Wahyu Astridya

Department of Linguistics, Faculty of Humanity, Airlangga University

farawahyuastridya@yahoo.com

Abstract

Learning a foreign language is often related to knowing the words which called lexical richness. This study investigates the lexical richness in senior high school. The data was taken from the students' writing on expository essay. Writing an expository essay is a challenge for students where they should be able to write based on a given theme using a varied vocabulary. The students will be capable of making a good writing. The students are gathered from three grades that consist of 30 students in each grades. The total amount of the students will be measured by some types of lexical richness; lexical density, lexical sophistication, and lexical variation. After investigation the results showed a significant increase in each class, starting from grade 10 to grade 12. According to all those three measurements, grade 12 is the highest number among others and it concludes that students in grade 12 is the most prepared and have the richest lexical between grade 10 and 11.

Keywords: Language, Lexical Richness, English Language Teaching (ELT).

I   INTRODUCTION

Learning a foreign language is often related to knowing the words and the sentences. The knowledge of words is called lexical richness. Kyle & Crossley (2016) stated that lexical richness is the measurement of how rich the students' have in writing and composing the words or lexical in a good essay. In relation to the occurrences of lexical richness, Djiwandono (2016) asserted that these occurrences were majorly triggered by the second language use that can be further acknowledged from its sophistication and L2 learner's productive vocabulary. Moreover, the occurrences of lexical richness were majorly analysed in the field of applied linguistics. In this study, the wide range of tests were utilised to measure the lexical use in children and ESL learners.

Laufer & Nation (1995) developed a measurement which is specifically designed to evaluate the student's lexical proficiency level by conducting some comparison in terms of their lexical richness and external reference point. Among the measurement test that were discovered by many scholars, a type-token ratio (TTR) test gained its popularity due to its insight in evaluating the students' lexical language proficiency level. Koizumi & In'nami (2012) emphasises that the statistical analysis of TTR can be only further calculated based on the length of the texts that are used as the data. As stated previously, the utilization of lexical richness measurement strives towards the assessment of the students' lexical proficiency level by comparing their lexical richness with an external reference point.

It should be noted that in the analysis of lexical richness, the text needs to be transcribed and formatted in advance for easy processing of data (Daller, 2010). Gharibi & Boers (2017) said that by using lexical richness, researchers can identify the weaknesses and advantages of the object of the research. It can also make it easier to calculate the lexical property that is controlled by a person. The use of lexicon as a research object is based on the asumsion that every person would need good words to write a whole good sentence.. Word selection required vocabulary richness so that later there is no repetition of words in each sentence because it will affect the calculation of lexical richness (Caselli, Caselli, & Goldberg, 2016; Gharibi & Boers, 2017; Suggate & Stoeger, 2017).

The use of lexical richness is very useful in calculating or measuring in large numbers of data to simplify the collection of the data such as collecting some data from high school students as an ESL learner. Lexical richness analysis can only be done through the use of some measurement tools in order to analyse the raw materials or data. Laufer & Nation (1995) classify the lexical richness measurement tools into four distinctive types that were majorly utilised as the effective tools in measuring the description of the productive lexicon such as Lexical Originality (LO), Lexical Density (LD), Lexical Sophistication (LS), and Lexical Variation (LV). Lexical richness provides some options of techniques to researchers who want to collect the data in many ways, such as lexical originality, lexical density, lexical diversity and lexical sophistication. Those terms are also having each tool to determine the lexical richness on each student's writing knowledge.

Lexical richness measures can be used to compare students' writing. Some previous studies discussed about non-native speaker (NNS) and native-speaker (NS) as also mentioned by Lu & Ai

(2015). They compared among writers with diverse L1 backgrounds in college-level English writing. They found out that there are significant differences between EFL learners who were grouped and those who were not grouped by their L1 backgrounds with NNS and NS. Unlike previous studies that more often compare between NNS and NS, the current study more focused on NNS. Therefore, the researcher choosed Indonesian EFL senior high school students to perceive the lexical richness level based on their writing.

Basically, the measurement of lexical richness strives towards the calculation of how the spoken or written text contains various distinctive words that are used in text that is utilised as the data. Gregori-Signes & Clavel-Arroitia (2015) and Hanafiah & Yusuf (2016) further formulate four fundamental aspects in measuring the lexical richness such as Lexical Originality (LO) that concerns in calculating the words' proportion that is utilised by only one learner in a group, Lexical Density (LD) which examines the lexical words' proportion in the data, Lexical Variation (LV) which strives to calculate the occurrences of various distinctive lexical words that appeared in the data, and Lexical Sophistication (LS) that focuses the analysis in calculating the use of advanced words in the data.

In relation to the definition of Lexical Originality (LO), Laufer & Nation (1995) define LO as the measurement in lexical richness which focuses on the calculation between the learner's performance relative to the group and the written composition. In this case, there is a strong correlation between two variables namely the group and the index. If the number of the group gradually changes, it will automatically affect the index number respectively. This method provides information about the use of individual vocabulary with respect to fellow creatures. Such statistics is useful but it cannot stand alone as it varies across different compositions of an individual or it would change as the group changes. This can only be determined by adding the unique words number associated with one learner in a group that will be further divided by the calculation of the whole tokens. The umber of tokens is "the total number of word forms, which means any word occurring more than once in the text is calculated each time it is used".

In Indonesia, the teaching of English as a foreign language seems to be on the crossroads due to different curriculum implementation (Tantra, 2015). This means that every school in Indonesia has their own way to teach English to the students. One of the lessons that must be taught in every curriculum in Indonesia is writing. In high school, writing has been taught from grade 10 and continues until grade 12 so that the students are required to be able to write an essay in English. There are four main types of writing; expository, descriptive, persuasive, and narrative. The researcher thinks that expository can be a perfect type of writing due to the fact that expository writing is a kind of genre that trigger the learner to present and support a point of view with several reasons and evidences (Manik & Simurat, 2015).

In relation to the investigation of lexical richness, there are several numbers of researches that examine the occurrences of lexical richness in the ESL/EFL learners' composition. First is a thesis by Rebecca White from University of Wellington who conducted a study in examining the lexical richness of the adolescent writing. White (2014) focused her study on how the vocabulary knowledge of the adolescence develops through this period by further relating this factor to the subjects' background that were in New Zealand secondary schools. This study further utilised a mixed method. A quantative approach was applied to determine the vocabulary use in authentic essays written by the subjects that were further classfied into three different groups: 13-14, 15-16, and 17-18 that are ranged from eight different schools in New Zealand. The implementation of qualitative approach is used to identify the teacher's perspectives on the impact of the secondary school bakground of the students in the development of vocabulary. This study analysed the essays using three distinctive lexical richness measurements such as Lexical Variation (LO), Lexical Sophistication (LS), and Lexical Density (LD). The result of this study reveals that the time between years 11 and 13 (age 15-18) constitutes a period of significant lexical development in the areas of lexical variation, lexical sophistication, and lexical density. In contrast, the time between years 9 and 11 (age 13-16) only shows development in the area of use of lower-frequency words (beyond the first 3,000 words of English).

Second thesis by Pritomo (2012), he examined the lexical richness in teacher talk of a native and non-native English teacher in oral production in teaching English. The result revealed that 28% of the occurrences of lexical richness were derived from the teacher's oral production which further categorized as fair where the occurrences of lexical richness in the teacher's written production also yielded the total number of occurrences as same as the teacher's oral production. Another study was conducted by Lu (2012) who analysed the relationship of lexical richness to the quality of ESL learners'. He identifies the measurements which strongly correlate to the quality of ESL learners' oral narrative, and also to understand the correlation between these measurements. In his research, he provides ESL

teachers and researchers with a robust tool to assess the lexical richness of ESL language data samples which may be effectively used as indices of the quality of ESL learners' speaking performance.

Based on the explanation and some previous studies above, this study has a totally different subject. The previous studies concern more on the lexical richness and syntactic complexity on NS (native speaker). It is quite different with this current study because although the measuring instruments are the same, this study focuses more on EFL learners as non-native speakers in high school. Nevertheless, the researcher will develop about how the implementation on using lexical richness in applying on their writing. Therefore, the aim of this study is finding out the level of lexical richness in students' writing skill in making expository essay in Indonesian High School students.

In using lexical richness, each technique is useful in measuring the students' proficiency of the vocabulary in learning English as a Foreign Language (EFL). The aim of this study is focused on the Lexical Richness of the Expository Writing in Senior high school students and the comparison between those three levels in lexical richness.

Based on the previous studies, this present study has different tools in analysing the data. The previous studies are more concern about the students' proficiency level by identifying the writing texts. Most of them are not using spesific tools to analyze the data. Here in the current study, the researcher uses lexical richness as a measurement tool to identify the correlation between narrative and expository essays written by EFL students. The reason the researcher retrieves data by comparing the narrative and the expository since the narrative is considered more connected to everyday life, so that the students find it easier to tell the story by using common words (Cummins & Quiroa, 2012). Therefore, the aim of this study is finding out the lexical richness in students' writing skill of narrative and expository writing by Indonesian Senior High School students.

II   MATERIALS AND METHOD

Lexical richness has four different types of tool to measure the data; they are lexical originality, lexical diversity, lexical density and lexical sophistication. In this study, the researcher only use three types of lexical richness, exclude lexical originality as explained in introduction. Since there are so many steps to achieve the goals of this study, it needs an approach to guide how to reach the impeccable research. It is very compatible for this study, since it uses lexical richness techniques in data retrieval. Both techniques are using some tools to retrieve the data.

In this research, the respondents are high school students. The students are selected through a selection held annually by selecting a rating with a minimum score of 35.00. The high school students are proved capable in English as Foreign Language. The sample was based on three academic year level groups representing three different stages of senior school education, the beginning, middle years and final year of senior high school. The participants of the study are grade 10, 11 and 12 which in each class has different number of students. All grades are divided into 13 classes which consist of 7 science programs, 5 social programs and one language program. Each class has approximately 30 to 35 students with total amount of all classes is around 100 students. The selected class is a class that represents each class in the category that has high competency class among the others. In addition, in the selection of the data students are asked to write an expository essay with a predetermined theme. Each student will be assessed based on the results of their writing ability whether it meets the criteria of lexical richness. The assessment is based on how rich the vocabulary they use in the sentences.

In collecting the data, the researcher applied two steps. First, the researcher started to choose the class which has the highest rank among other classes. The last procedure, the researcher asked the students to write an expository essay with a specified theme in a spesific time. Furthermore, the time required to work on the essay is 90 minutes for 350 words according to what their teacher usually did before. Those terms apply for all the classes from grade 10 to 12.

There are two ways that are followed by the researcher in examining data of the present study. Firstly, the researcher analyzes the lexical richness which has three different ways of measuring the data; lexical density, lexical sophistication and lexical variation. Analyze the lexical density is using Lexical Complexity Analyser (LCA) that strives to calculate the ratio of content words to the total number of words in students' written texts. Next is lexical sophistication that measured using two ways of measurement: mid- and low-frequency words also academic words. Mid- and low- frequency words are measured by using Vocabprofiler while academic words is using Academic Word List (AWL) to analyze the data. The last is type-token ratio (TTR) which is used to measure lexical variation. Secondly, the researcher compares the lexical richness in grade 10, 11 and 12 based on those three types of

measurements. Lastly, the researcher interprets the results based on all the measurements of lexical richness.

III   RESULTS AND DISCUSSION

In this section, the results and discussion are presented. After identifying and selecting the data which found in students' expository writing done by 95 students, it found out that there are a significant increase based on the three measurements; lexical density, lexical sophistication, and lexical variation. Those three types of lexical richness have their own way to measure the sample of data as mention below.

The sample data obtained from students of each grade 10, 11, and 12. Below is a table with details of words and types of words of each student's essays.

Table 1. number of words from each grades

	Grades		
	X	XI	XII
Token	8.370	9.765	9.943
Type	7.450	7.102	7.036
TTR	0.89	0.72	0.70
Token per type	1.12	1.37	1.41
Lexical Density	0.89	0.93	1.13

As seen in table 1, it showed that among the three levels, the 12th grade write with the most number of words (tokens) which eventually led to the appearance of the most types among others. In determining the token, the researcher uses the measurement by submitting the full of students' essays. In contrast, type is the total number of different words that occur in a text and it shows in the 10th grade. The type token ratio (TTR) shows that 10th grade has the highest percentage among others. TTR is the ratio obtained by dividing the type (number of different words) that occurs in text or speech with the token (total word count). Token per type is calculated from the number of token and divided by the number of type. High TTR shows high lexical variations while low TTR indicates otherwise. The result shows that 12th grade also has the highest number since the number of words in the essay is the highest of all. Means that grade 12 has the highest lexical variation among others.

TTR is also connected to lexical density which can be determined by how many words in the students' texts. After the findings, then the researcher separates the each of content words in each grade. Based on the table above, grade 12 indicates that has the highest percentage and it implies that 12th grade students write the essay in some variety of words. Selection of words used in text can also affect the percentage of lexical density. As mentioned before, lexical density is determined by the content words in the text. Various content words consist of verb, noun, adjective and adverb used in the text. As more variety of content words used, the more likely it will affect the quality of the texts. By seperating the each of content words, it can simplify the calculation of lexical density.

After finding the type token ratio of each essay from each grade, the researcher calculates based on each type of lexical richness. Each type uses a calculation application, Lexical Complexity Analyzer (LCA). By simply entering data from each essay of each student, later will appear the results of each type as listed below.

Table 2. Comparison of Each Grade

	X	XI	XII
Lexical Density	0.89	0.93	1.13
Lexical Sophistication	0.15	0.26	0.32
Lexical Variation	0.59	0.53	0.51

After calculating the data of each text above, it indicates that in grade 10 lexical density shows as the most frequent types compared to the others. Lexical density basically calculated from noun, verb, and adverb of each text and then grouped into one. Then, summed and calculated the average per sentences. The measurement focuses on the amount of content words such as nouns, lexical verbs, adjectives and non-grammaticalized adverbs which is normally used in a text, while the function words are included into the remaining words, which serve a grammatical purpose (Schmitt & Schmitt, 2013). It also counts the percentage of lexical word present in a text divided by the total number of words. Lexical words stands for noun, adjectives, verb and adverbs. Then, lexical density of a text can be calculated by expressing the number of content carrying words in a text/sentence as a proportion of all the words in the text/sentence (Bestgen, 2017). The short coming of this method is that it ignores syntactic structure of composition and other cohesive devices present in a given composition.

Next is lexical sophistication which measured by looking for what text that has the academic words and to determine low and high frequency, and it shows up that grade 10 has the lowest score. It is possible since grade 10 has less knowledge of writing an essay. Lexical sophistication is calculated by the proportion of low-frequency or advanced words in a text out of the total number of words (Milton, 2009). Lexical Sophistication, it is obtained by comparing the advanced level words used in the composition with the total number of the words used in it. This method has a drawback. It is based on the number of advanced words, which are relative to the exposure and learning. It means this is not a reliable measure of productive vocabulary.

The last is lexical variation where it can be found by finding out the type token ration in each essay. As explained before, type token ration is measured by dividing the number of words (token) by the number of types. Grade 12 shows the lowest score and it can be concluded that student in grade 12 rarely repeating words in a sentence. This technique is quite similar with lexical density since both of them are measuring about the content words in an essay, but lexical variation is focusing on the range of different words used across a text, or in other terms, the extent to which repetition is lacking (White, 2014). It is measured by type/token ratio. If a composition has greater number of different words, it would be a better composition. This method is only sensitive for different numbers of words but it does not point out the quality of different words used in composition.
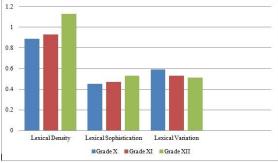

Figure 1. Comparison of Each Grade

From the chart above, it can be seen more clearly that the three types of lexical richness is a significant increase from grade 10 to 12. Lexical density is the highest percentage among other types. It could be due to the easy calculation. Basically to determine the lexical density of a text, we only need to sort out between verbs, nouns, adverbs, and adjectives. After grouping each type of content words, the author calculates using a tool called LCA. Without these tools, the data can actually be calculated manually by simply adding up each content words and then dividing it by the total number of words in a text. The disadvantage is to use the manual way later will take a lot of time. In lexical density, it can also be seen that from grade 10 to grade 12 there is a steady rise. It can be interpreted that students' writing ability can be measured from the grade level. The higher their grade, so does their ability to write.

Slightly different from lexical density, this type of lexical richness has a fairly low percentage level compared to other types. It concluded that there are few high school students in using academic words. Selection of words used by the students is still common words, so writing the sentence does not have a good meaning. In fact, it is expected to write an essay expository, students can learn to find and write with words that are not common so that it can provide lessons for students to later have a wealth of vocabulary. Yet, on the chart above shows the same thing with lexical density where there is a steady increase of classes 10 to 12.

In lexical variation, the measurement tool used is TTR where the TTR counts the number of words in a single text, and the number of words repeated several times in a sentence. The comparison is called lexical variation. TTR (type token ratio) is to calculate the number of tokens divided by type in one text. After that the result of this division, the number or percentage can be called as lexical variation. Based on the Figure 1, lexical variation includes high enough percentage. It can be interpreted that high school students make texts by using words that are quite a lot where from the beginning they were told to write 350 words within one hour. On average all students comply with the rules by writing according to the provisions. It can not be ascertained that the more tokens, the more types.

Different from the previous two types of lexical richness, the percentage of lexical variation decreases. The decrease here is good, as the percentage decreases, indicating that less repetition of the word is used in every sentence. Here it can be concluded that 10th graders most often repeat words in a single sentence. Students in grade 11 show different things where the students are not too many in using word repetition. Decreasing percentage also occurs in the 12th grade where the students in the class already know how to write a good essay. The selection of words they use also includes academic words. They also understand the rules to avoid repeating the word in every sentence.

IV CONCLUSION

Based on the results and discussion above, the lexical richness can be measured by three types; lexical density, lexical sophistication, and lexical variation. Each of three types have a measurements which can be used in measuring how rich the lexical on high school students' writing. Writing basically is one of the most arduous aspects on high school curriculum. In writing, students can explore their thoughts or ideas to express their feeling with their own words. Yet, in high school curriculum, writing can be so difficult for students since it becomes one of the requirements to pass their exam. In general, writing has four types; expository, descriptive, persuasive, and narrative. The reason why expository is chosen to be the sample of the data is because expository is the hardest of them all. In writing expository, students are asked to write based on the theme that chosen by the teacher, so the students somehow has limited words to write the essay. It challenges the students to explore the new vocabulary through their writing.

This is why the lexical richness has the important roles in students' writing. Using lexical richness to be the measurements is one of the simple way to find out how rich the lexical that students' have in their writing. As explained before in results and discussion, it showed that according to three types of lexical richness, all of them showed the significant increas from grade 10 to 12. It means that as the grade progresses, the students also show the progress on the vocabulary. Students in grade 10 still master a little vocabulary so the choice of words they use is still limited and often written repeatedly. Students in grade 11 have started to use vocabulary that is quite in accordance with the given theme. In the 12th grades, students more concern about the choice of words they use in their writing. They improve their writing skills by choosing some words worthy of use in their writing. By measuring the lexical richness, the aim of this study is the teacher and students will know how good and rich their writing based on the selected words they use.

The significance of the study is that the research may useful for the next students and also the teacher. For students, it can be useful for them in the future so that they can be more creative in putting their ideas into their writing. By knowing their level of lexical richness, they can use their error or lack of vocabulary especially to make their writing much more interesting. In addition, this study is expected to be useful for teachers who teach English in high school to improve students' writing skills by knowing their level of lexical richness. Moreover, this can also be expected to the next researcher to be their reference if they later want to examine the level of lexical richness in different subjects.

REFERENCES

Bestgen, Y. (2017). *Beyond single-word measures: L2 writing assessment, lexical richness and formulaic competence*. An International Journal of Educational Technology and Applied Linguistics.

Caselli, N., Caselli, M., & Goldberg, A. (2016). *Inflected words in production: Evidence for a morphologically rich lexicon*. Quartely Journal of Experimental Psychology, 69(3), 432–454.

Cummins, S., & Quiroa, R. . (2012). *Teaching writing expository responses to narrative texts*. The Reading Teacher, 66(6), 381–386.

Daller, J. T.-. (2010). *Operationalizing and measuring language dominance*. International Journal of Bilingualism, 15(2), 147–163.

Djiwandono, P. (2016). *Lexical Richness In Academic Papers: A Comparison Between Students' And Lecturers' Essays*. Journal of Applied Linguistics, 5(2).

Gharibi, K., & Boers, F. (2017). *Influential factors in lexical richness of young heritage speakers' family language: Iranians in New Zealand*. International Journal of Bilingualism.

Gregori-Signes, C., & Clavel-Arroitia, B. (2015). *Analysing Lexical Density and Lexical Diversity in University Students' Written Discourse*. 7th International Conference on Corpus Linguistics, 198(Cilc), 546–556. https://doi.org/10.1016/j.sbspro.2015.07.477

Hanafiah, R., & Yusuf, M. (2016). *Lexical Density And Grammatical Intricacy In Linguistic Thesis Abstract: A Qualitative Content Analysis*. Consortium of Asia Pacific.

Koizumi, R., & In'nami, Y. (2012). *Effects of text length on lexical diversity measures: Using short texts with less than 200 tokens*. System, 40(4), 522–532. https://doi.org/10.1016/j.system.2012.10.017

Kyle, K., & Crossley, S. (2016). *The relationship between lexical sophistication and independent and source-based writing*. Journal of Second Language Writing, 34, 12–24. https://doi.org/10.1016/j.jslw.2016.10.003

Laufer, B., & Nation, P. (1995). *Vocabulary Size and Use: Lexical Richness in L2 Written Production*. Oxford University Press.

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

Lu, X., & Ai, H. (2015). *Syntactic complexity in college-level English writing: Differences among writers with diverse L1 backgrounds*. Journal of Second Language Writing, 29, 16–27. https://doi.org/10.1016/j.jslw.2015.06.003

Manik, S., & Simurat, J. D. (2015). *Improve Students' Narrative Writing Achievement Through Film at SMA NEGERI I PALIPI*. International Journal of English Linguistics, 5(2).

Pritomo. (2012). *Lexical Richness in Teacher Talk of a Non Native English Teacher in the Foreign Language Classroom*. Universitas Negeri Malang.

Schmitt, N., & Schmitt, D. (2013). *A reassessment of frequency and vocabulary size in L2 vocabulary teaching*. Language Teaching.

Suggate, S., & Stoeger, H. (2017). *Fine Motor Skills Enhance Lexical Processing of Embodied Vocabulary: A Test of the Nimble-Hands, Nimble-Minds Hypothesis*. Quartely Journal of Experimental Psychology, 70(1), 2169–2187.

Tantra, D. (2015). *Teaching English As A Foreign Language In Indonesia: A Literature Review*. Lingual: Journal of Language and Culture, 4(1), 1–5. Retrieved from https://ojs.unud.ac.id/index.php/languange/article/view/19259

White, R. (2014). *Lexical Richness In Adolescent Writing, Insights From The Classroom: An L1 Vocabulary Development Study*. Victoria University of Wellington.