Lexical Errors Produced by High And Low Proficiency Learners of Writing 2 Classes

Ellara Yusea Ananda[1], Henny Putri Saking Wijaya[2]
English Department, Faculty of Languages and Literature, Petra Christian University, Siwalankerto 121-131, Surabaya, 60236, INDONESIA
E-mail: ellara.ananda@gmail.com[1]

ABSTRACT

This study was conducted to find out the types of lexical errors that high and low proficiency learners produced in their writing, and find the differences and similarities from the production of the two groups. To answer the research questions, the writers used the theory by James (2013) about lexical error classification. This study was a qualitative study. The sources of the data of this study were the lexical errors from 32 writing drafts. The writers divided the students into two groups: eight students from high proficiency level and eight students from low proficiency level. The findings showed that the high proficiency learners produced five types of formal errors and four types of semantic errors, while the low proficiency learners produced two types of formal errors and six types of semantic errors. In conclusion, the high and the low proficiency learners' lexical error productions are due to the learners' lack of knowledge in sense relation and collocation, and the learners' wrong terms of near-synonyms production.

Keywords: lexical errors, high proficiency learners, low proficiency learners, error analysis

INTRODUCTION

Errors in various types that are produced by foreign language learners are considered as "the process of learning a language" (Hemchua & Schmitt, 2006, p.3). One of the most frequent errors made by the learners is lexical errors, which is an incorrect selection of words that are intended to say. This error may cause misunderstanding and misinterpretation of the message as a whole since the lexicon selection carries the proposed message, as stated by Hemchua and Schmitt (2006). In written text, lexical errors that occur may be more disruptive as the messages delivered are all written with no oral explanation, thus the interpretation depends heavily on the writing including the selection of words. Therefore, lexical errors are important to be acknowledged. As a part of learning, making errors and evaluating the errors can help the teachers as well as the learners to improve their language teaching and learning better.

In this study, the writers analyzed the types of lexical errors that were produced in the writing texts of the second-semester students of the English Department in Petra Christian University. *Writing 2* is one of the skill courses in the second semester and is the subsequent class to *Writing 1* where the learners first get the writing course in the English Department. The writers chose *Writing 2* course because the learners were expected to have learned about writing the descriptive and exemplification paragraphs with minimal errors in grammar, vocabulary, and mechanics (English Department, 2016); thus, they had enough knowledge and skill for their subsequent course. In *Writing 2* course, there were approximately 16 meetings and two exams (mid-term and final test). This course was divided into three classes, class A, B, and C, according to the overall quota of 62 students who took the course in the second term of 2020/2021 academic year.

The writers divided the proficiency level of the learners based on their final grade of *Writing 1*: the learners whose final grade was A are the high proficiency learners as it is the highest score in the Writing course; and those whose final grade was C+ are the learners with

low proficiency, as it is one of the lowest scores and the passing grades that the learners can achieve to pass into *Writing 2* class. The passing grade of *Writing 1* course is actually C; however, according to the data, there was no student who got final grade of C. Therefore, final grade of C+ was taken for this study instead.

The lexical errors were from the learners' first draft of the first assignment and the writings of the midterm test in *Writing 2* class of English Department, Petra Christian University. Furthermore, the writers only used the students' two drafts that had not been corrected and given feedback by the lecturers since the drafts still contain the originality in terms of writing skill and production of the students.

The study aims is to examine the types of lexical errors produced by the high and the low proficiency learners in their writings based on the theory proposed by James (2013), about the classification of lexical errors. The classification consists of two major types: formal errors, which consist of formal misselection (suffix type, prefix type, vowel-based type, and consonant-based type), misinformation (borrowing, coinage, calque), and distortions (omission, overinclusion, misselection, misordering, and blend); and semantic errors, which consist of confusion of sense relations (superonym for hyponym, hyponym for superonym, inappropriate co-hyponym, and wrong term of near-synonyms) and collocational errors (semantically determined word selection, statistically weighted preferences, and arbitrary combinations). Additionally, this study also aims to find out if there are differences or similarities in the types of lexical errors produced by the high and the low proficiency learners.

METHODS

This study used qualitative approach with a validity of quasi-statistic (Maxwell, 2005). The participants were 16 second-semester learners from *Writing 2* classes of the English Department, Petra Christian University. The learners were divided into two groups according to the proficiency level based on their final grade of *Writing 1*: the learners whose final grade was A are the high proficiency learners as it is the highest score in the *Writing* course; and those whose final grade was C+ are the learners with low proficiency, as it is one of the lowest scores grades that the learners can achieve to pass into *Writing 2* class. The passing grade of *Writing 1* course was actually C; however, according to the data, there was no student who got final grade of C. Therefore, final grade of C+ was taken for this study instead.

The source of the data was the errors found in the high and low proficiency learners' writings. The data were the lexical errors produced by the learners in their drafts of the writing assignments. The students' writings were their two writing drafts that had not been corrected and given feedback by the lecturers since the drafts still contain the originality in terms of writing skill and production of the students.

After the data were collected, the writers began to analyze the students' writings and highlighted the words or phrases in each sentence that were identified as lexical errors. The writers put different highlight colors to differentiate the learners group: purple for high proficiency learners and green for low proficiency learners. Each error found in the learners' writing was counted even though the same word appeared more than once in one passage. Then, the errors produced by the students were written down in a table, and both levels were divided into two different tables with the same format. When the tables were completed, the writers evaluated the differences and similarities of the errors that high and low proficiency learners produced. The results were presented in the next section about findings and discussions.

FINDINGS AND DISCUSSIONS

The Types of Lexical Errors in the High Proficiency Learners' Production

Out of 19 types of lexical errors proposed by James (2013), the learners produced nine types of lexical errors: semantically determined word selection, wrong term of near-synonym, suffix, calque, statistically weighted preference, omission, overinclusion, consonant-based, and arbitrary combination. From all errors, the most frequent types to occur is semantically determined word selection and wrong term of near-synonym with six occurrences. Meanwhile, the error types that were not found were prefix, vowel-based, borrowing, coinage, misselection, misordering, blend, hyponym for superonym, superonym for hyponym, and inappropriate co-hyponym.

Below, the writers review some of the examples of the highest three of lexical error types that are produced by the high proficiency students.

- **Semantically Determined Word Selection Type**

Semantically determined word selection was produced when the learners chose the words that are not semantically relevant to the context of the sentence.

"The book has a light blue paperback cover ordained with floral patterns." (1.1.1.1.)

The learner chose the wrong word which is not even related to the context in the sentence. The highlighted word here means 'to invest or grant ministerial or priestly authority in someone'. Therefore, it should be 'decorated', or the learner could just omit the word 'ordained' from the sentence.

- **Wrong Term of Near-Synonym Type**

Wrong term of near-synonym occurred when the students chose the synonyms of the words they intended to say that were actually inappropriate or not suitable for the context of their sentence.

"During night time, the light on my desk that can change into different hues from warm to cool light keeps me companied while I study." (1.5.2.2.)

'Keeps me companied' actually means being there to keep someone from feeling lonely; meanwhile, for this context the learner meant to say that her light accompanied her during her studying. Thus, it was supposed to be 'keep accompanying me' as it was included as the near-synonym of 'companied'.

- **Suffix Type**

Suffix type occurred when the learners wrote the wrong suffix in the words they intended to say.

"These roughen soles assisted me when I enacted coupés, assemblies, and glissades." (1.2.1.3.)

It should be 'rough' since 'roughen' is the verb form. It is considered a suffix error because the learner chose the verb form which has a different suffix with the correct term that must be in adjective form.

The Types of Lexical Errors in the Low Proficiency Learners' Production

Out of 19 types of lexical errors proposed by James (2013), the learners produced eight types of lexical errors: wrong term of near-synonym, calque, statistically weighted preference, inappropriate co-hyponym, suffix, hyponym for superonym, semantically determined word selection, and arbitrary combination. From all errors, wrong term of near-synonym was the most frequent type to occur with 23 occurrences. Meanwhile, the error types that were not found were prefix, vowel-based, consonant-based, borrowing, coinage, omission, overinclusion, misselection, misordering, blend, and superonym for hyponym.

The writers review some of the examples of three highest lexical error types produced by the low proficiency students below.

- **Wrong Term of Near-Synonym Type**

Wrong term of near-synonym occurred when the students chose the synonyms of the words they intended to say that are actually inappropriate or not suitable for the context of their sentence

"So, I was very happy to live with grandfather." (2.9.1.2.)

It should be a conjunctive adverb, such as 'therefore/thus/hence', because 'so' is a conjunction and cannot be used in the beginning of a sentence. Although they basically have the same meaning, their functions in sentence structure are different.

- **Calque Type**

Calque type was produced when the learners literally translated the word from their first language into the target language (TL) which caused error as the translated word was not suitable or inappropriate for the context . In this study, the learners' first language is Bahasa Indonesia, while the target language is English.

"Even though my Aunt's house is not a big house me and my family always find a way to have fun right there like playing with my younger cousin they were about 3 and 8 years old children, and we also spend time by watching a movies together." (2.5.1.2.)

The learner literally translated '*saya dan keluarga saya*' into 'me and my family' which is structurally misordered in the target language. The correct phrase is 'my family and I' as it should be in the form of subjective pronouns.

- **Statistically Weighted Preference Type**

Statistically weighted preference was produced when the learners chose words or phrases that were structurally and/or semantically correct in TL; however, there were other referred words or phrases that were more preferred and more common to say.

"The iPad color is silver that makes it looks more manly. " (2.7.1.1.)

'More manly' is not entirely wrong in terms of the structure. However, the comparative form of manly is 'manlier', so it is more preferred to use.

Comparison of High and Low Proficiency Learners' Lexical Error Production

The overall lexical errors produced by the learners are 66 errors, which consisted of 17 (25.76%) formal errors and 49 (74.24%) semantic errors. The high proficiency learners produced 11 (37.93%) formal errors and 18 (62.07%) semantic errors, whereas the low proficiency learners produced six (15.79%) formal errors and 32 (84.21%) semantic errors. Furthermore, based on variation of the types, the overall occurrence in both learner groups was 11 types, which consisted of five formal error types and six semantic error types. The high proficiency learners produced five types of formal errors and four types of semantic errors, while the low proficiency learners produced two types of formal errors and six types of semantic errors.

The differences and similarities of lexical error production by high and low proficiency learners are summarized in Table 1 below.

Table 1: Comparison of lexical error type occurrences in high and low proficiency learners' production

Types of Lexical Errors	Occurrence in the Learners' Production	
	High Proficiency Learners	Low Proficiency Learners
Formal Errors		
Formal misselection		
Suffix	13.79%	2.63%
Prefix	-	-
Vowel-based	-	-
Consonant-based	3.45%	-
Misinformation		
Borrowing	-	-
Coinage	-	-
Calque	10.35%	13.16%
Distortions		
Omission	6.90%	-
Overinclusion	3.45%	-
Misselection	-	-
Misordering	-	-
Blend	-	-
Semantic Errors		
Confusion of sense-relations		
Superonym for hyponym	-	-
Hyponym for superonym	-	2.63%
Inappropriate co-hyponym	-	5.26%
Wrong term of near-synonym	20.69%	63.16%
Collocational errors		
Semantically determined word selection	20.69%	2.63%
Statistically weighted preference	13.79%	7.90%
Arbitrary combination	3.45%	2.63%

Based on the result above, the high proficiency learners produced more varied formal error types than semantic error types. This is possibly because the learners had more problems distinguishing the correct and appropriate form of the words they intended. This finding is consistent with the studies by Hemchua and Schmitt (2006), Andre and Jurianto (2015), and Saud (2018), in which the formal errors were more varied found in the students' writings. Meanwhile, the low proficiency learners produced more varied semantic error types. This might be due to them having more problems in expressing the words they intended as they still lack the knowledge of sense relation, collocation, connotation, and register. This finding turns out to be contrast with the findings in the related studies by Hemchua and Schmitt (2006), Andre and Juranto (2015), and Saud (2018)

Furthermore, most of the lexical errors produced by high and low proficiency learners tended to be influenced by intralingual error. James (2013) states that distortions, confusion of sense-relation, and collocational errors are caused by intralingual transfer and the error occurrences in this study are mostly from those three types. Thus, intralingual error is the common cause of the lexical errors in this study. Both high and low proficiency learners seem to have problems in writing the correct word form by deleting (omission) and adding (overinclusion) some of the lexical items. Moreover, they also have some difficulties distinguishing the words they intended to deliver with other words that are: 1) close or similar in meaning but different in certain usage (wrong term of near-synonym); 2) from the same group of hyponyms (inappropriate co-hyponym); or 3) too specific (hyponym for superonym). Last but not least, the learners also found it hard acknowledging: 1) the words with the exact or appropriate meaning they intended to write (semantically determined word selection); 2) the more preferred terms in general (statistically weighted preference); and 3) the correct combination of the words with arbitrary combinations. This result is also found in Andre and Jarianto's study (2015).

The next point that the writers wanted to highlight is the disruption scale between formal and semantic errors in the learners' production. According to the analysis, semantic errors presumably to be more disruptive than the formal errors since the semantic aspects carry the meaning or the context of what were actually being said by the learners in their written text. Meanwhile, the formal errors were more likely understandable as the errors still have close-like forms with the correct word forms. This is based on James' (2013) measurement of deviance, which states that to fulfill the acceptability the piece of a language has to refer to the right context although it is already correct in terms of grammaticality.

Another thing that the writers found important to highlight in this study is the usage of conjunctions 'and', 'but', and 'so' in the beginning of sentences that mostly occurred in the learners of both levels' lexical error production. This might be due to the influence of the learners' verbal communication on a daily basis. Usually, in verbal daily conversations that are casual occasions, the three conjunctions are often spoken in the beginning of sentences. This habit then influences the learners to also use the conjunctions in their writing tasks at class which are actually academic papers and thus have to be in formal style.

CONCLUSION

Based on the findings, the high proficiency learners produced five types of formal errors and four types of semantic errors, which consist of semantically determined word selection, suffix, wrong term of near-synonym, calque, statistically weighted preference, omission, overinclusion, consonant-based, and arbitrary combination. From all errors, the most frequent type to occur in the high proficiency learners' production were wrong term of near-synonym and semantically determined word selection with six occurrences of each (20.69%). Meanwhile,

the low proficiency learners produced two types of formal errors and six types of semantic errors, such as wrong term of near-synonym, calque, statistically weighted preference, inappropriate co-hyponym, suffix, hyponym for superonym, semantically determined word selection, and arbitrary combination. From all errors, wrong term of near-synonym was the most frequent type to occur in the low proficiency learners' production with 24 occurrences (63.16%). The differences between high and low proficiency learners' lexical error production are found in: 1) the occurrence of two major types of lexical errors; 2) the most and the least frequent lexical error subtypes occurred in the two groups' production; 3) the occurrences of the group cases as well as the percentage of each lexical error type; and 4) the types that only occurred in one group of learners. On the other hand, the similarities between high and low proficiency learners' lexical error production are found in: 1) the same frequencies of several lexical error types that are equally produced in both groups; and 2) the lexical error types that are not produced by the learners of both groups.

To sum up, the learners presumably lack the knowledge in sense-relation and collocation. Based on the percentage, both high and low proficiency learners produced most errors on semantic errors with 74.24% of overall total. This is consistent with Hemchua and Schmitt's (2006) study, in which the semantic errors were the most produced overall. Semantic knowledge was indicated as more complex to master according to Hemchua and Schmitt (2006). Moreover, in terms of disruption, semantic errors presumably to be more disruptive than the formal errors since the semantic aspects carry the meaning or the context of what were actually being said by the learners in their written text. Meanwhile, the formal errors were more likely understandable as the errors still have close-like forms with the correct word forms. As stated by James (2013), to fulfill the acceptability, the piece of a language has to refer to the right context although it is already correct grammatically. Lastly, the learners also presumably had problems in distinguishing the right terms from a set of near-synonyms in their writings, as the wrong term of near-synonyms were found to be the most produced error type in the overall of the learners' writings.

REFERENCES

Andre, R. & Jurianto. (2015). An analysis of lexical errors in the English narrative writing produced by the tenth grade students of SMA Negeri 9 Surabaya in EFL classroom. *Anglicist*, *4*(2), 69-76.

English Department. (2016). *English Department catalog 2017-2018* [Exhibition catalogue]. Petra Christian University.

Hemchua S. & Schmitt, N. (2006). An analysis of lexical errors in the English compositions of Thai learners. *Prospect*, (*21*)3, 2-25. https://www.researchonline.mq.edu.au/

James, C. (2013). *Errors in language learning and use: Exploring error analysis*. Routledge.

Maxwell, J. A. (2005). *Qualitative research design: An interactive approach* (2nd ed.). Sage Publications.

Saud, W. I. (2018). Lexical errors of third year undergraduate students. *English Langugae Teaching, 11*(11), 161-168. http://doi.org/10.5539/elt.v11n11p161